

Area laws for classical entropies in a spin-1 Bose-Einstein condensate

Yannick Deller,¹ Martin Gärttner,² Tobias Haas,^{3,*} Markus K. Oberthaler,¹ Moritz Reh,^{1,2} and Helmut Strobel¹

¹*Kirchhoff-Institut für Physik, Universität Heidelberg,
Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*

²*Institut für Festkörpertheorie und Optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany*

³*Centre for Quantum Information and Communication, École polytechnique de Bruxelles,
CP 165, Université libre de Bruxelles, 1050 Brussels, Belgium*

We investigate the information extractable from measurement distributions of two non-commuting spin observables in a multi-well spin-1 Bose-Einstein condensate. We provide a variety of analytic and numerical evidence that suitably chosen classical entropies and classical mutual informations thereof contain the typical feature of quantum entropies known in quantum field theories, that is, the area law, even in the non-Gaussian regime and for a non-zero temperature. Towards a feasible experimental implementation, we estimate entropic quantities from a finite number of samples without any additional assumptions on the underlying quantum state using k -nearest neighbor estimators.

I. INTRODUCTION

The scaling of entropic measures associated with the quantum state of a spatial subregion is one of the central characteristics describing how quantum information and entanglement are encoded in spacetime. Originally discovered by Bekenstein in the context of black hole physics [1, 2] (see also [3–6]), the entropy of a subregion is, to leading order, proportional to the area of its enclosing surface – rather than the subregion’s volume – which is conveniently referred to as the *area law*. Over the last two decades, much theoretical evidence has been gathered for its appearance in various contexts, including, for instance, quantum field theory [7–12], quantum many-body systems [13–15], tensor networks [16] and thermalization [17]. Also, it has been shown to occur for measures of correlations between the subregion of interest and its complement, e.g. the quantum mutual information [18].

Considerably less evidence is available on the experimental side, which can be traced back to the notoriously difficult task of extracting the local quantum state, i.e. performing full quantum state tomography. So far, the area law has been observed only for a handful of degrees of freedom for which entropic measures can be reduced to direct observables. This includes, for example, a study of a few-site Bose-Hubbard system where the Rényi-2 quantum entropy is read out via two-copy interference [19, 20] and a trapped ion quantum simulator of ten qubits [21], in which the same quantity was extracted using random measurements [22].

The task of accessing the quantum state is even more challenging for continuous quantum many-body systems, which are in principle described by quantum fields over continuous positions as well as infinite-dimensional local Hilbert spaces. A recent ultracold atom experiment reported an area law for the quantum mutual information under the assumption that the underlying quantum state

is of Gaussian form, in which case measurements of two-point correlation functions suffice to calculate entropic quantities [23]. Nevertheless, studies beyond the Gaussian case have remained elusive so far.

What all aforementioned approaches, theoretical as well as experimental, have in common, is their reliance on a *quantum* entropy as an indicator for the area law-like behavior of quantum correlations. However, it has been shown recently that the appearance of the area law is by no means restricted to such quantum entropies: when considering phase-space representations and measurement distributions of the quantum state instead, their corresponding *classical* entropies reveal the area law in the *next-to-leading* order terms, i.e. when classical contributions are subtracted properly [24].

Prominent examples of such distributions are the Wigner W -distribution [25], its marginals and the Husimi Q -distribution [26, 27] (see [28–30] for reviews on phase-space methods). Especially the latter is of particular interest as it is a non-negative and normalized function in phase space, which allows for well-defined entropic descriptions [31–36]. Further, its usefulness for witnessing entanglement has already been demonstrated theoretically in terms of entropic measures [37–39] as well as experimentally [40, 41].

Although all phase-space distributions contain the very same information as the density operator, estimating their associated classical differential entropies is a significantly simpler task than reconstructing the full many-body quantum state and computing its quantum entropy. The outcome of an experiment can be thought of as a sample drawn from a given distribution (usually the marginals of the Wigner W - or the full Husimi Q -distribution are considered), see e.g. [40–48]. In principle, one could then attempt to infer the corresponding distribution from these samples, which is equivalent to state tomography when considering a full phase-space distribution, and calculate its entropy thereafter. However, the need for reconstructing the underlying distribution can be bypassed by employing sophisticated methods to estimate the entropy from the sampled data *directly*. In this context, a widely-used approach is the k -nearest neighbor (k NN) method,

* tobias.haas@ulb.be

which produces an asymptotically unbiased estimate for the entropy without any assumptions on the underlying distribution [49–63].

In this work, we consider an experimentally friendly scenario in which the build-up of an area law over time is expected: the post-quench dynamics of a multi-well spin-1 Bose-Einstein condensate (BEC). In what follows, we simulate the dynamics using the truncated Wigner approximation (TWA) and employ the k NN method to estimate subtracted classical entropies as well as classical mutual informations of phase-space distributions from sampled data. We provide detailed numerical evidence for the appearance of the area law for classical entropies in a variety of scenarios, including, for instance, non-Gaussian states and varying system sizes, and under typical experimental constraints, e.g., thermal fluctuations and finite sample size. Our simulation results are supported by an analytical model covering the early-time dynamics, which allows for a straightforward evaluation of all entropic quantities. With our findings, we pave the ground for an experimental observation of the area law in a continuous quantum many-body system without any prior assumptions on the underlying quantum state.

The present work describes a detailed description of how the area law can be accessed from classical entropies, both from a theoretical and a practical perspective. In the accompanying manuscript [64], we provide a concise summary of our findings and additionally discuss the process of local thermalization in terms of such entropies.

The remainder of this paper is organized as follows. In section II, we introduce the model system of our interest – a spin-1 Bose-Einstein condensate in the multi-well setup – with a special focus on the Hamiltonian and the spin observables. We proceed with mapping this system to a continuous-variable quantum system described by canonical commutation relations using the so-called undepleted pump approximation in section III. Therein, we also put forward phase-space representations of quantum states and show how these are related to the measurement distributions of the spin observables. The two approaches for simulating the dynamics, i.e. the truncated Wigner approximation for the full Hamiltonian and an approximate, analytically solvable model valid for early times are discussed in section IV. For the latter, we employ methods from Gaussian quantum information theory, which we describe in detail. Thereupon, in section V, we introduce the necessary background on the area law of quantum and classical entropies in phase space. Then, we describe the k NN machinery for estimating entropies from sampled data and benchmark the method for several cases relevant for its application to the system of our interest in section VI. Finally, we provide our main results in section VII. After a comparison of the analytical model with the TWA approach (section VII A) we show plenty of evidence for the emergence of the area law over time for a variety of classical entropies (section VII B), which is followed by systematic studies of its robustness. More precisely, we consider strongly non-Gaussian distributions

(section VII C), a thermal initial state (section VII D), boundary effects (section VII E), varying system size (section VII F), varying sample size (section VII G) and different types of boundary conditions (section VII H). A comprehensive discussion of our results and some future prospects are given in section VIII.

Notation. We employ natural units $\hbar = k_B = 1$, denote quantum operators by bold letters, e.g. $\boldsymbol{\phi}$, and classical variables by normal letters, e.g. ϕ , (similarly for operations on operators / matrices, e.g. $\text{Tr}\{\boldsymbol{\rho}\} / \det\{\boldsymbol{\gamma}\}$) and equip vacuum expressions with a bar, e.g. \bar{Q} . Further, we use upper indices to refer to a specific well, e.g. Q^j for the j -th well, or to a subsystem of a bipartition AB , e.g. Q^A , and put lower indices for the hyperfine levels, e.g. a_0 , or relative modes, e.g. a_{\pm} .

II. MULTI-WELL SPIN-1 BEC

We begin with a discussion of the multi-well setup of coupled spin-1 Bose-Einstein condensates. Then, we introduce the Hamiltonian governing the corresponding dynamics and the typical observables in such systems.

A. Setup

We consider an experimental setup consisting of a spin-1 BEC with ferromagnetic spin coupling, based on Lithium-7 [65, 66]. We assume the condensate’s geometry to be quasi one-dimensional, realizable via tight confinement along the radial direction, which can be implemented experimentally by ensuring that the radial trapping frequency is much larger than the longitudinal one, i.e. $\omega_r \gg \omega_l$. In the longitudinal direction, the BEC is further subjected to an optical lattice potential, which divides the overall BEC into N smaller-size BECs, which we refer to as wells. The depth of the optical lattice controls the potential barrier between neighboring wells. When tuned to sufficiently small energies, atoms can tunnel between neighboring wells, which establishes a nearest-neighbor interaction between the local degrees of freedom. This setup constitutes a discretized approximation to a continuous quantum field, see section IV B 4.

The internal degrees of freedom encoding the spin-1 system are the three Zeeman levels $m_F = 0, \pm 1$ of the $F = 1$ hyperfine manifold of the electronic ground state. Initially, we consider all atoms to be prepared in the $m_F = 0$ mode. Thereafter, spin-changing collisions triggered by off-resonant microwave dressing lead to the side modes being dynamically populated [65, 67, 68]. As this process is symmetric with respect to the side modes, the linear Zeeman shift cancels for all populated states and hence only the quadratic Zeeman shift is relevant for the dynamics.

We consider tunneling to only take place in the side modes and not in the initially macroscopically populated zero mode, such that information exchange between wells

through tunneling only takes place after the generation of side-mode population. Experimentally, this can be achieved, for example, by working in a state-dependent lattice, such that the side modes are trapped more weakly than the zero mode. Another possibility is to tune the spin-changing collisions into resonance with excited states of the side modes and work in a deep optical lattice, such that atoms transferred to the side modes acquire enough kinetic energy to tunnel to neighboring sites of the deep optical lattice, whereas the kinetic energy of atoms in the zero mode remains insufficient for tunneling.

B. Modes and Hilbert space

We are interested in the dynamics of the internal degrees of freedom characterized by the three hyperfine levels $m_F \in \{-1, 0, 1\}$. To each hyperfine level m_F and well j we associate a set of bosonic creation and annihilation operators

$$[a_{m_F}^j, a_{m_F'}^{j'\dagger}] = \delta^{jj'} \delta_{m_F m_F'}, \quad (1)$$

where the upper index labels the $N \in \mathbb{N} < \infty$ wells such that $j \in \{1, \dots, N\}$. For a single well j the underlying Hilbert space \mathcal{H}^j is a Fock space constructed from multi-particle states

$$|n_1^j, n_0^j, n_{-1}^j\rangle \equiv \frac{(a_1^{j\dagger})^{n_1^j} (a_0^{j\dagger})^{n_0^j} (a_{-1}^{j\dagger})^{n_{-1}^j}}{\sqrt{n_1^j! n_0^j! n_{-1}^j!}} |0, 0, 0\rangle, \quad (2)$$

with the tensor product notation $|\cdot, \cdot, \cdot\rangle = |\cdot\rangle \otimes |\cdot\rangle \otimes |\cdot\rangle$ understood and the number of particles $n_{m_F}^j$ in mode j and level m_F being defined as the eigenvalue of the corresponding particle number operator

$$N_{m_F}^j = a_{m_F}^{j\dagger} a_{m_F}^j. \quad (3)$$

Accordingly, the total number of particles in well j is measured by the operator

$$N^j = \sum_{m_F=-1}^1 N_{m_F}^j, \quad (4)$$

which has the eigenvalue $n^j = \sum_{m_F=-1}^1 n_{m_F}^j$. Then, the full Hilbert space is obtained by taking the tensor product with respect to all wells, i.e. $\mathcal{H} = \otimes_{j=1}^N \mathcal{H}^j$, such that a generic element of the full Fock basis reads

$$|n_1^1, n_0^1, n_{-1}^1; \dots; n_1^N, n_0^N, n_{-1}^N\rangle \equiv \bigotimes_{j=1}^N |n_1^j, n_0^j, n_{-1}^j\rangle. \quad (5)$$

C. Hamiltonian

The full Hamiltonian is composed of two main terms: a single-well Hamiltonian H_{sw}^j describing the on-site dynamics of well j as well as a tunneling Hamiltonian H_t^j

encoding the coupling between neighboring wells j and $j+1$, such that in total we have

$$H = \sum_{j=1}^N H_{\text{sw}}^j + \sum_{j=1}^{N \text{ or } N-1} H_t^j. \quad (6)$$

For the sake of generality, we keep the type of boundary conditions open at this point. More precisely, we allow for periodic or open boundary conditions, which are implemented by the second sum in (6) running up to N or $N-1$, i.e. coupling the first and the last well or not, respectively.

The single-well Hamiltonian for well j is given by

$$\begin{aligned} H_{\text{sw}}^j &= c_0 N^j (N^j - 1) \\ &+ c_1 \left[\left(N_0^j - (1/2)1 \right) \left(N_1^j + N_{-1}^j \right) \right. \\ &\quad \left. + a_0^{j\dagger} a_0^j a_1^j a_{-1}^j + a_1^{j\dagger} a_{-1}^j a_0^j a_0^j \right] \\ &+ q \left(N_1^j + N_{-1}^j \right), \end{aligned} \quad (7)$$

and has been investigated in great detail, see e.g. [40, 41, 65, 67, 69, 70]. It contains three contributions: First, the on-site density-density interaction $c_0 > 0$ describing the repulsive interaction of atoms regardless of their hyperfine levels (petrol ellipse in Figure 1). Second, the spin-changing collision interaction $c_1 < 0$ encoding the generation of spin pairs in the $m_F = \pm 1$ hyperfine levels from the $m_F = 0$ level (red arrows), or the reverse process. For clarity, we omitted the mean-field shifts in the second line of Eq. (7). Third, we included the quadratic $q > 0$ Zeemann shift (green arrows).

The tunneling between equal hyperfine levels $m_F = \pm 1$ of neighboring wells j and $j+1$ is described by

$$H_t^j = -J \sum_{m_F=\pm 1} \left(a_{m_F}^{j\dagger} a_{m_F}^{j+1} + a_{m_F}^{(j+1)\dagger} a_{m_F}^j \right), \quad (8)$$

with a non-negative tunnel rate $J \geq 0$ (blue arrows). Note again that the $m_F = 0$ mode does not couple to neighboring wells and also that in the limit $J \rightarrow 0$ the wells evolve independently.

D. Observables: Spin operators

Among the local observables in every well are the eight spin operators that form a representation of the Lie algebra $\mathfrak{su}(3)$ [65, 67]. They are constructed following the Jordan-Schwinger map [71, 72]: starting from an irreducible representation of the eight three-dimensional matrices G_α of $\mathfrak{su}(3)$ defined via

$$[G_\alpha, G_\beta] = f_{\alpha\beta\gamma} G_\gamma, \quad (9)$$

where $f_{\alpha\beta\gamma}$ are the so-called structure constants (we omit their explicit form here) and $\alpha \in \{1, \dots, 8\}$ labels the

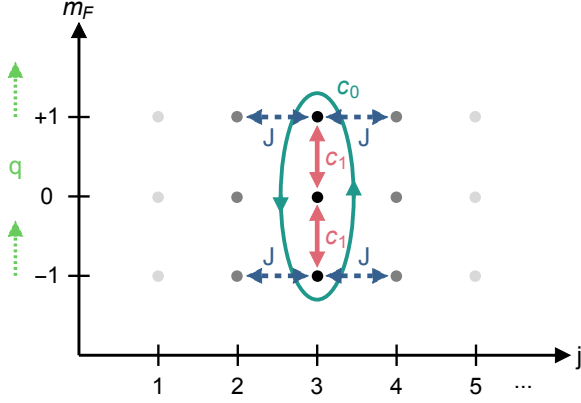


Figure 1. Illustration of the dynamics generated by the five contributions to the full Hamiltonian (6) for the $j = 3$ well. First-order (tunneling) and second-order (collisions) processes are indicated by dashed and solid arrows, respectively, while the energy shift due to the Zeemann effect is depicted by a dotted arrow.

matrices, the local quantum operators G_α^j obeying the very same algebra are constructed following

$$G_\alpha^j = \sum_{m_F, m_{F'}}^1 a_{m_F}^{j\dagger} (G_\alpha)_{m_F m_{F'}} a_{m_{F'}}^j, \quad (10)$$

with $(G_\alpha)_{m_F m_{F'}}$ denoting the $(m_F, m_{F'})$ -th. entry of the spin matrix G_α .

A common choice for the $\mathfrak{su}(3)$ matrices G_α is made by first constructing a $\mathfrak{su}(2)$ subspace (see [67] for explicit expressions for the corresponding matrices) via

$$[S_\alpha^j, S_\beta^j] = i\epsilon_{\alpha\beta\gamma} S_\gamma^j, \quad (11)$$

where $\alpha, \beta, \gamma \in \{x, y, z\}$, and thereupon deriving the three corresponding spin operators from (10), leading to

$$\begin{aligned} S_x^j &= \frac{1}{\sqrt{2}} \left[a_0^{j\dagger} (a_1^j + a_{-1}^j) + (a_1^{j\dagger} + a_{-1}^{j\dagger}) a_0^j \right], \\ S_y^j &= \frac{i}{\sqrt{2}} \left[a_0^{j\dagger} (a_1^j - a_{-1}^j) - (a_1^{j\dagger} - a_{-1}^{j\dagger}) a_0^j \right], \\ S_z^j &= a_1^{j\dagger} a_1^j - a_{-1}^{j\dagger} a_{-1}^j. \end{aligned} \quad (12)$$

The remaining five operators are the so-called quadrupole operators defined as

$$Q_{\alpha\beta}^j = \{S_\alpha^j, S_\beta^j\} - \frac{4}{3}\delta_{\alpha\beta}\mathbb{1}, \quad (13)$$

with $\{.,.\}$ denoting the anticommutator. Here Q_{yz}^j and Q_{xz}^j are of special interest, which read

$$\begin{aligned} Q_{yz}^j &= \frac{i}{\sqrt{2}} \left[a_0^{j\dagger} (a_1^j + a_{-1}^j) - (a_1^{j\dagger} + a_{-1}^{j\dagger}) a_0^j \right], \\ Q_{xz}^j &= \frac{1}{\sqrt{2}} \left[a_0^{j\dagger} (a_1^j - a_{-1}^j) + (a_1^{j\dagger} - a_{-1}^{j\dagger}) a_0^j \right]. \end{aligned} \quad (14)$$

E. Readout schemes

We will analyze the system by gathering information about the measurement distributions of the variable pairs (S_x^j, Q_{yz}^j) over multiple wells. We discuss methods for reading out two types of such distributions in the following.

1. Separate detection

In BEC experiments the population of each mode, $n_{m_F}^j = \langle N_{m_F}^j \rangle$ is detected through absorption imaging after Stern-Gerlach separation of the different m_F components. From this the z -component of the local spin can be extracted as $S_z^j = n_1^j - n_{-1}^j$. Other spin components can be measured by applying a radio-frequency (rf) magnetic field prior to the absorption imaging. For a frequency matching the linear Zeeman shift, resonant Rabi oscillations are driven, described by the Hamiltonian (in the rotating wave approximation) [69, 70]

$$H_{\text{rf}} = \sum_{j=1}^N \Omega_{\text{rf}} [\cos(\phi_{\text{rf}}) S_y^j - \sin(\phi_{\text{rf}}) S_x^j] \quad (15)$$

Here, Ω_{rf} is the Rabi frequency of the drive and ϕ_{rf} is a tunable phase. Applying the drive for a time $t = \pi/(2\Omega_{\text{rf}})$ thus allows one to map the spin along an arbitrary direction of the equatorial plane of the spin sphere onto S_z^j . In the case where the $m_F = 0$ mode is a coherent state with population far larger than the side modes, this scheme is analogous to homodyne detection in optics, where the signal mode(s) are mixed with a local oscillator mode on a beam splitter to extract the field quadratures [73].

To toggle between measurements of spin operators and quadrupole operators one can add a time delay before the application of the rf-rotation. By tuning the quadratic Zeeman shift and the microwave dressing field in the single-well Hamiltonian Eq. (7) the $m_F = 0$ mode acquires a relative phase. For instance, measuring Q_{yz}^j instead of S_x^j requires an additional phase of $\pi/2$.

2. Simultaneous detection

Sophisticated methods to extract joint distributions over (S_x^j, Q_{yz}^j) , or equivalently (S_y^j, Q_{xz}^j) , in close analogy to the heterodyne detection protocol in quantum optics [29, 30, 74, 75] have been experimentally realized rather recently [40, 41, 70]. For this, one exploits the availability of additional, initially unoccupied, internal levels, in our example the $F = 2$ hyperfine manifold with five additional Zeeman levels. Using microwave fields resonantly coupling the levels of the $F = 1$ manifold with those of the $F = 2$ manifold, one can realize analogs of beam splitter operations between them. Splitting each mode equally and subsequently applying different rf-rotations

in the two manifolds realizes the simultaneous detection of, for example, \mathbf{S}_x^j and \mathbf{Q}_{yz}^j (see Ref. [40] for details). In the case of the $m_F = 0$ mode being macroscopically occupied, this corresponds to sampling from the Husimi Q -distribution in the phase space spanned by \mathbf{S}_x^j and \mathbf{Q}_{yz}^j (see section III for details).

III. FROM A SPIN-1 BEC TO A CONTINUOUS-VARIABLE QUANTUM SYSTEM

We point out the connection between the system of our interest, a multi-well spin-1 BEC, and a continuous-variable quantum system describing coupled bosonic oscillator modes. Thereupon, we introduce several kinds of phase-space descriptions.

A. Undepleted pump approximation

The undepleted pump regime captures the early-time dynamics under the Hamiltonian (6) when initially preparing the polar ground state of the spin-1 BEC in every well, i.e.

$$|\psi(0)\rangle = |0, \alpha, 0\rangle^{\otimes N} \quad (16)$$

where $|\alpha\rangle$ is the local coherent state of the zero mode with mean particle number $n = |\alpha|^2$. For large values of n the $m_F = 0$ mode is macroscopically populated compared to the side modes $m_F = \pm 1$ in every well, i.e.

$$\langle \mathbf{N}_0^j \rangle \gg \langle \mathbf{N}_{\pm 1}^j \rangle, \quad (17)$$

for all j , which is equivalent to

$$n^j = \langle \mathbf{N}^j \rangle \cong \langle \mathbf{N}_0^j \rangle. \quad (18)$$

Under the above assumptions, we can approximate the $m_F = 0$ mode operators by their norms

$$\mathbf{a}_0^j = \mathbf{a}_0^{j\dagger} \cong \sqrt{n^j}, \quad (19)$$

such that $\mathbf{N}_0^j \cong n^j$. Then, the full Hamiltonian (6) simplifies to

$$\begin{aligned} H_{\text{up}} = & \sum_{j=1}^N \left[c_0 \left(\mathbf{a}_1^{j\dagger} \mathbf{a}_1^{j\dagger} \mathbf{a}_1^j \mathbf{a}_1^j + \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^j \mathbf{a}_{-1}^j + 2\mathbf{N}_1^j \mathbf{N}_{-1}^j \right) \right. \\ & + \tilde{c}_1^j \left(\mathbf{a}_1^j \mathbf{a}_{-1}^j + \mathbf{a}_1^{j\dagger} \mathbf{a}_{-1}^{j\dagger} \right) \\ & \left. + (\tilde{q}^j + c_0 n^j) \left(\mathbf{N}_1^j + \mathbf{N}_{-1}^j \right) \right] \\ & - J \sum_{j=1}^{N \text{ or } N-1} \left(\mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^{j+1} + \mathbf{a}_{-1}^{(j+1)\dagger} \mathbf{a}_{-1}^j \right. \\ & \quad \left. + \mathbf{a}_1^{j\dagger} \mathbf{a}_1^{j+1} + \mathbf{a}_1^{(j+1)\dagger} \mathbf{a}_1^j \right) \\ & + H_{\text{offset}}. \end{aligned} \quad (20)$$

Here, we defined the rescaled couplings

$$\tilde{c}_1^j = c_1 n^j, \quad \tilde{q}^j = c_1 \left(n^j - \frac{1}{2} \right) + q, \quad (21)$$

as well as a constant offset term

$$H_{\text{offset}} = c_0 \sum_{j=1}^N (n^j)^2, \quad (22)$$

which we will drop in the following.

B. Relative modes and canonically conjugate variables

In the undepleted pump regime, the relevant degrees of freedom and their Hamiltonian (20) can be mapped to a continuous-variable quantum system. To that end, we introduce the relative mode operators between the side modes

$$\mathbf{a}_{\pm}^j = \frac{1}{\sqrt{2}} \left(\mathbf{a}_1^j \pm \mathbf{a}_{-1}^j \right), \quad (23)$$

which also represent independent bosonic modes, since (1) implies

$$[\mathbf{a}_{\pm}^j, \mathbf{a}_{\pm}^{j'\dagger}] = \delta_{jj'}, \quad [\mathbf{a}_{\pm}^j, \mathbf{a}_{\mp}^{j'\dagger}] = 0. \quad (24)$$

Their associated canonical operators are defined as

$$\phi_{\pm}^j = \frac{1}{\sqrt{2}} \left(\mathbf{a}_{\pm}^{j\dagger} + \mathbf{a}_{\pm}^j \right), \quad \pi_{\pm}^j = \frac{i}{\sqrt{2}} \left(\mathbf{a}_{\pm}^{j\dagger} - \mathbf{a}_{\pm}^j \right), \quad (25)$$

which fulfill the canonical commutation relations

$$[\phi_{\pm}^j, \pi_{\pm}^{j'}] = i\delta^{jj'}, \quad [\phi_{\pm}^j, \pi_{\mp}^{j'}] = 0. \quad (26)$$

Interestingly, in the undepleted pump regime, the two pairs of canonical operators are equivalent to pairs of spin operators up to normalization, to wit

$$\begin{aligned} \mathbf{S}_x^j &= \sqrt{2n^j} \phi_+^j, & \mathbf{Q}_{yz}^j &= -\sqrt{2n^j} \pi_+^j, \\ \mathbf{S}_y^j &= \sqrt{2n^j} \phi_-^j, & \mathbf{Q}_{xz}^j &= -\sqrt{2n^j} \pi_-^j, \end{aligned} \quad (27)$$

which follows from using (19) in (12) and (14). This shows that, for short times, the relevant degrees of freedom mimic two pairs of canonical variables and hence the local Hilbert spaces \mathcal{H}^j decompose as $\mathcal{H}^j = \mathcal{H}_+^j \otimes \mathcal{H}_-^j$.

Next, we express the Hamiltonian (20) in terms of the newly defined operators (see Appendix A for details). We find the general decomposition

$$H_{\text{up}} = \mathbf{H}_{\text{up}}^+ + \mathbf{H}_{\text{up}}^- + \mathbf{H}_{\text{up}}^{\text{mix}}, \quad (28)$$

where the form-equivalent \mathbf{H}_{up}^+ and \mathbf{H}_{up}^- contain all terms with only + or - modes, respectively, while $\mathbf{H}_{\text{up}}^{\text{mix}}$ contains all terms mixing the two relative modes \pm . In terms of

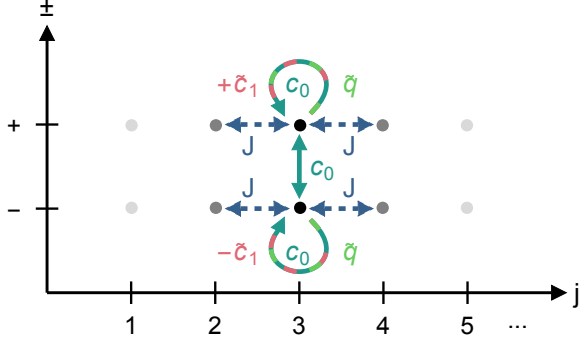


Figure 2. Sketch of the dynamics in the undepleted pump regime generated by (28) (in analogy to Figure 1).

relative mode operators, these two types of Hamiltonians read

$$\begin{aligned}
 H_{\text{up}}^{\pm} &= \sum_{j=1}^N \left[c_0 \mathbf{a}_{\pm}^{j\dagger} \mathbf{a}_{\pm}^j \mathbf{a}_{\pm}^{j\dagger} \mathbf{a}_{\pm}^j \right. \\
 &\quad \left. + [\tilde{q}^j + c_0 (n^j - 1)] N_{\pm}^j \right. \\
 &\quad \left. \pm \frac{\tilde{c}_1^j}{2} \left(\mathbf{a}_{\pm}^{j\dagger} \mathbf{a}_{\pm}^{j+1} + \mathbf{a}_{\pm}^j \mathbf{a}_{\pm}^{j+1\dagger} \right) \right] \\
 &\quad - J \sum_{j=1}^{N \text{ or } N-1} \left(\mathbf{a}_{\pm}^{j\dagger} \mathbf{a}_{\pm}^{j+1} + \mathbf{a}_{\pm}^{(j+1)\dagger} \mathbf{a}_{\pm}^j \right), \\
 H_{\text{up}}^{\text{mix}} &= 2c_0 \sum_{j=1}^N N_+^j N_-^j,
 \end{aligned} \tag{29}$$

whose dynamics are sketched in Figure 2. From the latter formulas, it becomes apparent that in the undepleted pump regime the relative mode operators, or equivalently the two pairs of canonical operators, constitute a complete set of observables for characterizing the early-time dynamics. Further, the dynamics within the \pm phase spaces differ only by a sign in the $\pm \tilde{c}_1$ term.

C. Canonical phase-space

As the observables of our interest now correspond to continuous variables, we can faithfully apply the powerful concepts of phase-space descriptions. The two sets of canonical operators $(\phi_{\pm}^j, \pi_{\pm}^j)$ span the two two-dimensional canonical phase spaces for well j , which are known to be isomorphic to the Euclidean plane \mathbb{R}^2 with measure [76, 77]

$$\int d\phi_{\pm}^j d\pi_{\pm}^j. \tag{30}$$

The field operators are conveniently combined into a single vector χ_{\pm}^j by defining $\chi_{\pm}^j = \phi_{\pm}^j$ for $j \in [1, N]$ and $\chi_{\pm}^j = \pi_{\pm}^{j-N}$ for $j \in [N+1, 2N]$, which we formally write

as [74, 75]

$$\chi_{\pm} = (\phi_{\pm}, \pi_{\pm})^T = (\phi_{\pm}^1, \dots, \phi_{\pm}^N, \pi_{\pm}^1, \dots, \pi_{\pm}^N)^T. \tag{31}$$

The canonical commutation relations (26) become

$$[\chi_{\pm}^j, \chi_{\pm}^{j'}] = i\Omega^{jj'} \mathbb{1}, \tag{32}$$

where

$$\Omega = (i\sigma_2) \otimes \mathbb{1}_N \tag{33}$$

denotes the symplectic metric revealing the symplectic structure of the canonical phase space and σ_2 is the second Pauli matrix.

For every well j we define the set of canonical coherent states associated with the relative modes \pm as displaced vacuum states [29, 30, 74, 75], i.e.

$$|\alpha_{\pm}^j\rangle = D(\alpha_{\pm}^j) |0_{\pm}^j\rangle, \tag{34}$$

where $D(\alpha_{\pm}^j)$ is the unitary displacement operator

$$D(\alpha_{\pm}^j) = e^{\alpha_{\pm}^j \mathbf{a}_{\pm}^{j\dagger} - \alpha_{\pm}^{*j} \mathbf{a}_{\pm}^j}, \tag{35}$$

with the complex-valued phase fields being parameterized in terms of cartesian coordinates as

$$\alpha_{\pm}^j = \frac{1}{\sqrt{2}} (\phi_{\pm}^j + i\pi_{\pm}^j). \tag{36}$$

Importantly, the set of coherent states constitutes an overcomplete basis, i.e. the coherent states defined in (34) resolve the identity in the \pm subspaces

$$\mathbb{1} = \int \frac{d\phi_{\pm}^j d\pi_{\pm}^j}{2\pi} |\alpha_{\pm}^j\rangle \langle \alpha_{\pm}^j|, \tag{37}$$

but are not orthogonal to each other.

D. Phase space distributions

The phase-space picture enables the description of the system's state ρ^j in terms of classical phase-space distributions. In the following, we will consider the distributions associated with either the $+$ or the $-$ mode, whose corresponding density operators are obtained via the partial trace $\rho_{\pm}^j = \text{Tr}_{\mp} \{\rho^j\}$.

1. Wigner W -distribution

The arguably most prominent phase-space representation is the Wigner W -distribution [25], which is defined as the Fourier transform of the characteristic function, namely [74]

$$\begin{aligned}
 \mathcal{W}_{\pm}^j &\equiv \mathcal{W}_{\pm}^j(\phi_{\pm}^j, \pi_{\pm}^j) \\
 &= \int \frac{d\tilde{\phi}_{\pm}^j d\tilde{\pi}_{\pm}^j}{2\pi} e^{-i(\phi_{\pm}^j, \pi_{\pm}^j) \Omega (\tilde{\phi}_{\pm}^j, \tilde{\pi}_{\pm}^j)^T} \\
 &\quad \times \text{Tr} \left\{ \rho_{\pm}^j e^{i(\phi_{\pm}^j, \pi_{\pm}^j) \Omega (\tilde{\phi}_{\pm}^j, \tilde{\pi}_{\pm}^j)^T} \right\}.
 \end{aligned} \tag{38}$$

Here, we have chosen the normalization such that \mathcal{W}_{\pm}^j is normalized to unity with respect to the phase-space integral measure (30), i.e.

$$1 = \text{Tr} \left\{ \rho_{\pm}^j \right\} = \int d\phi_{\pm}^j d\pi_{\pm}^j \mathcal{W}_{\pm}^j. \quad (39)$$

It is well-known that the Wigner W -distribution can become negative when the underlying state is non-classical [78]. However, we note that both our simulation approaches are naturally based on positive Wigner W -distributions (see section IV).

2. Marginal distributions

Next, we introduce the measurement distributions over the canonical operators

$$\begin{aligned} f_{\pm}^j &\equiv f_{\pm}^j(\phi_{\pm}^j) = \text{Tr} \left\{ \rho_{\pm}^j |\phi_{\pm}^j\rangle \langle \phi_{\pm}^j| \right\}, \\ g_{\pm}^j &\equiv g_{\pm}^j(\pi_{\pm}^j) = \text{Tr} \left\{ \rho_{\pm}^j |\pi_{\pm}^j\rangle \langle \pi_{\pm}^j| \right\}, \end{aligned} \quad (40)$$

with the eigenvalue equations

$$\phi_{\pm}^j |\phi_{\pm}^j\rangle = \phi_{\pm}^j |\phi_{\pm}^j\rangle, \quad \pi_{\pm}^j |\pi_{\pm}^j\rangle = \pi_{\pm}^j |\pi_{\pm}^j\rangle, \quad (41)$$

understood. The distributions (40) are true probability density functions with normalizations

$$1 = \text{Tr} \left\{ \rho_{\pm}^j \right\} = \int d\phi_{\pm}^j f_{\pm}^j = \int d\pi_{\pm}^j g_{\pm}^j, \quad (42)$$

and can be identified with the marginals of the Wigner W -distributions

$$f_{\pm}^j = \int d\pi_{\pm}^j \mathcal{W}_{\pm}^j, \quad g_{\pm}^j = \int d\phi_{\pm}^j \mathcal{W}_{\pm}^j. \quad (43)$$

As such, they encode the full information of ϕ_{\pm}^j and π_{\pm}^j , but no information about their correlations. Hence, in the special case when ϕ_{\pm}^j and π_{\pm}^j are uncorrelated, the corresponding Wigner W -distribution decomposes into a product $W_{\pm}^j = f_{\pm}^j g_{\pm}^j$.

The separate detection scheme, or homodyne detection [79], described in section II E 1, approximately corresponds to sampling from the marginal distributions. Fundamentally, of course, this cannot generally be an exact correspondence for any finite atom number, as the detected quantities $\mathbf{n}_{m_F}^j$ have a discrete spectrum (detected atom numbers), while the desired Wigner marginals are continuous. However, in the undepleted pump approximation, the Wigner marginals are generally well-approximated by the homodyne statistics [69, 70].

3. Husimi Q -distribution

Since the coherent state projectors $|\alpha_{\pm}^j\rangle \langle \alpha_{\pm}^j|$ are non-negative and resolve the identity, they constitute a posi-

tive operator-valued measure (POVM). The corresponding measurement distribution is the so-called Husimi Q -distribution [26]

$$\mathcal{Q}_{\pm}^j \equiv \mathcal{Q}_{\pm}^j(\phi_{\pm}^j, \pi_{\pm}^j) = \text{Tr} \left\{ \rho_{\pm}^j |\alpha_{\pm}^j\rangle \langle \alpha_{\pm}^j| \right\}. \quad (44)$$

While the Wigner W -distribution can be negative, the Husimi Q -distribution is *always* non-negative [27], since it stems from a POVM. More specifically, it is bounded by $0 \leq \mathcal{Q}_{\pm}^j \leq 1$ [29, 30]. The normalization is induced by (37), leading to

$$1 = \text{Tr} \left\{ \rho_{\pm}^j \right\} = \int \frac{d\phi_{\pm}^j d\pi_{\pm}^j}{2\pi} \mathcal{Q}_{\pm}^j. \quad (45)$$

When operating in the regime $n_0^j \gg n_{\pm 1}^j$ the simultaneous detection scheme described in section II E 2 can be identified with the so-called eight-port-homodyne detection in quantum optics (also referred to as heterodyne detection). In this scheme, each measurement corresponds to drawing a sample from the Husimi- Q -distribution of the signal field [80].

IV. SIMULATING THE DYNAMICS

We simulate the dynamics of the spin-1 BEC in two different approximations. To capture most of the full Hamiltonian and to check the validity of the undepleted pump approximation, we perform a truncated Wigner simulation, see section IV A. Additionally, we derive an analytically solvable model based on (28) in section IV B.

A. Truncated Wigner Approximation

In the regime of high occupations, it is usually infeasible to exactly solve the exponentially complex dynamics in the full Hilbert space \mathcal{H} . Instead, as the populations are increased, the role of quantum fluctuations compared to the mean-field dynamics grows smaller and hence the predictions made by semiclassical approximation methods become more and more accurate. One such semiclassical technique is the truncated Wigner approximation (TWA). At its core, all mode operators (we drop all indices in this subsection for brevity) are replaced by complex numbers [81]

$$\mathbf{a}^{(\dagger)} \rightarrow \alpha^{(*)}, \quad (46)$$

which amounts to a lowest-order expansion of the Wigner-Weyl correspondence rules. Consequently, the Hamiltonian operator $\mathbf{H}(\mathbf{a}, \mathbf{a}^{\dagger})$ reduces to a classical Hamiltonian function $H(\alpha, \alpha^*)$ of the complex-valued phase-space coordinates (α, α^*) .

In general, the von Neumann equation

$$i\partial_t \mathbf{O}(t) = [\mathbf{O}(t), \mathbf{H}], \quad (47)$$

for some time-dependent operator \mathbf{O} , can be translated into an equation of motion for the operator's Weyl symbol O in phase space via the Moyal bracket

$$\{A, B\}_{\text{MB}} = \frac{A}{2} \sin \left[2 \left(\overleftarrow{\partial}_\alpha \overrightarrow{\partial}_{\alpha^*} - \overleftarrow{\partial}_{\alpha^*} \overrightarrow{\partial}_\alpha \right) \right] B, \quad (48)$$

which results in [81]

$$i\partial_t O(t) = \{O(t), H\}_{\text{MB}}. \quad (49)$$

In TWA, one is interested in the evolution of the elementary c -numbers α representing the mode operators to leading order, for which Eq. (49) simplifies to the classical Poisson bracket

$$i\partial_t \alpha(t) = \partial_{\alpha^*(t)} H, \quad (50)$$

since all higher-order derivatives vanish.

In the case when multiple modes are coupled as for the Hamiltonian of interest in Eq. (6), Eq. (50) defines a system of coupled differential equations that may be solved numerically using a suited integrator. Before being propagated in time, the initial set of coordinates $\alpha(0)$ of a given well j and mode m_F is obtained as Monte Carlo samples from the initial Wigner W -distribution, for which two cases are relevant. When considering the vacuum $|0\rangle$, the Wigner W -distribution takes the form of a Gaussian, which is centered at the origin and contains no correlations, i.e.

$$|\psi\rangle = |0\rangle \iff \alpha \sim \mathcal{N} \left[0, \frac{1}{2} \right] + i\mathcal{N} \left[0, \frac{1}{2} \right]. \quad (51)$$

The samples of coherent states $|\gamma\rangle$ have the same variance but are displaced by the square root of their mean particle number $\langle \mathbf{N} \rangle = |\gamma|^2$ from the origin

$$|\psi\rangle = |\gamma\rangle \iff \alpha \sim \mathcal{N} \left[\text{Re}(\gamma), \frac{1}{2} \right] + i\mathcal{N} \left[\text{Im}(\gamma), \frac{1}{2} \right], \quad (52)$$

since $|\gamma\rangle$ itself is obtained by acting with the displacement operator $\mathbf{D}(\gamma)$ on the vacuum, see (34).

In the presence of thermal fluctuations, the pure vacuum state $|0\rangle$ is replaced by the thermal ensemble $\rho \propto \exp(-\beta \mathbf{a}^\dagger \mathbf{a})$ with the inverse temperature $\beta = 1/T$. Accordingly, the standard deviations of the Wigner W -distribution are rescaled as

$$\frac{1}{2} \rightarrow \sqrt{1 + 2n_{\text{BE}}(\beta)} \frac{1}{2}, \quad (53)$$

which is equivalent to adding $n_{\text{BE}}(\beta)$ to both variances, where

$$n_{\text{BE}}(\beta) = \frac{1}{e^\beta - 1} \quad (54)$$

denotes the Bose-Einstein distribution.

Then, the expectation value of some observable \mathbf{O} is obtained as the stochastic average over all generated samples

$$\langle \mathbf{O} \rangle_{\text{TWA}} = \frac{1}{|\mathcal{S}|} \sum_{\alpha \in \mathcal{S}} O(\alpha, \alpha^*), \quad (55)$$

where \mathcal{S} denotes the set of all samples. Note here that we again rely on the correspondence between \mathbf{O} and O given in Eq. (46).

Having introduced all necessary tools in TWA, we want to briefly comment on its regimes of applicability. In the limit of high occupations $n \gg 1$ and initial states with minimal fluctuations, i.e. coherent states, the operator to c -number correspondence Eq. (46) is a justified simplification, since the relative fluctuations scale as $1/\sqrt{n}$. The fluctuations of the initial state are herein represented accurately, as they correspond to exact samples of the Wigner W -distribution of the initial state $|\psi\rangle$, while only those fluctuations that build up during the evolution due to the unitary evolution under \mathbf{H} are not captured exactly. In the limit $\hbar \rightarrow 0$, the dynamics generated by TWA become exact. For a more thorough picture of this matter, see Ref. [82], in which the impossibility of quantum phase-space trajectories for generic (i.e. anharmonic) quantum systems is discussed. For further reading regarding TWA and other semiclassical techniques we refer the reader to [81] and [83].

B. Gaussian Model

Interestingly, the short-time dynamics can be described by a simple Gaussian model to a reasonable extent, which we discuss next.

1. Gaussian Hamiltonian

Starting from the undepleted pump Hamiltonian (28), we make two simplifying assumptions to arrive at a quadratic model: First, we neglect the atomic collisions between the atoms, i.e. set $c_0 \approx 0$, which amounts to dropping all fourth- and one second-order contribution in (29). This results in a vanishing mixing Hamiltonian $\mathbf{H}_{\text{up}}^{\text{mix}} \approx 0$, which disentangles the \pm phase spaces and hence allows us to study their dynamics independently. Second, despite no explicit translational invariance, an approximate invariance exists in the center of the lattice. Since boundary effects are negligible for the subsystem of our interest (see section V A), we can assume all parameters to be equal across all wells, i.e. $\tilde{q}^j \approx \tilde{q}$, $\tilde{c}_1^j \approx \tilde{c}_1$ and $n^j \approx n$.

The resulting Hamiltonian is Gaussian, i.e. of second order in the relative mode operators, and reads for the $+$ mode (the $-$ mode can be treated on equal footing)

$$\begin{aligned} \mathbf{H}_{\text{up, Gauss}}^+ &= \sum_{j=1}^N \left[\tilde{q} \mathbf{N}_+^j + \frac{\tilde{c}_1}{2} \left(\mathbf{a}_+^j \mathbf{a}_+^j + \mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j\dagger} \right) \right] \\ &+ J \sum_{j=1}^{N \text{ or } N-1} \left(\mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j+1} + \mathbf{a}_+^{(j+1)\dagger} \mathbf{a}_+^j \right). \end{aligned} \quad (56)$$

The dynamics generated by the latter Hamiltonian is illustrated in Figure 3. As a result of its simple form,

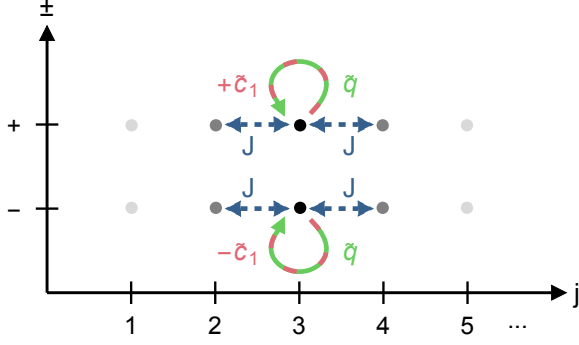


Figure 3. Sketch of the dynamics in the Gaussian regime generated by (56) (in analogy to Figure 1).

we can provide a more detailed description of how the different terms in (56) contribute to the overall dynamics (see Figure 4): The first term proportional to \tilde{q} causes a local rotation (green dashed circle in **d**), while the second term proportional to \tilde{c}_1 generates local squeezing (red dashed arrows in **d**). Both act locally on the j -th. $+$ phase space. The coupling term proportional to J builds up correlations between the wells (blue solid arrow in **i**) by stretching the local distributions proportionally to each other (see **g**) versus **h**)).

One may also express (56) through the canonical field operators (see Appendix A for details), leading to

$$\begin{aligned} H_{\text{up,Gauss}}^+ &= \frac{1}{2} \sum_{j=1}^N \left[\lambda^+ (\phi_+^j)^2 + \lambda^- (\pi_+^j)^2 \right] \\ &\quad - J \sum_{j=1}^N \text{or } N-1 \left(\phi_+^j \phi_+^{j+1} + \pi_+^j \pi_+^{j+1} \right) \quad (57) \\ &\equiv \frac{1}{2} \sum_{j,j'=1}^{2N} \chi_+^{jT} \Lambda^{jj'} \chi_+^{j'}, \end{aligned}$$

where we introduced the two new couplings

$$\lambda^\pm = \tilde{q} \pm \tilde{c}_1 \quad (58)$$

for brevity¹ and the so-called Hamiltonian matrix Λ . The latter is a $2N \times 2N$, real and symmetric matrix of block-diagonal form

$$\Lambda = \begin{pmatrix} \Lambda^+ & 0 \\ 0 & \Lambda^- \end{pmatrix}, \quad (59)$$

where Λ^+ and Λ^- describe the dynamics of the fields ϕ_+^j and the momentum fields π_+^j , respectively. Both have the same form with non-vanishing entries only on the three

leading diagonals for open boundary conditions, i.e.

$$\Lambda^\pm = \begin{pmatrix} \lambda^\pm & -J & 0 & 0 \\ -J & \lambda^\pm & -J & 0 \\ 0 & -J & \lambda^\pm & \ddots \\ 0 & 0 & \ddots & \ddots \end{pmatrix}, \quad (60)$$

and additionally a contribution $-J$ in the first and last anti-diagonal entries for periodic boundary conditions. Hence, the diagonal entries in (60) describe single wells, while the next-to-leading diagonals (and possibly the outer two entries of the anti-diagonal) contain the coupling between neighboring wells. We also note that the Hamiltonian (57) is bounded from below by a positive number, i.e. has a positive-energy ground state, when $\lambda^\pm \geq 0$, which is fulfilled if $q \geq c_1/2$ and hence in particular for non-negative $q \geq 0$.

2. Mean and covariance matrix

It is well known that quadratic Hamiltonians map Gaussian states to Gaussian states [74, 75]. Since the initial state $|0, \alpha, 0\rangle^{\otimes N}$ from Eq. (16) corresponds to an uncorrelated set of local vacua in the $+$ phase spaces, the initial state in the undepleted pump approximation is indeed Gaussian. Hence, the dynamics generated by (57) is fully encoded in the two lowest-order moments of the quantum state ρ_+ , i.e. its field expectation values

$$\chi_+^j = \text{Tr}\{\rho_+ \chi_+^j\} \quad (61)$$

and its covariance matrix [74, 75]

$$\gamma_+^{jj'} = \frac{1}{2} \text{Tr}\{\rho_+ \{\chi_+^j - \chi_+^j, \chi_+^{j'} - \chi_+^{j'}\}\}. \quad (62)$$

The latter is a real, symmetric, and non-negative $2N$ -dimensional matrix of block form

$$\gamma_+ = \begin{pmatrix} \mathcal{M}_+ & \mathcal{T}_+ \\ \mathcal{T}_+^T & \mathcal{N}_+ \end{pmatrix}, \quad (63)$$

and contains the two fundamental two-point correlation functions

$$\begin{aligned} \mathcal{M}_+^{jj'} &= \text{Tr}\{\rho_+ \phi_+^j \phi_+^{j'}\} - \phi_+^j \phi_+^{j'}, \\ \mathcal{N}_+^{jj'} &= \text{Tr}\{\rho_+ \pi_+^j \pi_+^{j'}\} - \pi_+^j \pi_+^{j'}, \end{aligned} \quad (64)$$

as well as the mixed correlator

$$\mathcal{T}_+^{jj'} = \text{Tr}\{\rho_+ \phi_+^j \pi_+^{j'}\} - \frac{i}{2} \delta_{jj'} - \phi_+^j \pi_+^{j'}. \quad (65)$$

Importantly, the covariance matrix is constrained by the Robertson-Schrödinger uncertainty relation in the form of a lower bound to its determinant [84–86]

$$\det \gamma_+ \geq \frac{1}{2^{2N}}, \quad (66)$$

¹ Note here that λ^\pm does *not* refer to the \pm phase spaces.

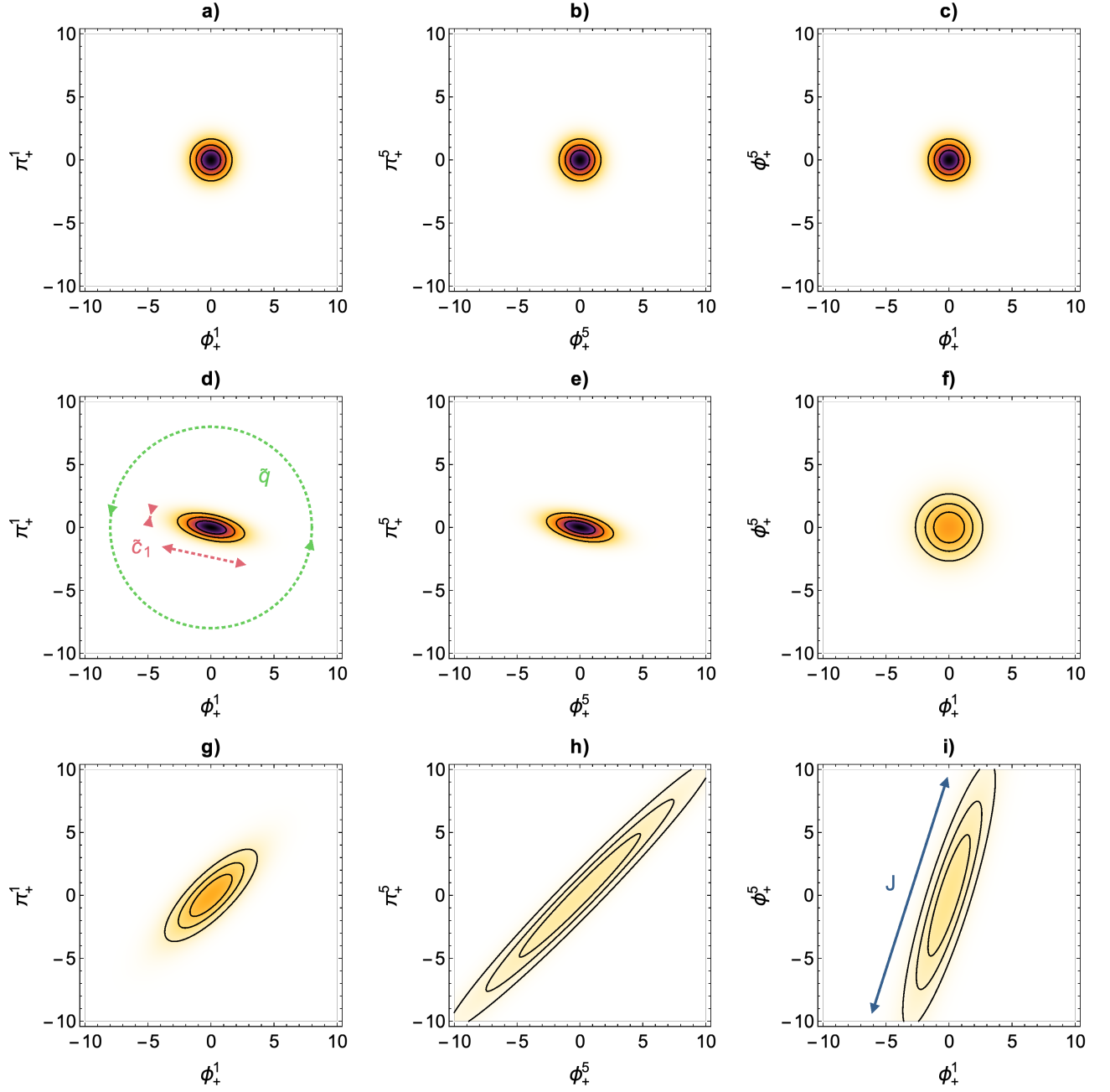


Figure 4. Illustrations of the effects of the three terms in the Gaussian Hamiltonian (56) on the $+$ phase-space distributions for $N = 9$ wells and open boundary conditions. The rows show the initial state, an evolved state for uncoupled wells, i.e. $J = 0$, and an evolved state for coupled wells, i.e. $J > 0$, respectively, while the columns depict the phase-space distributions of the first well, the fifth well in the center as well as the correlations in the fields between these two wells, respectively. For $t = 0$, see **a)**-**c)**, all wells are in the same state, and the wells are uncorrelated. When the uncoupled system has evolved, see **d)**-**f)**, all local distributions are equally squeezed (red dashed arrows) and rotated (green dashed arrow) by \tilde{c}_1 and \tilde{q} , respectively, and remain uncorrelated. For a non-zero tunnel rate $J > 0$, see **g)**-**i)**, the local distributions are stretched proportionally (note that the stretching is equal across the wells for periodic boundary conditions), see **g)** and **h)**, such that correlations between the wells build up, see **i)** (blue solid arrow).

with equality if and only if ρ_+ corresponds to a pure Gaussian state [74, 75]. Thus, for our analytical model, $\det \gamma_+ > 1/2^{2N}$ encodes mixedness of ρ_+ (see also [section VB 2](#)).

Importantly, χ_+ and γ_+ correspond to the first- and second-order moments of the Wigner W -distribution \mathcal{W}_+ , and the diagonal entries of γ_+ are encoded in f_+ and g_+ . While χ_+ is also the mean of the Husimi Q -distribution

\mathcal{Q}_+ , its covariance matrix V_+ acquires an additional contribution, which is a consequence of \mathcal{Q}_+ being a convolution of \mathcal{W}_+ , resulting in

$$V_+ = \gamma_+ + \frac{1}{2} \mathbb{1}. \quad (67)$$

3. Symplectic time evolution

The effect of the evolution under the Hamiltonian (57) on the two lowest-order moments can be described by symplectic transformations in phase space [74, 75]. To that end, we recall a few basics of Gaussian quantum information theory. The Hamiltonian $\mathbf{H}_{\text{up,Gauss}}^+$ generates the time evolution of the initial state $\rho_+(0)$ at $t = 0$ via the unitary time evolution operator

$$\mathbf{U}_{\text{up,Gauss}}^+(t) = e^{-it\mathbf{H}_{\text{up,Gauss}}^+}, \quad (68)$$

in the sense that

$$\rho_+(t) = \mathbf{U}_{\text{up,Gauss}}^+(t) \rho_+(0) \mathbf{U}_{\text{up,Gauss}}^{+\dagger}(t). \quad (69)$$

In phase space, the unitary transformation $\mathbf{U}_{\text{up,Gauss}}^+(t)$ corresponds to a symplectic transformation $\mathcal{S}_{\text{up,Gauss}}^+(t)$, that is, a transformation which preserves the commutator in phase space (see Eq. (32)), i.e. leaves the symplectic form Ω invariant

$$\mathcal{S}_{\text{up,Gauss}}^+(t) \Omega \mathcal{S}_{\text{up,Gauss}}^{+T}(t) = \Omega. \quad (70)$$

When the Hamiltonian is written as a bilinear form as in the second line in (57), the symplectic time evolution matrix can be expressed through the Hamiltonian matrix Λ via [74, 75]

$$\mathcal{S}_{\text{up,Gauss}}^+(t) = e^{t\Omega\Lambda}. \quad (71)$$

This matrix can directly be applied to the two lowest-order moments of ρ_+ , yielding the simple time evolution equations [74, 75]

$$\begin{aligned} \chi_+(t) &= \mathcal{S}_{\text{up,Gauss}}^+(t) \chi_+(0), \\ \gamma_+(t) &= \mathcal{S}_{\text{up,Gauss}}^+(t) \gamma_+(0) \mathcal{S}_{\text{up,Gauss}}^{+T}(t). \end{aligned} \quad (72)$$

Since the initial state is the uncoupled vacuum state, we have $\chi_+(0) = 0$, and hence all phase-space distributions remain centered around the origin, i.e.

$$\chi_+(t) = \chi_+(0) = 0. \quad (73)$$

Instead, the initial covariance matrix

$$\gamma_+(0) = \frac{1}{2} \mathbb{1}, \quad (74)$$

evolves non-trivially following the second line of (72), which can be computed analytically for a given set of parameters.

Note that for a non-zero initial temperature, the initial covariance matrix is also proportional to the identity, but the variances increase due to thermal fluctuations, which results in

$$\gamma_+(0) = \left[\frac{1}{2} + n_{\text{BE}}(\beta) \right] \mathbb{1}. \quad (75)$$

Note also that adding half the identity to $\gamma_+(t)$ gives the time-evolved covariance matrix $V_+(t)$ of the Husimi Q -distribution. Since $\mathcal{S}\mathcal{S}^T = \mathbb{1}$ for all symplectic matrices \mathcal{S} one may equally apply the symplectic transformation directly to the defining equation of V_+ , i.e. (67).

4. Comment on field theory correspondence

At last, we comment on the field theory the Gaussian Hamiltonian (57) reproduces when taking the continuum limit. To that end, we introduce a lattice spacing ϵ and define the field operators

$$\phi_+(x) \equiv \frac{1}{\sqrt{\epsilon}} \phi_+^j, \quad \pi_+(x) \equiv \frac{1}{\sqrt{\epsilon}} \pi_+^j, \quad (76)$$

which fulfill the distribution-valued commutation relations

$$[\phi_+(x), \pi_+(x')] = i\delta(x - x'), \quad (77)$$

in the continuum limit $\epsilon \rightarrow 0$. Then, using (76) in (57) together with

$$\phi_+(x + \epsilon) = \phi_+(x) + \epsilon \partial_x \phi_+(x) + \mathcal{O}(\epsilon^2), \quad (78)$$

analogously for $\pi_+(x + \epsilon)$, results in

$$\mathbf{H}_{\text{up,Gauss}}^+ = \frac{1}{2} \int dx [\kappa^+(x) \phi_+^2(x) + \kappa^-(x) \pi_+^2(x)], \quad (79)$$

with the differential operators

$$\kappa^\pm(x) = \lambda^\pm + 1 + J\epsilon \partial_x, \quad (80)$$

to leading order in ϵ . Hence, a reasonable continuum limit requires the limit of infinite coupling $J \rightarrow \infty$ with the product $J\epsilon$ kept fixed, which effectively implements $\epsilon = 1/J \rightarrow 0$.

V. AREA LAW IN PHASE SPACE

We now discuss the scaling behaviors of quantum entropies associated with a subsystem. Then, we define *classical* entropies of the phase space distributions introduced in section III D and provide mathematical as well as heuristic arguments for why these entropies encode the area law.

A. Quantum entropies

We consider a subsystem of five wells from the N total wells, which can exchange energy and particles with its complement and therefore should be considered an open quantum system. We partition this subsystem into a subregion A consisting of the first $0 \leq M \leq 5$ wells and its complement B composed out of the remaining $5 - M$ wells (see Figure 5). The multi-well setup offers the possibility to study the discretized version of a continuous quantum field theory with the wells corresponding to lattice points. The local states associated with the subregions are, as usual, defined via the partial trace

$$\rho_+^A = \text{Tr}_B\{\rho_+\} \quad (81)$$

and analogous for ρ_+^B . The local mixedness of A is conveniently measured by the von Neumann entanglement entropy [87]

$$S(\rho_+^A) = -\text{Tr}\{\rho_+^A \ln \rho_+^A\}, \quad (82)$$

which serves as an entanglement measure if and only if the bipartite state ρ_+ is pure, in which case $S(\rho_+^A) > 0$ if and only if A and B are entangled and $S(\rho_+^A) = 0$ if not. However, when the bipartite state is mixed, for example, due to the presence of thermal fluctuations or coupling to the environment as considered here, one typically studies the quantum mutual information instead. It is defined as [88, 89]

$$I(\rho_+^A : \rho_+^B) = \text{Tr}\{\rho_+ (\ln \rho_+ - \ln \rho_+^A \otimes \rho_+^B)\}, \quad (83)$$

or equivalently as

$$I(\rho_+^A : \rho_+^B) = S(\rho_+^A) + S(\rho_+^B) - S(\rho_+), \quad (84)$$

if all three terms on the right are finite, and serves as the measure for the total correlations between A and B being zero if and only if $\rho_+ = \rho_+^A \otimes \rho_+^B$. Note that when the initial state is pure, the entanglement entropy and the quantum mutual information are proportional to each other, i.e. $I(\rho_+^A : \rho_+^B) = 2S(\rho_+^A)$.

Both quantities have been studied extensively for a large variety of quantum field theories and quantum many-body systems and it is well-known that for *typical* states of Hamiltonians with local interactions, they scale with the *area* of the entangling surface separating A and B , see e.g. [7, 8, 10, 11, 15] for reviews. In this context, states are called typical for example when they lie sufficiently close to the ground state of the Hamiltonian or when they are generated from an initial product state on short time scales after a quench [15]. In the following, we shall be concerned with the latter scenario.

For our setup, three competing length scales are relevant for the precise scaling with subsystem size M of the aforementioned quantum entropies: the system size N , the (initial) inverse temperature $\beta = 1/T$, and the subsystem size M itself. When the dominating scale is M , i.e. when

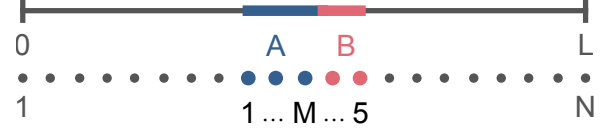


Figure 5. The open subsystem of interest (colored) is partitioned into A (blue) and B (red).

we consider the subinterval size being small compared to the system size and the inverse temperature $M \ll N, \beta$, then the entanglement entropy scales logarithmically (blue solid curve in Figure 6) [7, 8, 10, 11]

$$S(\rho_+^A) \sim \kappa_1 \ln(M + \kappa_2) + \kappa_3, \quad (85)$$

with κ_i being real constants. We normalize all entropic quantities such that they vanish at $M = 0$, resulting in the constraint

$$\kappa_2 = e^{-\kappa_3/\kappa_1}, \quad (86)$$

and hence only κ_1 and κ_3 are independent. Typically, κ_1 is universal, i.e. regularization independent, especially in conformal field theories where it corresponds to the central charge, while κ_3 is usually not [7, 8, 10].

When instead N is dominant, i.e. when the subinterval size is of the order of the system size $M \sim N$ and the temperature is still small $N \ll \beta$, the so-called finite-size area law holds (petrol dashed curve) [7, 8]

$$S(\rho_+^A) \sim \kappa_1 \ln \left[\frac{N}{\pi} \sin \left(\frac{\pi M}{N} \right) + \kappa_2 \right] + \kappa_3, \quad (87)$$

again subject to the constraint (86). Since the quantum state of the whole system is close to being pure in this case, we have $S(\rho_+^A) \rightarrow 0$ when $M \rightarrow N$. Further, in the limit of small $M \ll N$, the latter equation reduces to (85) since $(N/\pi) \sin(\pi M/N) = M$ to first order in M .

If finite-temperature effects become relevant, i.e. for $\beta \sim N$, we instead expect a scaling of the form (red dotted curve) [7, 8]

$$S(\rho_+^A) \sim \kappa_1 \ln \left[\frac{\beta}{\pi} \sinh \left(\frac{\pi M}{\beta} \right) + \kappa_2 \right] + \kappa_3, \quad (88)$$

again constrained by (86). In this case, the entropy remains finite if computed for the whole system due to additional (classical) mixedness. While for small temperatures $\beta \gg N$ we obtain back (85), large temperatures $\beta \ll N$ lead to an extensive entropy obeying a volume law (green dot-dashed curve)

$$S(\rho_+^A) \sim \kappa_1 M, \quad (89)$$

in which case classical correlations dominate over the quantum ones.

Remarkably, the quantum mutual information follows the finite-size area law (87) in all of the four considered

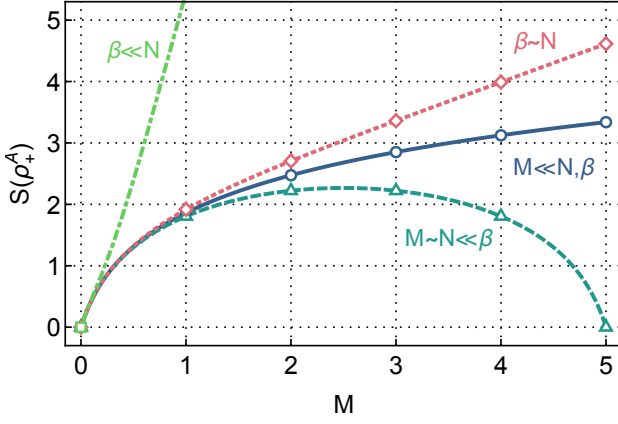


Figure 6. Scaling of the entanglement entropy $S(\rho_+^A)$ with subsystem size M for the four archetypal scenarios: large system size at zero temperature (blue curve), finite system size at zero temperature (petrol curve), finite temperature (red curve) and high temperature (green curve). For small M , the first three curves converge.

cases [18]. In particular, even if thermal fluctuations prevent the appearance of the area law in the entropies, the extensive contributions drop out when considering the quantum mutual information, which demonstrates its utility for observing an area law in actual experiments.

B. Classical entropies

We now introduce various entropic measures in the phase space associated with subsystem A (subsystem B and the full subsystem AB can be treated on equal footing). To that end, we shall group the set of corresponding local field variables according to

$$\chi_+^A = (\phi_+^A, \pi_+^A)^T = (\phi_+^1, \dots, \phi_+^M, \pi_+^1, \dots, \pi_+^M)^T. \quad (90)$$

Then, the distributions corresponding to A are obtained by integrating out all degrees of freedom belonging to subsystem B , for instance

$$\mathcal{W}_+^A = \int d\phi_+^B d\pi_+^B \mathcal{W}_+, \quad (91)$$

and similarly for the other two types of phase space distributions.

1. Standard entropies

Let us start with the Wigner entropy, which is defined as the differential entropy of the Wigner W -distribution, namely

$$S(\mathcal{W}_+^A) = - \int d\phi_+^A d\pi_+^A \mathcal{W}_+^A \ln \mathcal{W}_+^A. \quad (92)$$

It is well-defined, i.e. real, provided that $\mathcal{W}_+^A \geq 0$, otherwise it becomes complex-valued. A lower bound encoding the uncertainty principle has been conjectured in [90–92] and reads

$$S(\mathcal{W}_+^A) \geq S(\bar{\mathcal{W}}_+^A) = M(1 + \ln \pi), \quad (93)$$

with equality if and only if \mathcal{W}_+^A corresponds to a product of pure Gaussian states.

The marginal entropy of f_+^A is defined as

$$S(f_+^A) = - \int d\phi_+^A f_+^A \ln f_+^A, \quad (94)$$

and analogously for g_+^A . Since f_+^A and g_+^A are true probability density functions, their entropies are always well-defined. The corresponding entropic uncertainty relation has been put forward by Białynicki-Birula and Mycielski [93–96] (see also [97, 98] for reviews)

$$S(f_+^A) + S(g_+^A) \geq S(\bar{f}_+^A) + S(\bar{g}_+^A) = M(1 + \ln \pi), \quad (95)$$

which contains the same bound as the Wigner entropy conjecture (93). In fact, (95) would be a simple consequence of (93) when using the subadditivity of entropy, namely

$$S(\mathcal{W}_+^A) \leq S(f_+^A) + S(g_+^A), \quad (96)$$

with equality if and only if $\mathcal{W}_+^A = f_+^A g_+^A$. However, the bound in (95) is less tight since it is attained only for products of squeezed vacuum states.

At last, we introduce the Wehrl entropy [31, 32]

$$S(\mathcal{Q}_+^A) = - \int \frac{d\phi_+^A d\pi_+^A}{(2\pi)^M} \mathcal{Q}_+^A \ln \mathcal{Q}_+^A, \quad (97)$$

which is also always well-defined since $\mathcal{Q}_+^A \geq 0$. It is bounded from below by the Wehrl-Lieb inequality [31–34]

$$S(\mathcal{Q}_+^A) \geq S(\bar{\mathcal{Q}}_+^A) = M, \quad (98)$$

which is tight if and only if the state is a product of pure coherent states. It is also bounded by the Wigner and quantum entropies [37]

$$S(\mathcal{Q}_+^A) \geq S(\mathcal{W}_+^A) - M \ln \pi, \quad S(\mathcal{Q}_+^A) \geq S(\rho_+^A). \quad (99)$$

Note here that the additional constant $-M \ln \pi$ is a result of the different normalizations of the Wigner W - and Husimi Q -distributions. Therefore, (98) would also be a direct consequence of (93).

The three aforementioned types of entropies have in common that their lower bounds (93), (95) and (98) are attained for a product of pure vacuum states and scale with the volume of the subsystem A . In this sense, all these entropies are *classical* to leading order for arbitrary states with the main contribution coming from the vacuum, in contrast to the quantum entropies introduced in section V A which vanish for all pure states. In what follows, we will argue that quantum features such as the area law are still present in classical entropies, but hidden in the next-to-leading order terms.

2. Subtracted entropies

Recently, it has been argued that quantum features such as the area law of the entanglement entropy are also present in classical entropies when the leading order extensive contributions are subtracted [24]. Based on this idea, we define *subtracted* classical entropies by subtracting the extensive vacuum contributions appearing in the bounds of the corresponding entropic uncertainty relations, to wit

$$\begin{aligned}\Delta S(\mathcal{W}_+^A) &\equiv S(\mathcal{W}_+^A) - S(\bar{\mathcal{W}}_+^A), \\ \Delta S(f_+^A, g_+^A) &\equiv S(f_+^A) + S(g_+^A) - S(\bar{f}_+^A) - S(\bar{g}_+^A), \quad (100) \\ \Delta S(\mathcal{Q}_+^A) &\equiv S(\mathcal{Q}_+^A) - S(\bar{\mathcal{Q}}_+^A).\end{aligned}$$

Let us provide some simple arguments for why these entropies encode the very same features as their quantum analogs in the following.

First, we consider Gaussian quantum states, which provide a reasonable first-order approximation to the states generated by the Hamiltonian (6), see [section IV B](#). A general Gaussian Wigner W -distribution is of the form

$$\mathcal{W}_+^A = \frac{1}{Z_+^A} e^{-\frac{1}{2}(\mathbf{x}_+^A)^T (\gamma_+^A)^{-1} \mathbf{x}_+^A}, \quad (101)$$

with $Z_+^A = (2\pi)^M \sqrt{\det \gamma_+^A}$ being a normalization constant. Computing the subtracted Wigner entropy for such a state gives [75, 99]

$$\Delta S(\mathcal{W}_+^A) = \frac{1}{2} \ln (2^{2M} \det \gamma_+^A) = S_2(\boldsymbol{\rho}_+^A), \quad (102)$$

where

$$S_2(\boldsymbol{\rho}_+^A) = -\ln \text{Tr}\{(\boldsymbol{\rho}_+^A)^2\} \quad (103)$$

denotes the quantum Rényi-2 entropy of the state $\boldsymbol{\rho}_+^A$ that corresponds to \mathcal{W}_+^A . Hence, the subtracted Wigner entropy *coincides* with a quantum entropy for Gaussian states. For completeness, let us also give the Gaussian expressions for the Husimi-based quantities, which read

$$\mathcal{Q}_+^A = \frac{1}{Z_+^A} e^{-\frac{1}{2}(\mathbf{x}_+^A)^T (V_+^A)^{-1} \mathbf{x}_+^A}, \quad (104)$$

with $Z_+^A = \sqrt{\det V_+^A}$ and

$$\Delta S(\mathcal{Q}_+^A) = \frac{1}{2} \ln \det V_+^A. \quad (105)$$

Note, however, that these entropies can not be related with quantum entropies in general.

For arbitrary states, including Wigner-negative states, a general relation between subtracted classical and quantum entropies can only be established for the subtracted Rényi-2 Wigner entropy [100, 101]

$$\Delta S_2(\mathcal{W}_+^A) = S_2(\boldsymbol{\rho}_+^A). \quad (106)$$

Note here that the entropic order alters the lower bound in the corresponding entropic uncertainty relation and hence a different term compared to $\Delta S(\mathcal{W}_+^A)$ has to be subtracted. Although the relation (106) can *not* be generalized to other subtracted Rényi-Wigner entropies, it has been shown in [24] that the crucial feature of quantum entropies, i.e. the area law, is present for the entire family of subtracted Rényi-Wigner entropies beyond the Gaussian case, indicating that the area law may also appear for the classical entropies defined in (100), which we investigate in detail in [section VII](#).

C. Classical mutual informations

Let us also introduce classical mutual informations as measures for correlations in phase space. For Wigner-positive states, we define the Wigner mutual information

$$I(\mathcal{W}_+^A : \mathcal{W}_+^B) = S(\mathcal{W}_+^A) + S(\mathcal{W}_+^B) - S(\mathcal{W}_+), \quad (107)$$

while for general states we define the marginal mutual informations

$$I(f_+^A : f_+^B) = S(f_+^A) + S(f_+^B) - S(f_+), \quad (108)$$

(similarly for g_+) and the Wehrl mutual information [37]

$$I(\mathcal{Q}_+^A : \mathcal{Q}_+^B) = S(\mathcal{Q}_+^A) + S(\mathcal{Q}_+^B) - S(\mathcal{Q}_+). \quad (109)$$

All of them are non-negative functionals being zero if and only if the two local distributions on A and B are uncorrelated.

It has been shown that every classical mutual information is a lower bound to the quantum mutual information [37, 102]. In particular, we have

$$I(\mathcal{W}_+^A : \mathcal{W}_+^B), I(\mathcal{Q}_+^A : \mathcal{Q}_+^B) \leq I(\boldsymbol{\rho}_+^A : \boldsymbol{\rho}_+^B), \quad (110)$$

indicating that classical mutual informations do *not* capture all correlations in general. However, the bound (110) is expected to be tighter than the standard second-order lower bounds on the quantum mutual information in terms of two-point correlation functions (which only faithfully describe Gaussian correlations).

In contrast to classical entropies, extensive contributions naturally cancel when considering classical mutual informations. For example, decomposing (107) in the sense of (84) shows that

$$I(\mathcal{W}_+^A : \mathcal{W}_+^B) = \Delta S(\mathcal{W}_+^A) + \Delta S(\mathcal{W}_+^B) - \Delta S(\mathcal{W}_+), \quad (111)$$

since the vacuum is uncorrelated, i.e. $S(\bar{\mathcal{W}}_+) = S(\bar{\mathcal{W}}_+^A) + S(\bar{\mathcal{W}}_+^B)$. For Gaussian states, we can find a relation analogous to (102), which reads [99]

$$\begin{aligned}I(\mathcal{W}_+^A : \mathcal{W}_+^B) &= \frac{1}{2} \ln \frac{\det \gamma_+^A \det \gamma_+^B}{\det \gamma_+} \\ &= I_2(\boldsymbol{\rho}_+^A : \boldsymbol{\rho}_+^B),\end{aligned} \quad (112)$$

with the Rényi-2 mutual information being defined as

$$I_2(\rho_+^A : \rho_+^B) = S_2(\rho_+^A) + S_2(\rho_+^B) - S_2(\rho_+). \quad (113)$$

Fortunately, the partly heuristic arguments for the appearance of the area law for classical entropies in [section V B 2](#) can be made rigorous when considering classical mutual informations by adapting the argument in [\[18\]](#). To that end, we first note that all three Hamiltonians [\(6\)](#), [\(28\)](#) and [\(56\)](#) can be decomposed according to

$$\mathbf{H} = \mathbf{H}^A + \mathbf{H}^B + \mathbf{H}^\partial, \quad (114)$$

where \mathbf{H}^A and \mathbf{H}^B denote the Hamiltonians of subsystems A and B , respectively, while \mathbf{H}^∂ contains the local interactions coupling A to B . Further, we consider the rather general class of thermal states for a given Hamiltonian \mathbf{H} , i.e.

$$\rho_+ = \frac{1}{Z} e^{-\beta \mathbf{H}}, \quad (115)$$

which are typically of non-Gaussian form. Then, every distribution \mathcal{O}_+ over measurement outcomes o_+ corresponding to some POVM \mathbf{O}_+ can be written as

$$\mathcal{O}_+ = \frac{1}{Z} e^{-\beta H}, \quad (116)$$

with the classical Hamiltonian H being defined implicitly via

$$\text{Tr}\{e^{-\mathbf{H}} \mathbf{O}_+\} = e^{-H}. \quad (117)$$

For such a distribution, the classical entropy evaluates to

$$S(\mathcal{O}_+) = \ln Z + \beta \langle H \rangle_{\mathcal{O}_+}, \quad (118)$$

where

$$\langle H \rangle_{\mathcal{O}_+} = \int do_+ \mathcal{O}_+(o_+) H \quad (119)$$

is the classical energy expectation value with respect to the distribution $\mathcal{O}_+(o_+)$. Using [\(118\)](#), we find that the (classical) free energy is nothing but

$$F(\mathcal{O}_+) \equiv -\frac{\ln Z}{\beta} = \langle H \rangle_{\mathcal{O}_+} - \frac{S(\mathcal{O}_+)}{\beta}. \quad (120)$$

The latter is minimized by the thermal distribution \mathcal{O}_+ , a fact which can be derived, for instance, using the non-negativity of the classical relative entropy of any given distribution with respect to the thermal distribution \mathcal{O}_+ . As a special case, we can conclude that the thermal free energy is bounded from above by the free energy of the corresponding product distribution, i.e.

$$F(\mathcal{O}_+) \leq F(\mathcal{O}_+^A \mathcal{O}_+^B). \quad (121)$$

Expanding the latter inequality in the sense of [\(120\)](#) and using the additivity of classical entropies for product distributions yields

$$\begin{aligned} S(\mathcal{O}_+^A) + S(\mathcal{O}_+^B) - S(\mathcal{O}_+) \\ \leq \beta \left(\langle H \rangle_{\mathcal{O}_+^A \mathcal{O}_+^B} - \langle H \rangle_{\mathcal{O}_+} \right). \end{aligned} \quad (122)$$

The decomposition [\(114\)](#) implies

$$\langle H \rangle_{\mathcal{O}_+^A \mathcal{O}_+^B} - \langle H \rangle_{\mathcal{O}_+} = \langle H^\partial \rangle_{\mathcal{O}_+^A \mathcal{O}_+^B} - \langle H^\partial \rangle_{\mathcal{O}_+}, \quad (123)$$

which, together with the definition of the classical mutual information, finally leads to

$$I(\mathcal{O}_+^A : \mathcal{O}_+^B) \leq \beta \left(\langle H^\partial \rangle_{\mathcal{O}_+^A \mathcal{O}_+^B} - \langle H^\partial \rangle_{\mathcal{O}_+} \right). \quad (124)$$

The derived upper bound on the classical mutual information is a function of the classical boundary Hamiltonian H^∂ only, which proves the area law for all variants of classical mutual information, including the three of our interest, for local interactions and thermal states.

VI. ENTROPY ESTIMATION

We now turn to the challenging task of estimating classical entropies from finitely many samples. Albeit being a general problem, it arises especially for the system of our interest, since experimental runs are particularly costly. To that end, we rely on suitable k -nearest-neighbor estimators and benchmark their validity for various scenarios. For our implementation, we utilize the ‘Non-parametric Entropy Estimation Toolbox’ for Python publicly available at [\[103\]](#).

A. k -nearest-neighbor estimator

The integrals that give the Wigner or Wehrl entropy, Eqs. [\(92\)](#) and [\(97\)](#), respectively, are in general hard to estimate from samples. Unless the distribution is of a specific form, such as a Gaussian, that allows reexpressing the entropy through low-order correlators, one must use the available samples to approximate local densities of the distribution. A straightforward approach is to bin the samples onto a grid and use the relative frequencies in the histogram as approximations for the local densities. While this procedure presents an asymptotically unbiased estimator in the limit of infinite samples and small bins, it is problematic in practice, since the binning procedure presents an information bottleneck as it coarse grains the information about the precise sample position.

A more elaborate approach is to use the k -nearest-neighbor (k NN) statistics of the sample set to approximate local densities [\[49–63\]](#). The intuition behind the nearest-neighbor statistics is that samples with short distances to their neighbors are located in regions of high probability density, while samples which are located in regions of low probability density have nearest neighbors that are far away. A formalization of this intuition allows to construct an asymptotically unbiased discretization-free estimator \hat{S} for differential entropies [\[49, 56\]](#)

$$\hat{S}(k, N_s) = g(k, N_s, d) + \frac{d}{N_s} \sum_{i=1}^{N_s} \ln \epsilon^i(k) \quad (125)$$

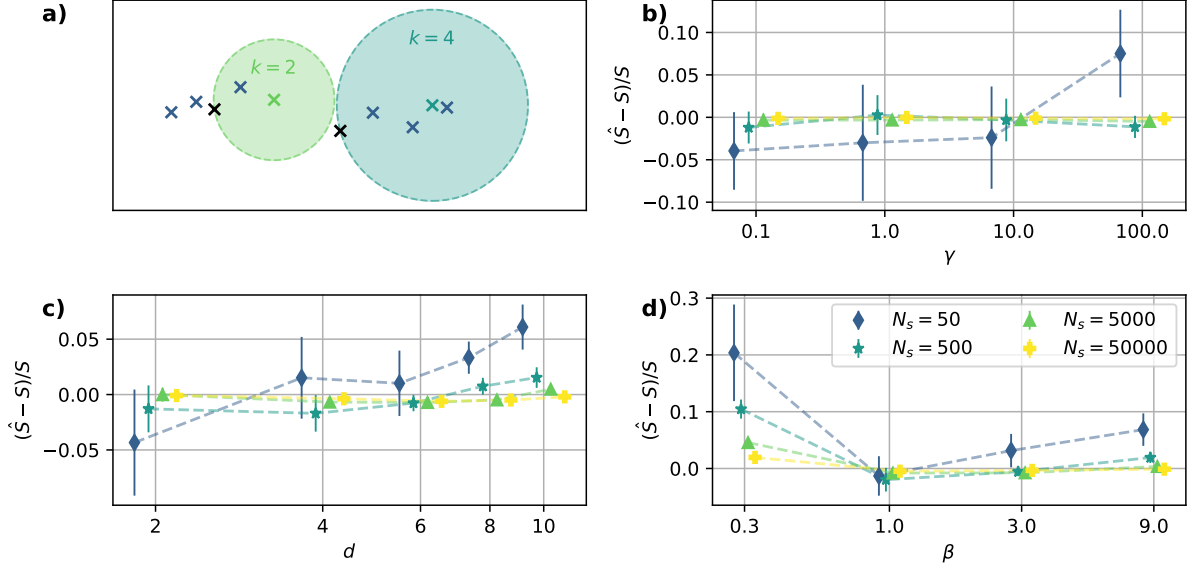


Figure 7. Benchmark of the entropy estimator \hat{S} for distributions that allow for an analytic comparison to S . In **a)**, we visualize the algorithm for a few Gaussian samples in $d = 2$ dimensions, with two random samples selected and two choices for the hyperparameter k , namely $k = 2$ and $k = 4$. The only assumption of the estimator is constant density in the light green and cyan circles respectively, an assumption less justified the larger k . Both k -th neighbors are shown in black. In **b)** we benchmark the entropy estimator on a squeezed covariance matrix in $d = 2$ dimensions with squeezing parameter γ , where large γ correspond to strong squeezing. The covariance matrix is given by $\Sigma = (\mathbf{1} + \gamma J) / \det(\mathbf{1} + \gamma J)^{1/d}$ where J is the matrix of all ones. All error bars are obtained from ten independent estimates and we use offsets on the x axis to aid in differentiating the different data points. In **c)** we let $\gamma = 0$ and vary the dimension d , observing decreasing performance with increasing dimension, as is expected. In **d)** we tune away from the Gaussian regime in $d = 4$ dimensions, using a generalized normal distribution [104] parameterized by β with scale α chosen such that the entropy S takes the value of a standard normal distribution in $d = 4$ dimensions. As can be seen from the density function in Eq. (127), small values of β correspond to heavy tails, $\beta = 2$ corresponds to a Gaussian shape, and the support of the distribution shrinks to the interval $[-1, 1]$ for $\beta \rightarrow \infty$. We denote the sample set size with N_s and choose the nearest neighbor parameter k to be three in all experiments.

where d is the dimension of the distribution, N_s is the number of samples and $\epsilon^i(k)$ is the distance of the i -th sample to its k -th neighbor. The term $g(k, N_s, d)$ is a constant that is independent of the samples and is given by [49, 56]

$$g(k, N_s, d) = -\psi(k) + \psi(N_s) + \ln(c_d), \quad (126)$$

where ψ is the digamma function and c_d is the volume of the d -dimensional unit ball.

The asymptotical unbiasedness of the estimator means that one is guaranteed to obtain the true entropy in the limit of infinitely many samples, which directly stems from the observation that the sample distances must shrink to zero in this limit. This is in contrast to approaches that fit a probability density function of a specific form to the data, which may result in more accurate estimates in the regime of few samples, but will generally not converge to the true value when the amount of samples is increased.

We aim to give an intuitive picture of the k NN method in Figure 7 a), where the spheres, for which constant density is assumed, are drawn for $k = 2$ and $k = 4$ for two different samples. This example reveals the only trade-off in the k NN algorithm, which is the so-called hyperparam-

eter choice k . For small k , the assumption of constant density has stronger justification as the spheres are necessarily smaller compared to larger k . This, however, comes at the expense of stronger fluctuations, meaning larger statistical uncertainties. These can be systematically reduced by choosing larger k , as the volume of the spheres shows a power law in the dimension of the distribution. Typically, k is set to $k = 3$ or similar, depending on the situation at hand [103].

Since the estimation of differential entropies from samples is a task of fundamental interest [105] with applications ranging from statistics [106–108] to signal processing [109, 110] to machine learning and pattern recognition [111–114], various other paths beside the aforementioned one have been explored [115]. There exists a rich literature for non-parametric entropy estimators, that do not assume a specific form of the underlying probability density [49–55, 57–63]. These are either based on nearest neighbor statistics or arrive at an estimate for the local density by using kernel density estimates. Recently, machine learning-inspired techniques have been explored. Giving up on the non-parametric property of the estimator results in the loss of the asymptotically unbiasedness

at possibly increased performance. Numerous works have been put forward in this regard, with applications of estimating entropies [116] and mutual informations [117].

B. Benchmarks in low dimensions

We showcase the performance of the estimator for Gaussian distributions in Figure 7 b). The data points are obtained from 10 independent runs, meaning that a total of $10N_s$ data points were generated, from which we extracted means and standard deviations on the relative deviations $(\hat{S} - S)/S$. We study the performance of the algorithm in $d = 2$ dimensions as a function of the squeezing parameter γ , where $\gamma = 0$ corresponds to a normal distribution and large γ implies higher squeezing. Here, $S = 2.83$ for all values of γ . A central assumption in the derivation of the expression in Eq. (125) is uniform density inside the d -dimensional ball with radius $\epsilon^i(k)$ between sample i and its k -th neighbor. For larger squeezing, this assumption is violated more strongly, since the squeezing introduces a preferred direction at odds with the assumed isotropy, resulting in worse performance for larger γ .

C. Scaling with dimension

In Figure 7 c), we benchmark the estimator on a standard normal distribution for increasing dimension. The higher dimension implies that the assumption of constant density inside the d -dimensional spheres of radius $\epsilon^i(k)$ must hold for larger volumes if the number of samples is held constant. This will generally result in worse performance, as demonstrated.

D. Tuning away from the Gaussian regime

To test the estimator's performance away from the Gaussian regime, we benchmark it on a generalized normal distribution [104]. The latter has a probability density of the form

$$f_{\text{GND}}(x; \mu, \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left(-\left|\frac{x - \mu}{\alpha}\right|^\beta\right), \quad (127)$$

with mean μ , scale $\alpha > 0$, shape $\beta > 0$ and Γ denoting the gamma function. In Figure 7 d), we vary α and β in such a way that the true entropy S is kept constant at the value corresponding to the standard normal distribution in $d = 4$ dimensions, for which $\alpha = 1/2, \beta = 2$, i.e. $S = 5.67$. We observe the worst performance for small values of β , i.e. in the regime of heavy tails, which is expected as the central assumption of constant density of the employed estimator needs to hold for larger regions in space. For large values of β and sufficiently many samples, the results improve (for $N_s = 5000$ (50000) the relative deviations for $\beta = 1$, $\beta = 3$ and $\beta = 9$ are -0.008, -0.007

and 0.003 (-0.004, -0.003 and -0.001)), in line with the intuition that the entropy of a uniform distribution should be easy to estimate. However, for $N_s = 50$ and $N_s = 500$ samples, this intuition seems to be misleading as the best performance is observed for $\beta = 1$ and $\beta = 3$, respectively, which resemble Gaussian-like distributions.

VII. RESULTS

We are now ready to present our main results: analytical and numerical observations of the area law for classical entropies. After a general comparison of the analytical model and TWA in terms of two-point correlators and entropic quantities, we systematically study typical effects, including non-Gaussian features, a thermal initial state, the influence of the subsystem's position within the total system, the total system size, the dependence of the estimated entropies on the number of samples and finally two types of boundary conditions.

Unless specified differently, we use the values $c_1 = -1/n$, $n = 1000$, such that our energy scale is given by $|nc_1| = 1$ (note that this renders *all* parameters and other quantities of our interest dimensionless). Further, we consider Lithium-7 for which $c_0 = -2c_1$, set $q = 2J$ with $J = 2$ and test $N = 20$ wells with open boundary conditions at the three different times $t = 0.5, 0.75, 1$. From the 20 wells, we study five wells located in the center of the lattice, i.e. wells 8, 9, 10, 11, 12, and neglect the remaining 15 wells from our examination, leaving us with the subsystem depicted in Figure 5. We assume the initial temperature to be zero, i.e. $T = 0$, and base the entropy estimation on 10^4 samples. We systematically vary these parameters one by one and investigate their influences from section VII C to section VII H. The plotted curves are either dashed or solid, corresponding to interpolations or fits of the (finite-size) area or the volume law.

A. Correlation matrices

As a first qualitative comparison between the analytical model and the truncated Wigner approximation that takes into account all interactions contained in Eq. (6), we show the correlation matrices between the phase-space variables ϕ_+ and π_+ in Figure 8 for the five wells under scrutiny at the three different times. The correlation matrix is given by the normalized covariances

$$C_+^{jj'} = \frac{\gamma_+^{jj'}}{\sqrt{\gamma_+^{jj} \gamma_+^{j'j'}}}, \quad (128)$$

thereby rendering all entries to lie between -1 and 1. We want to emphasize that the analytical model is *fully* characterized by the observables in Eq. (128), reducing the complexity of extracting entropies to an extraction of its second moments. In contrast to that, the full Hamiltonian Eq. (6) contains more complex, i.e. higher-order

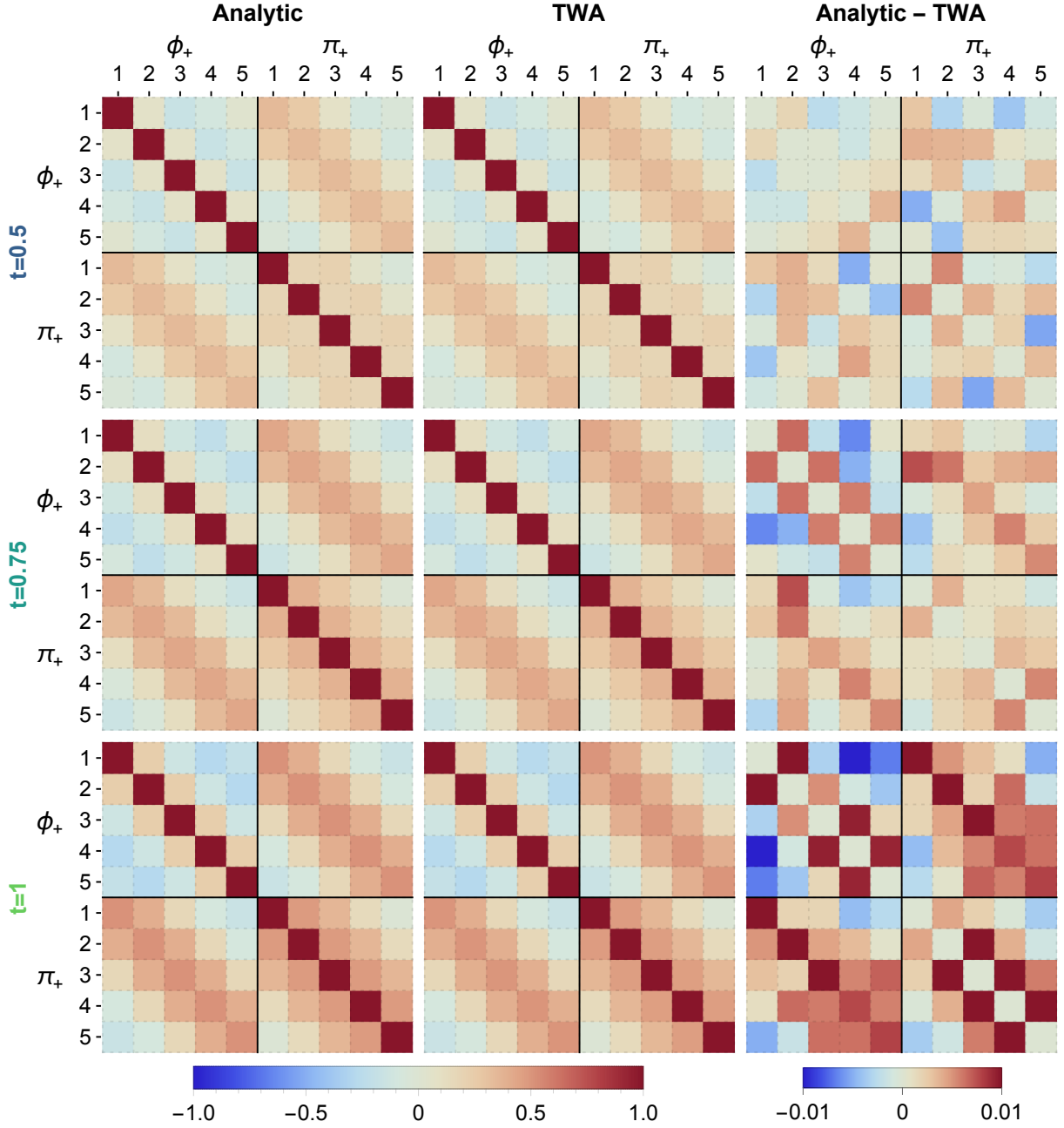


Figure 8. Time evolution of the correlation matrices (128) generated by the analytic Hamiltonian (57) (left column) and the TWA simulation (middle column) together with their difference (right column).

terms, meaning that observables that contain higher-order fluctuations are required to fully characterize entropies.

Since the initial polar state ($t = 0$) in (16) is uncorrelated, its correlation matrix is the identity matrix. During the evolution, correlations between different wells are building up through the transport of atoms permitted by the tunnel Hamiltonian given in Eq. (8). Simultaneously, correlations between ϕ_+^j and π_+^j within one well are generated through the internal dynamics in Eq. (7), leading to the complex structures visible in Figure 8. Remarkably,

the analytical model captures most of the features visible in the normalized second-order correlators for the three times considered, with the absolute value of deviations not exceeding 1% (see right column).

Note, however, that this analysis does not permit any statements about the agreement of higher-order moments. These can in principle deviate from a Gaussian prediction, due to the presence of fourth-order terms in Eq. (6). Crucially, we therefore do not assume the distribution to obey a specific functional form in our subsequent analysis

of sample data generated by the means of TWA with regards to entropic quantities.

The good agreement between the two models leads us to conclude that the analytical model is a justified approximation in regimes where the contribution of fourth-order terms to the full Hamiltonian Eq. (6) is negligible. These regimes are characterized by small values of c_0 , i.e. Lithium-7, and large values of q , such that the undepleted pump approximation (see section III A) is justified. Furthermore, we expect the approximations made by the analytical model to be valid, especially for early times, which is why we restrict ourselves to times up to $t = 1$. An extended discussion for later times can be found in section VII C and in [64].

B. Observation of the area law

Having gained a qualitative understanding of the dynamics generated by the Hamiltonian (6), we now turn our attention to the dynamics of entropic quantities associated with the quantum state's phase space and measurement distributions. In this context, we want to discuss our main result Figure 9, which demonstrates the numerical observation of the area law from experimentally extractable quantities without assuming a specific functional form of the quantum state. Crucially, we only require 10^4 samples to estimate entropies of up to ten-dimensional distributions, which we deem experimentally feasible.

In the following, we consider the three distributions of interest, namely

- the Wigner W -distribution, introduced in Eq. (38) which has mainly theoretical interest due to its connection to quantum entropies in the early-time regime and its restricted accessibility,
- its marginal distributions, introduced in Eq. (40), motivated by the fact that they allow for direct experimental extraction as described in section II E 1,
- and finally the Husimi Q -distribution, introduced in Eq. (44) and directly measurable in experiments using the readout techniques explored in [40, 41] and described in section II E 2,

for subsystem A of varying size M , such that B contains $5 - M$ wells (see Figure 5). For the aforementioned distributions, we compute the absolute and subtracted entropies as well as the mutual information for the evolution times $t = 0.5$, $t = 0.75$, and $t = 1$. We compare the numerical TWA data (open markers) to the analytical data (solid markers) and fit (finite-size) area laws (solid lines).

We first discuss the absolute entropies shown in the first row of Figure 9, i.e. a), b), c). All entropies show similar linear behavior and the area law is masked behind the leading order terms. These leading order contributions are subtracted in Figure 9 d), e) and f), as discussed in section V B 2, unveiling the area law in form of the typical

logarithmic growth in all three types of distributions. Finally, the finite-size area law becomes apparent in all three classical mutual informations, see Figure 9 g) - i).

At this point, let us emphasize the challenges associated with the task of estimating the subtracted entropies: We are interested in the estimation of entropies of up to ten-dimensional distributions at an accuracy that is between one and two orders of magnitude higher than the signal itself (consider e.g. $\Delta S(\mathcal{W}_+^A)/S(\mathcal{W}_+^A) \approx 0.03$ for $M = 5$ wells at $t = 0.5$). From this perspective, we consider the agreement between analytical and TWA results as remarkable.

This concludes the presentation of our main results. In the following, we carry out ablative studies to demonstrate the genericness of our experimental proposal. In particular, we explore later times (section VII C), where the distributions show stronger non-Gaussian features, consider imperfectly prepared, thermal initial states (section VII D) and investigate the influence of boundary effects by shifting the position of the five wells within the total system of $N = 20$ wells (section VII E). We also reduce the total system to $N = 5$ wells to hunt the *finite-size* area law for the subtracted entropies (section VII F). Finally, we aim to shed light on the sample complexity associated with these tasks when utilizing the k NN-estimator (section VII G) before considering the effect of periodic boundary conditions (section VII H).

C. Distributions: Gaussian versus non-Gaussian

To test the influence of non-Gaussian features of the distribution we consider the parameters $q = 2$ and $J = 0.5$ at later times, namely $t = 2$, $t = 3$, and $t = 4$. We expect stronger non-Gaussian features to emerge for smaller values of q , as we cross the polar to easy-plane ferromagnet phase transition [68], due to the (less detuned) squeezing generated by c_1 . Similar to Figure 8, we plot the discrepancies between the correlation matrices obtained with the analytical model and TWA for the three different times. Figure 10 shows the differences now become more pronounced, growing as large as 0.1, more than an order of magnitude larger compared to Figure 8. We particularly expect the analytical model to mainly give reliable results for the early time dynamics, so a disparity between the two is expected.

To gain more intuition about the characteristics of the distribution at those times, we show samples (ϕ_+^1, π_+^1) of the Wigner W -distribution of the first well in the second row of Figure 10, with all other degrees of freedom integrated out. While the local distribution at $t = 2$ still seems relatively Gaussian, an analysis of its 4th-moments reveals that Gaussianity is already strongly violated, with Isserlis' theorem [118] showing relative deviations as large as 4%. For $t = 3$ ($t = 4$) these deviations grow as large as 17% (60%). Hence, also the entropic quantities show stronger deviations. In the case of the Wigner W -distribution in Figure 10 g) substantial differences occur

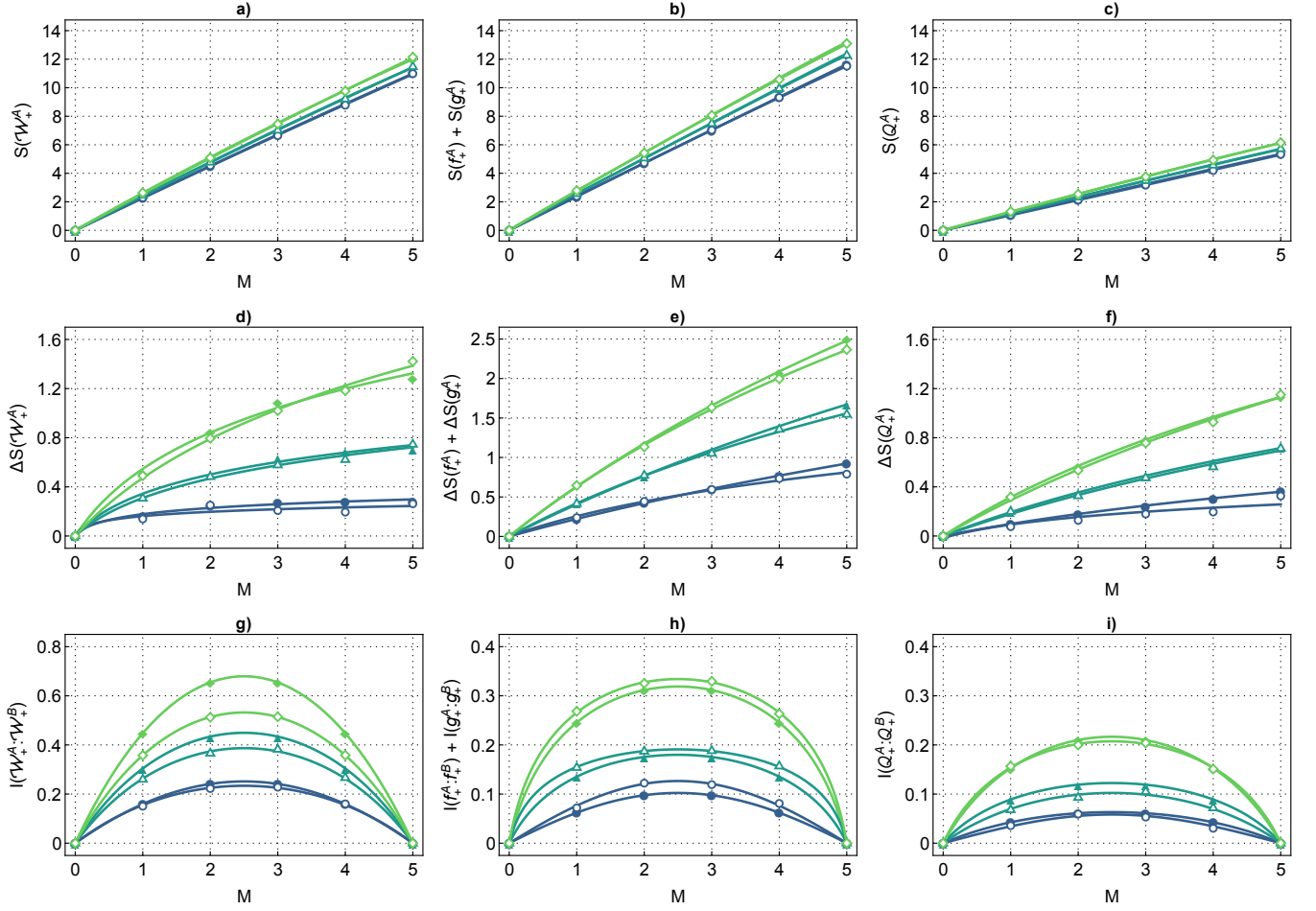


Figure 9. The upper, middle, and lower rows correspond to absolute entropies, subtracted entropies, and mutual information, respectively, while the left, middle, and right columns show Wigner, marginal, and Wehrl quantities. Solid (open) markers correspond to analytical (TWA) data, while solid lines depict fitted curves. The fit function for **d)** - **f)** is the area law (85), with an additional extensive term for **a)** - **c)**. For **g)** - **i)**, we fit the finite-size area law (87).

for more wells and late times, as the high dimensionality as well as the strong non-Gaussian features make it difficult for the estimator to reliably estimate the entropies using the 10^4 available samples. The mutual information shown in [Figure 10 h\)](#) and [i\)](#), show strong deviations between TWA and the analytical model, although a quantitative comparison does not seem meaningful given the strong non-Gaussian features shown in [Figure 10 d\)](#) - [f\)](#). A more informative observation is that the estimates of the mutual informations still change significantly upon increasing the sample size by one order of magnitude, albeit not changing their functional form. In the case of the Wigner mutual information, this underestimation is around 15% with respect to the currently given data, while it is around 7% for the Wehrl mutual information. We attribute this difference to the increased smoothness of the Husimi Q -distribution compared to the Wigner W -distribution.

D. Initial state: vacuum versus thermal

Up to now, our discussion has not been concerned with experimental imperfections such as thermal noise. We here want to address this issue by considering thermal initial states that undergo the unitary evolution. These initial states are characterized by a temperature scale β that we choose to be on the order of a tenth of the system size, such that $\beta = 2$. The thermal fluctuations alter the computations of the dynamics by adding thermal noise to the quantum one-half, see Eqs. (53) and (75).

As expected, the subtracted Wigner entropy $\Delta S(\mathcal{W}_+^A)$ rather features an extensive volume law than an area law, as can be seen in [Figure 11 a\)](#). When computing both the Wigner and Wehrl mutual information in [Figure 11 b\)](#) and [c\)](#), we observe the finite-size area law in both instances, just as in [Figure 9](#). The Wigner mutual information shows the same quantitative behavior observed in [Figure 9 g\)](#), since the contributions due to the initial thermal noise cancel out. In contrast, the Wehrl mutual

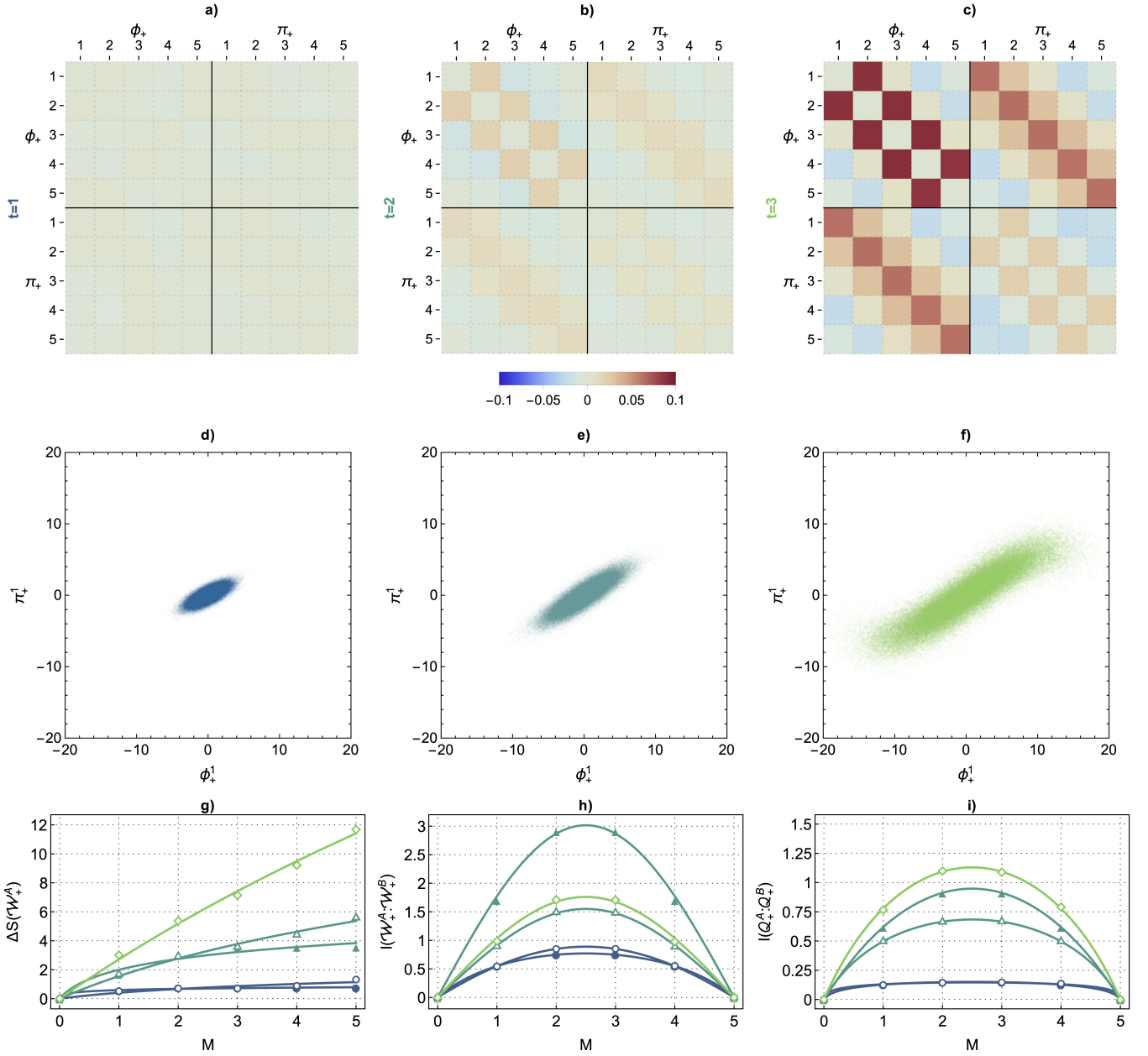


Figure 10. The differences between the analytical model and TWA in the correlation matrices are shown in **a)**, **b)** and **c)** for the three times $t = 1$, $t = 2$ and $t = 3$, respectively, and $q = 2$, $J = 0.5$. The discrepancies grow as large as 0.1, such that we can no longer meaningfully speak about agreement between the methods. The non-Gaussian form of the underlying phase-space distributions is illustrated in **d)** - **f)** in terms of sample points (ϕ_+^1, π_+^1) of the first well. The corresponding subtracted Wigner entropy, Wigner mutual information, and Wehrl mutual information are shown in **g)** - **i)**, for the analytical model (solid markers, only up to $t = 2$) as well as TWA (open markers). We observe the area law for the former and the finite-size area law for the two latter quantities, also in the non-Gaussian regime.

information, is increased compared to Figure 9 i). We attribute this to the inequality (110) becoming tight in the infinite temperature limit [102] (this is also evident from (67): the additional term $(1/2)\mathbb{1}$ becomes irrelevant in this case). Importantly, the finite-size area law (87) describes the measurable and noisy data well, rendering the proposal robust against thermal noise. Hence, the considered measurement distributions indeed describe a

suited setup to observe the area law experimentally.

E. Subsystem position: center vs. outward

In this subsection, we will choose the position of the subsystem within the bigger system to be wells $M = 1 - 5$ and $M = 3 - 7$, to investigate how the boundary affects

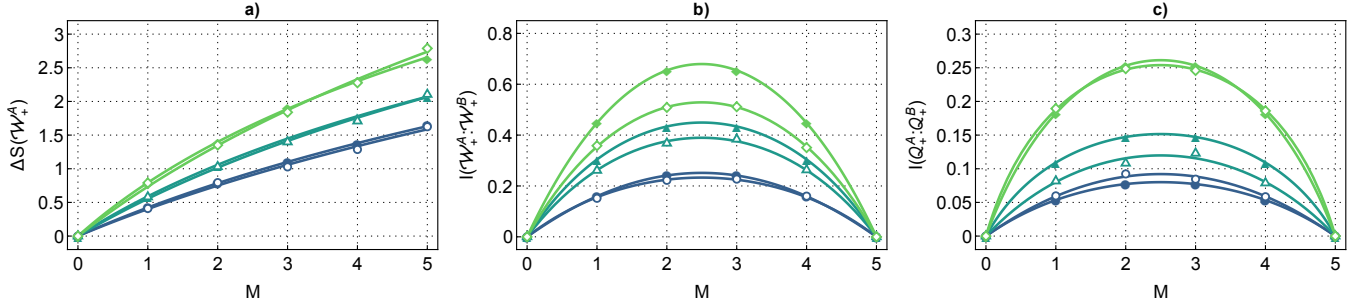


Figure 11. Entropic curves for a thermal initial state with $T = 1/2$. The fits correspond to the area law in **a)** and the finite-size area law in **b)** and **c)**, respectively.

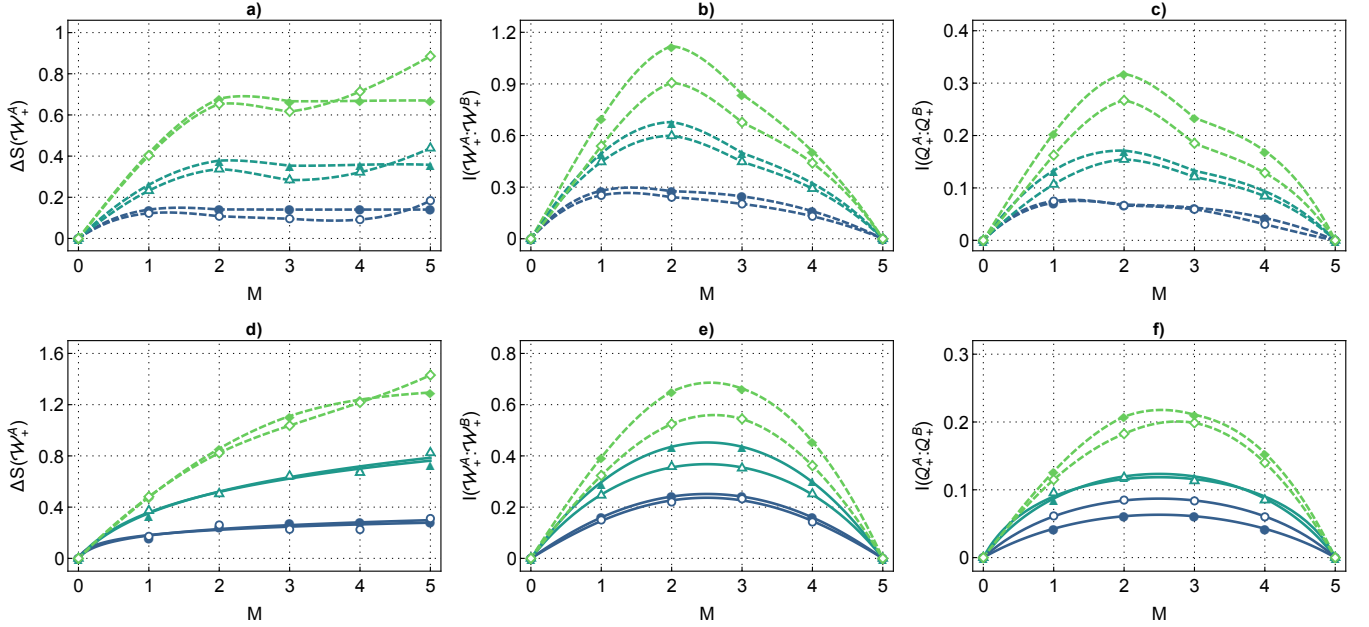


Figure 12. Entropic curves for the wells $M = 1 - 5$ (upper row) and $M = 3 - 7$ (lower row), i.e. when picking five wells out of $N = 20$ and partitioning these five wells into two subsystems as shown in [Figure 5](#). Near the boundary, the area laws are strongly distorted by boundary effects, illustrated by the interpolated curves in **a)** - **c)**. Further away from the boundary, boundary effects become only relevant when the boundary information had enough time to affect local quantities (note that all curves at $t = 1$ are interpolations), see e.g. the asymmetry of the mutual informations in **e)** and **f)**.

the area law.

After computing the subtracted Wigner entropy as well as the Wigner and Wehrl mutual informations for the two subsystem positions, we find the data to be well described by the area law for those instances where boundary effects are irrelevant. The upper row in [Figure 12](#) shows that the subsystem 1-5 is strongly affected by the boundary on the left, and no area law can be observed. In contrast, the data shown in the lower for the subsystem 3-7 is captured well by the fitted curves, albeit the mutual informations are showing some slight asymmetry for $t = 1.0$, i.e. when the boundary information has had sufficient time to propagate to the considered subsystem.

F. System size: large vs. small

In contrast to the previous discussion, we now restrict the total system to $N = 5$ wells such that the evolution within the considered region is unitary. The small system size implies that the dynamics is dominated by boundary effects and the regular logarithmic area law is replaced by the finite-size area law also in the case of the subtracted classical entropies. More precisely, since the density matrix of the five-well system remains pure throughout the evolution, the subtracted Wigner entropy returns to the value zero once the entire system is considered, as shown in [Figure 13 a\)](#). The reason why this is not captured with the TWA data is due to an insufficient amount of samples. This is in contrast with local distributions, which are likely to be smooth (since local states are likely to

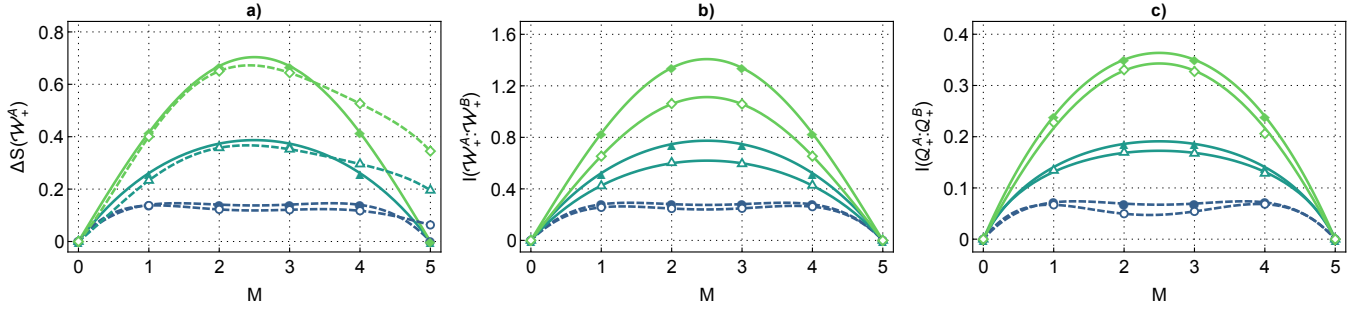


Figure 13. When the total system size equals the subsystem size, i.e. when $N = 5$, the area law for the subtracted classical entropies translates into a finite-size area law reflecting the purity of the global state. We remark the underperformance of the entropy estimator for the subtracted Wigner entropies in **a)** and that the build-up of the finite-size area law only occurs after a transient phase around $t = 0.5$. In both cases, the dashed curves correspond to interpolations, which we show for illustrative purposes.

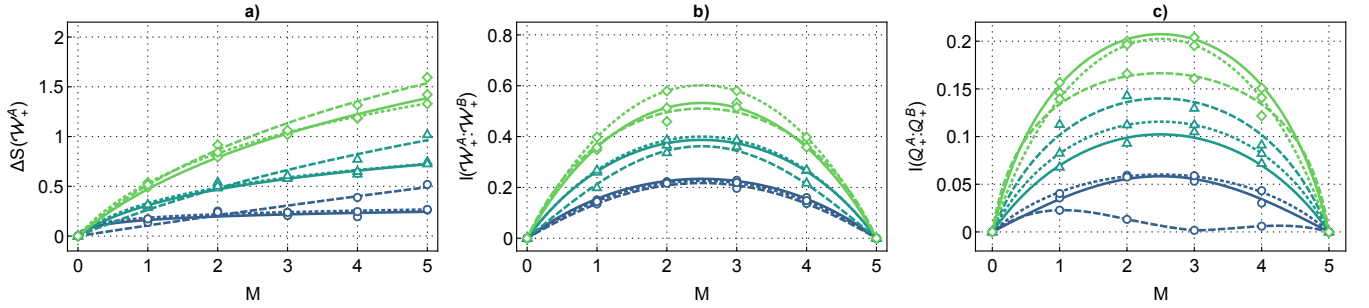


Figure 14. The feasibility of the entropy estimator is checked by varying sample size, where dashed, solid, and dotted lines correspond to 10^3 , 10^4 , and 10^5 samples. All curves are fits, the only exception being the 10^3 curve at $t = 0.5$ in **c)**. The convergence between 10^4 and 10^5 samples becomes visible especially for the subtracted Wigner entropy in **a)** and the Wehrl mutual information in **c)**.

be mixed), and hence require fewer samples to be reconstructed adequately. Upon increasing the sample number to 10^5 , the subtracted Wigner entropy is much closer to zero for $M = 5$ wells (0.00, 0.06, and 0.18 instead of 0.06, 0.20, and 0.34 for $t = 0.50$, $t = 0.75$ and $t = 1.00$). The Wigner and Wehrl mutual information both feature the area law, but only starting at later times $t \geq 0.75$.

G. Sample size: small versus large

Thus far, we estimated entropies of the distributions generated through TWA using 10^4 samples, without questioning whether the k NN estimator had converged for the given sample set. Similar to the benchmarks presented in [section VIA](#), we want to better understand the sample complexity of the entropy estimation using the k NN-estimator. To this end, we carry out the estimations of the subtracted Wigner entropy, the Wigner mutual information and the Wehrl mutual information in [Figure 14 a\), b\) and c\)](#), respectively, for 10^3 , 10^4 and 10^5 samples, indicated through dashed, solid and dotted lines respectively.

The subtracted Wigner entropy seems to be converged

for all times when using 10^4 samples, as no significant differences exist to the data points that were obtained using 10^5 samples. The biggest visual differences exist for the Wigner mutual information for late times, which however mainly stem from the different scale between [Figure 14 a\)](#) and **b)**. For the experimentally relevant Wehrl mutual information, we observe convergence for 10^4 samples for almost all times, with an even further reduced scale compared to **b)**, rendering a potential experimental implementation feasible.

H. Boundary conditions: open versus periodic

While experimental setups will typically use open boundary conditions, a question of theoretical interest is whether the presented framework is sensitive to the type of boundaries that are employed. Therefore, we modify the tunnel Hamiltonian (8) to also allow atoms to jump from the last to the first well and show the same quantities as previously in [Figure 15](#). Also in this setting, we find that the area law persists. More precisely, the curves resemble those obtained for open boundary conditions ([Figure 9](#)), since the the considered subsystem is

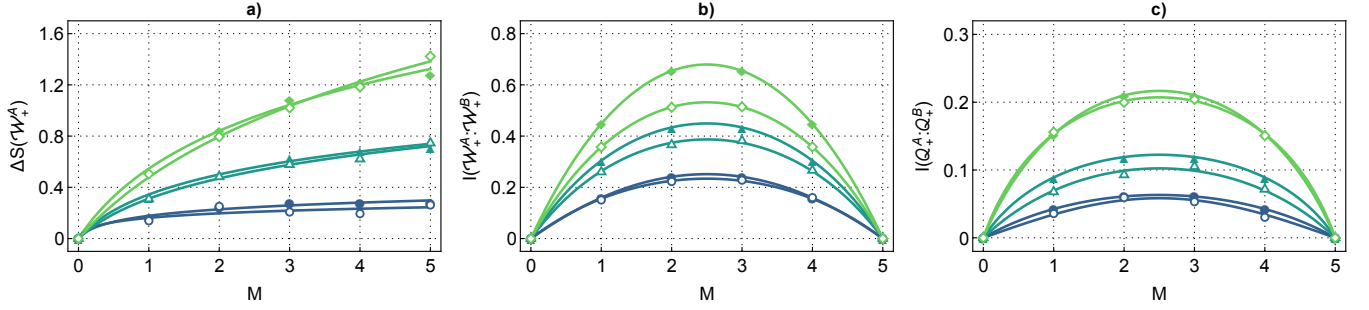


Figure 15. Entropic curves for periodic boundary conditions. As boundary effects are negligible for the subsystem under consideration, all curves agree with those in Figure 9 to a good extent.

sufficiently far away from the total system’s boundaries.

VIII. DISCUSSION

We have demonstrated the experimental feasibility of observing the area law in the subtracted classical entropies and mutual informations of measurable distributions in a multi-well spin-1 Bose-Einstein condensate using numerical techniques. As a testbed, we chose a system consisting of 20 wells that feature internal dynamics as well as tunneling among wells and read out two observables that obey a canonical commutation relation for early times. Focusing on a subregion of five wells, we find (finite-size) area laws in the Wigner distribution, its marginals as well as the Husimi-Q distribution, where the latter two lend themselves to experimental investigation using independent subsequent measurements [119] and a readout technique proposed in [40, 41], respectively. We have shown that all types of subtracted entropies as well as mutual informations are quantities worth studying in order to observe the area law and substantiated its robustness with respect to changing various parameters of the experimental setup from section VII C to section VII H. The discussed area laws are observed dynamically, meaning that the underlying distributions are generated through quenches rather than ground-state preparations, simplifying the experimental requirements. The proposed procedure makes no assumptions regarding the functional form of the distribution under scrutiny, by estimating its differential entropy using a suited k -nearest-neighbor estimator to locally estimate the density at each sample point. In section VII G we demonstrated that for the estimator to converge we require on the order of 10^4 samples, which we deem experimentally feasible.

We simulated the system’s dynamics using an analytical model derived from the full Hamiltonian and TWA. While the analytical model only features quadratic terms corresponding to an integrable system, the dynamics predicted by the TWA are more intricate and lead to late-time equilibration due to finite particle number and the non-integrable nature of the mean-field equations of motion. We accordingly found agreement between the two meth-

ods at early times and found pronounced differences at times $t \geq 2$. Particularly for parameter choices that feature stronger non-Gaussian features, such as a lower value of q violating the undepleted pump approximation more strongly, we observed greater disparities between TWA and the analytical model.

In the non-Gaussian regime, we found the k NN-estimator to require more samples to reach convergence, hinting at the difficulty of estimating entropies of ten-dimensional distributions while being agnostic to their functional form. As the k NN estimator is asymptotically unbiased, it is always possible to increase or decrease the sample size, to check whether the estimator has converged. We wish to state that while the k NN-estimator has a great appeal due to the guarantee of converging to the true entropy using sufficiently many samples, the estimation of differential entropies is a long-standing challenge and various ideas have been put forward to tackle the task more efficiently [49–63, 103, 105–117]. Of particular interest is the variational approach, which consists of adapting the parameters of a defined function by maximizing the likelihood of the observed samples. While this approach can work well, it is i) generally difficult to build a class of functions that includes the observed density distribution and ii) challenging to converge to the global minimum, rendering the approach uncontrolled. However, future research may aim to build functional forms that are directed at estimating the densities generated by specific classes of Hamiltonians, for which one may hope to arrive at more efficient estimators. Nevertheless, this would come at the expense of losing generality, which may be a significant drawback when applying the methodology to an experiment that suffers from undesired noise effects that one cannot account for analytically.

Having established the presence of quantum features in entropies of classical distributions, one can envision various ways forward. It will be interesting to further investigate which other platforms lend themselves to the herein-discussed approach and obey the area law. Particularly systems that are fully described by the set of observables being read out and systems with other types of degrees of freedom that are constrained by algebras different from canonical commutation relations are of

interest. For instance, one may consider the hitherto discussed spin-1 BEC in the pseudo-spin 1/2 configuration, in which case three spin operators forming an $\mathfrak{su}(2)$ algebra constitute a complete representation of the system's state with the phase space distribution being represented on a sphere [76, 119]. But also non-atomic setups, such as integrated photonic waveguides may present promising candidate systems. We emphasize that the herein presented approach and analyses are applicable to any setup capable of preparing or dynamically generating an area law.

It will be equally interesting to see which other quantum phenomena can be accessed using classical entropies of measurement distributions. In this context, let us also highlight the two accompanying publications that study thermalization and the long time limit of the herein discussed system through the lens of classical entropies [64] and a work that is concerned with showing the generality of the area law for classical entropies [24].

ACKNOWLEDGEMENTS

The authors want to thank Thomas Gasenzer and Ido Siovitz for helpful discussions during the development of the manuscript. T. H. is supported by the European Union under project ShoQC within the ERA-NET Cofund

in Quantum Technologies (QuantERA) program, as well as by the F.R.S.- FNRS under project CHEQS within the Excellence of Science (EOS) program. This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC2181/1-390900948 (the Heidelberg STRUCTURES Excellence Cluster) and within the Collaborative Research Center SFB1225 (ISOQUANT). This work was partially financed by the Baden-Württemberg Stiftung gGmbH. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS [120] at Jülich Supercomputing Centre (JSC).

Appendix A: Operator identities for the undepleted pump regime

We list a few identities to reexpress the Hamiltonian (20) in terms of the relative modes as well as the canonical variables. For the step from the local modes to the relative modes we invert (23), which yields

$$\mathbf{a}_1^j = \frac{1}{\sqrt{2}} \left(\mathbf{a}_+^j + \mathbf{a}_-^j \right), \quad \mathbf{a}_{-1}^j = \frac{1}{\sqrt{2}} \left(\mathbf{a}_+^j - \mathbf{a}_-^j \right), \quad (\text{A1})$$

leading to the identities

$$\begin{aligned} \mathbf{a}_1^{j\dagger} \mathbf{a}_1^{j\dagger} \mathbf{a}_1^j \mathbf{a}_1^j + \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^j \mathbf{a}_{-1}^j + 2\mathbf{N}_1^j \mathbf{N}_{-1}^j &= \mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j\dagger} \mathbf{a}_+^j \mathbf{a}_+^j + \mathbf{a}_-^{j\dagger} \mathbf{a}_-^{j\dagger} \mathbf{a}_-^j \mathbf{a}_-^j + 2\mathbf{N}_+^j \mathbf{N}_-^j, \\ \mathbf{a}_1^{j\dagger} \mathbf{a}_1^{j\dagger} + \mathbf{a}_1^j \mathbf{a}_1^j + \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^{j\dagger} + \mathbf{a}_{-1}^j \mathbf{a}_{-1}^j &= \mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j\dagger} + \mathbf{a}_+^j \mathbf{a}_+^j + \mathbf{a}_-^{j\dagger} \mathbf{a}_-^{j\dagger} + \mathbf{a}_-^j \mathbf{a}_-^j, \\ \mathbf{a}_1^j \mathbf{a}_{-1}^j + \mathbf{a}_1^{j\dagger} \mathbf{a}_{-1}^{j\dagger} &= \frac{1}{2} \left(\mathbf{a}_+^j \mathbf{a}_+^j + \mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j\dagger} - \mathbf{a}_-^j \mathbf{a}_-^j - \mathbf{a}_-^{j\dagger} \mathbf{a}_-^{j\dagger} \right), \\ \mathbf{N}_1^j - \mathbf{N}_{-1}^j &= \mathbf{a}_+^{j\dagger} \mathbf{a}_-^j + \mathbf{a}_-^{j\dagger} \mathbf{a}_+^j, \\ \mathbf{N}_1^j + \mathbf{N}_{-1}^j &= \mathbf{N}_+^j + \mathbf{N}_-^j, \\ \mathbf{a}_1^{j\dagger} \mathbf{a}_1^{j+1} + \mathbf{a}_1^{(j+1)\dagger} \mathbf{a}_1^j + \mathbf{a}_{-1}^{j\dagger} \mathbf{a}_{-1}^{j+1} + \mathbf{a}_{-1}^{(j+1)\dagger} \mathbf{a}_{-1}^j &= \mathbf{a}_+^{j\dagger} \mathbf{a}_+^{j+1} + \mathbf{a}_+^{(j+1)\dagger} \mathbf{a}_+^j + \mathbf{a}_-^{j\dagger} \mathbf{a}_-^{j+1} + \mathbf{a}_-^{(j+1)\dagger} \mathbf{a}_-^j. \end{aligned} \quad (\text{A2})$$

For the step from relative modes operators to canonical operators we invert (25), which gives

$$\mathbf{a}_\pm^j = \frac{1}{\sqrt{2}} \left(\phi_\pm^j + i\pi_\pm^j \right), \quad (\text{A3})$$

implying the identities

$$\begin{aligned} \mathbf{N}_\pm^j &= \frac{1}{2} \left[\left(\phi_\pm^j \right)^2 + \left(\pi_\pm^j \right)^2 - 1 \right], \\ \mathbf{a}_\pm^{j\dagger} \mathbf{a}_\pm^{j\dagger} + \mathbf{a}_\pm^j \mathbf{a}_\pm^j &= \left(\phi_\pm^j \right)^2 - \left(\pi_\pm^j \right)^2, \\ \mathbf{a}_\pm^{j\dagger} \mathbf{a}_\pm^{j+1} + \mathbf{a}_\pm^{(j+1)\dagger} \mathbf{a}_\pm^j &= \phi_\pm^j \phi_\pm^{j+1} + \pi_\pm^j \pi_\pm^{j+1}. \end{aligned} \quad (\text{A4})$$

[1] J. D. Bekenstein, Black holes and the second law, *Lett. Nuovo Cimento* **4**, 737 (1972).

[2] J. D. Bekenstein, Black Holes and Entropy, *Phys. Rev.*

- D **7**, 2333 (1973).
- [3] S. W. Hawking, Particle creation by black holes, *Comm. Math. Phys.* **43**, 199 (1975).
 - [4] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Quantum source of entropy for black holes, *Phys. Rev. D* **34**, 373 (1986).
 - [5] M. Srednicki, Entropy and area, *Phys. Rev. Lett.* **71**, 666 (1993).
 - [6] C. Callan and F. Wilczek, On geometric entropy, *Phys. Lett. B* **333**, 55 (1994).
 - [7] P. Calabrese and J. Cardy, Entanglement entropy and quantum field theory, *J. Stat. Mech. Theo. Exp.* **2004**, P06002 (2004).
 - [8] P. Calabrese and J. Cardy, Entanglement Entropy and Quantum Field Theory: A Non-Technical Introduction, *Int. J. Quantum Inf.* **04**, 429 (2006).
 - [9] M. B. Hastings, An area law for one-dimensional quantum systems, *J. Stat. Mech.: Theo. Exp.* **2007**, P08024 (2007).
 - [10] P. Calabrese and J. Cardy, Entanglement entropy and conformal field theory, *J. Phys. A Math. Theo.* **42**, 504005 (2009).
 - [11] H. Casini and M. Huerta, Entanglement entropy in free quantum field theory, *J. Phys. A Math. Theo.* **42**, 504007 (2009).
 - [12] S. Hollands and K. Sanders, *Entanglement Measures and their Properties in Quantum Field Theory* (Springer, 2018).
 - [13] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Entanglement in many-body systems, *Rev. Mod. Phys.* **80**, 517 (2008).
 - [14] I. Peschel and V. Eisler, Reduced density matrices and entanglement entropy in free lattice models, *J. Phys. A: Math. Theo.* **42**, 504003 (2009).
 - [15] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277 (2010).
 - [16] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Entropy Scaling and Simulability by Matrix Product States, *Phys. Rev. Lett.* **100**, 030504 (2008).
 - [17] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
 - [18] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Area laws in quantum systems: Mutual information and correlations, *Phys. Rev. Lett.* **100**, 070502 (2008).
 - [19] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, Measuring entanglement entropy in a quantum many-body system, *Nature* **528**, 77 (2015).
 - [20] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, *Science* **353**, 794 (2016).
 - [21] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, *Science* **364**, 260 (2019).
 - [22] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, Rényi Entropies from Random Quenches in Atomic Hubbard and Spin Models, *Phys. Rev. Lett.* **120**, 050406 (2018).
 - [23] M. Tajik, I. Kukuljan, S. Sotiriadis, B. Rauer, T. Schweigler, F. Cataldini, J. Sabino, F. Møller, P. Schüttelkopf, S.-C. Ji, D. Sels, E. Demler, and J. Schmiedmayer, Verification of the area law of mutual information in a quantum field simulator, *Nat. Phys.* **19**, 1022 (2012).
 - [24] T. Haas, Area laws from classical entropies, See same arXiv posting (2024).
 - [25] E. P. Wigner, On the Quantum Correction For Thermodynamic Equilibrium, *Phys. Rev.* **40**, 749 (1932).
 - [26] K. Husimi, Some formal properties of the density matrix, *Proc. Phys.-Math. Soc. Jap. 3rd Ser.* **22**, 264 (1940).
 - [27] N. D. Cartwright, A non-negative Wigner-type distribution, *Phys. A* **83**, 210 (1976).
 - [28] H.-W. Lee, Theory and application of the quantum phase-space distribution functions, *Phys. Rep.* **259**, 147 (1995).
 - [29] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag Berlin, 2001).
 - [30] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 2013).
 - [31] A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50**, 221 (1978).
 - [32] A. Wehrl, On the relation between classical and quantum-mechanical entropy, *Rep. Math. Phys.* **16**, 353 (1979).
 - [33] E. H. Lieb, Proof of an entropy conjecture of Wehrl, *Commun. Math. Phys.* **62**, 35 (1978).
 - [34] 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).
 - [35] S. Ditsch and T. Haas, Entropic distinguishability of quantum fields in phase space, [arXiv:2307.06128](https://arxiv.org/abs/2307.06128) (2023).
 - [36] N. J. Cerf and T. Haas, Information and majorization theory for fermionic phase-space distributions, [arXiv:2401.08523](https://arxiv.org/abs/2401.08523) (2024).
 - [37] S. Floerchinger, T. Haas, and H. Müller-Groeling, Wehrl entropy, entropic uncertainty relations, and entanglement, *Phys. Rev. A* **103**, 062222 (2021).
 - [38] M. Gärttner, T. Haas, and J. Noll, Detecting continuous-variable entanglement in phase space with the Q distribution, *Phys. Rev. A* **108**, 042410 (2023).
 - [39] M. Gärttner, T. Haas, and J. Noll, General Class of Continuous Variable Entanglement Criteria, *Phys. Rev. Lett.* **131**, 150201 (2023).
 - [40] 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).
 - [41] 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).
 - [42] M. J. Collett, R. Loudon, and C. W. Gardiner, Quantum Theory of Optical Homodyne and Heterodyne Detection, *J. Mod. Opt.* **34**, 881 (1987).
 - [43] J. W. Noh, A. Fougères, and L. Mandel, Measurement of the quantum phase by photon counting, *Phys. Rev. Lett.* **67**, 1426 (1991).
 - [44] J. W. Noh, A. Fougères, and L. Mandel, Operational approach to the phase of a quantum field, *Phys. Rev. A* **45**, 424 (1992).
 - [45] S. Stenholm, Simultaneous measurement of conjugate variables, *Ann. Phys.* **218**, 233 (1992).
 - [46] U. Leonhardt and H. Paul, Phase measurement and Q function, *Phys. Rev. A* **47**, R2460 (1993).
 - [47] 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).
- [48] O. Landon-Cardinal, L. C. G. Govia, and A. A. Clerk, Quantitative Tomography for Continuous Variable Quantum Systems, *Phys. Rev. Lett.* **120**, 090501 (2018).
- [49] L. F. Kozachenko and N. N. Leonenko, Sample Estimate of the Entropy of a Random Vector, *Probl. Peredachi Inf.* **23**, 9 (1987).
- [50] L. Györfi and E. C. van der Meulen, Density-free convergence properties of various estimators of entropy, *Comp. Stat. Data Anal.* **5**, 425 (1987).
- [51] H. Joe, Estimation of entropy and other functionals of a multivariate density, *Ann. Inst. Stat. Math.* **41**, 683 (1989).
- [52] P. Hall and S. C. Morton, On the estimation of entropy, *Ann. Inst. Stat. Math.* **45**, 69 (1993).
- [53] L. Birge and P. Massart, Estimation of Integral Functionals of a Density, *Ann. Stat.* **23**, 10.1214/aos/1176324452 (1995).
- [54] J. Beirlant, E. Dudewicz, L. Györfi, and E. Meulen, Nonparametric Entropy Estimation: An Overview, *Int. J. Math. Stat. Sci.* **6**, 17 (1997).
- [55] H. Singh, N. Misra, V. Hnizdo, A. Fedorowicz, and E. Demchuk, Nearest Neighbor Estimates of Entropy, *Am. J. Math. Manag. Sci.* **23**, 301 (2003).
- [56] A. Kraskov, H. Stögbauer, and P. Grassberger, Estimating mutual information, *Phys. Rev. E* **69**, 066138 (2004).
- [57] M. N. Goria, N. N. Leonenko, V. V. Mergel, and P. L. N. Inverardi, A new class of random vector entropy estimators and its applications in testing statistical hypotheses, *J. Nonparametric Stat.* **17**, 277 (2005).
- [58] K. Sricharan, D. Wei, and A. O. Hero, Ensemble Estimators for Multivariate Entropy Estimation, *IEEE Trans. Inf. Theory* **59**, 4374 (2013).
- [59] K. Kandasamy, A. Krishnamurthy, B. Poczos, L. Wasserman, and J. M. Robins, Nonparametric von Mises Estimators for Entropies, Divergences and Mutual Informations, in *Adv. NIPS*, Vol. 28, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (2015).
- [60] S. Singh and B. Poczos, Finite-Sample Analysis of Fixed-k Nearest Neighbor Density Functional Estimators, in *Adv. NIPS*, Vol. 29, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (2016).
- [61] T. B. Berrett, R. J. Samworth, and M. Yuan, Efficient multivariate entropy estimation via k -nearest neighbour distances, *arXiv:1606.00304* (2016).
- [62] W. Gao, S. Oh, and P. Viswanath, Demystifying Fixed k -Nearest Neighbor Information Estimators, *IEEE Trans. Inf. Theory* **64**, 5629 (2018).
- [63] C. Lu and J. Peltonen, Enhancing Nearest Neighbor Based Entropy Estimator for High Dimensional Distributions via Bootstrapping Local Ellipsoid, *Proc. AAAI* **34**, 5013 (2020).
- [64] Y. Deller, M. Gärttner, T. Haas, M. K. Oberthaler, M. Reh, and H. Strobel, Area laws and thermalization from classical entropies in a Bose-Einstein condensate, Companion paper appearing in the same arXiv posting (2024).
- [65] Y. Kawaguchi and M. Ueda, Spinor Bose-Einstein condensates, *Phys. Rep.* **520**, 253 (2012).
- [66] S. Huh, K. Kim, K. Kwon, and J.-y. Choi, Observation of a strongly ferromagnetic spinor Bose-Einstein condensate, *Phys. Rev. Res.* **2**, 033471 (2020).
- [67] 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).
- [68] D. M. Stamper-Kurn and M. Ueda, Spinor Bose gases: Symmetries, magnetism, and quantum dynamics, *Rev. Mod. Phys.* **85**, 1191 (2013).
- [69] 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).
- [70] P. Kunkel, *Splitting a Bose-Einstein condensate enables EPR steering and simultaneous readout of noncommuting observables*, *Ph.D. thesis*, Ruprecht-Karls-Universität Heidelberg (2019).
- [71] J. Schwinger, On Angular Momentum, Unpublished Report **Report Number NYO-3071** (1952).
- [72] S. K. Kim, Theorems on the Jordan-Schwinger representations of Lie algebras, *J. Math. Phys.* **28**, 2540 (1987).
- [73] C. Gross, H. Strobel, E. Nicklas, T. Zibold, N. Bar-Gill, G. Kurizki, and M. K. Oberthaler, Atomic homodyne detection of continuous-variable entangled twin-atom states, *Nature* **480**, 219 (2011).
- [74] 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).
- [75] A. Serafini, *Quantum Continuous Variables* (CRC Press, 2017).
- [76] W.-M. Zhang, D. H. Feng, and R. Gilmore, Coherent states: Theory and some applications, *Rev. Mod. Phys.* **62**, 867 (1990).
- [77] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States, 2nd ed.* (John Wiley and Sons, 2017).
- [78] A. Kenfack and K. Życzkowski, Negativity of the Wigner function as an indicator of non-classicality, *J. Opt., B Quantum Semiclass. Opt.* **6**, 396 (2004).
- [79] S. L. Braunstein, Homodyne statistics, *Phys. Rev. A* **42**, 474 (1990).
- [80] U. Leonhardt and H. Paul, Measuring the quantum state of light, *Prog. Quantum. Electron.* **19**, 89 (1995).
- [81] A. Polkovnikov, Phase space representation of quantum dynamics, *Ann. Phys.* **325**, 1790 (2010).
- [82] M. Oliva, D. Kakofengitis, and O. Steuernagel, Anharmonic quantum mechanical systems do not feature phase space trajectories, *Phys. A: Stat. Mech. Appl.* **502**, 201 (2018).
- [83] P. Blakie, A. Bradley, M. Davis, R. Ballagh, and C. Gardiner, Dynamics and statistical mechanics of ultra-cold Bose gases using c-field techniques, *Adv. Phys.* **57**, 363 (2008).
- [84] H. P. Robertson, The Uncertainty Principle, *Phys. Rev.* **34**, 163 (1929).
- [85] H. P. Robertson, A general formulation of the uncertainty principle and its classical interpretation, *Phys. Rev.* **35**, 667 (1930).
- [86] E. Schrödinger, Zum Heisenbergschen Unschärfepinzip, *Sitzungsberichte der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse* **14**, 296 (1930).
- [87] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955).

- [88] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
- [89] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
- [90] Z. Van Herstraeten and N. J. Cerf, Quantum Wigner entropy, *Phys. Rev. A* **104**, 042211 (2021).
- [91] Z. Van Herstraeten, M. G. Jabbour, and N. J. Cerf, Continuous majorization in quantum phase space, *Quantum* **7**, 1021 (2023).
- [92] Z. Van Herstraeten, *Majorization theoretical approach to quantum uncertainty*, *Ph.D. thesis*, Université libre de Bruxelles (2021).
- [93] H. Everett, 'Relative State' Formulation of Quantum Mechanics, *Rev. Mod. Phys.* **29**, 454 (1957).
- [94] I. I. Hirschman, A Note on Entropy, *Am. J. Math.* **79**, 152 (1957).
- [95] W. Beckner, Inequalities in Fourier Analysis, *Ann. Math.* **102**, 159 (1975).
- [96] I. Białynicki-Birula and J. Mycielski, Uncertainty relations for information entropy in wave mechanics, *Commun. Math. Phys.* **44**, 129 (1975).
- [97] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, *Rev. Mod. Phys.* **89**, 015002 (2017).
- [98] A. Hertz and N. J. Cerf, Continuous-variable entropic uncertainty relations, *J. Phys. A Math. Theor.* **52**, 173001 (2019).
- [99] G. Adesso, D. Girolami, and A. Serafini, Measuring Gaussian Quantum Information and Correlations Using the Rényi Entropy of Order 2, *Phys. Rev. Lett.* **109**, 190502 (2012).
- [100] G. Manfredi and M. R. Feix, Entropy and Wigner functions, *Phys. Rev. E* **62**, 4665 (2000).
- [101] J. J. Włodarz, Entropy and Wigner Distribution Functions Revisited, *Int. J. Theor. Phys.* **42**, 1075 (2003).
- [102] E. H. Lieb and R. Seiringer, Stronger subadditivity of entropy, *Phys. Rev. A* **71**, 062329 (2005).
- [103] G. V. Steeg, Non-parametric Entropy Estimation Toolbox, <https://github.com/gregversteeg/NPEET> (2014).
- [104] S. Nadarajah, A generalized normal distribution, *J. Appl. Stat.* **32**, 685 (2005).
- [105] H. Hino, K. Koshijima, and N. Murata, Non-parametric entropy estimators based on simple linear regression, *Comput. Stat. Data Anal.* **89**, 72 (2015).
- [106] F. Tarasenko, On the evaluation of an unknown probability density function, the direct estimation of the entropy from independent observations of a continuous random variable, and the distribution-free entropy test of goodness-of-fit, *Proc. IEEE* **56**, 2052 (1968).
- [107] O. Vasicek, A Test for Normality Based on Sample Entropy, *J. R. Stat. Soc., B: Stat. Methodol.* **38**, 54 (1976).
- [108] H. Hino, K. Wakayama, and N. Murata, Entropy-based sliced inverse regression, *Comput. Stat. Data Anal.* **67**, 105 (2013).
- [109] E. G. Learned-Miller and J. W. F. III, ICA using spacings estimates of entropy, *J. Mach. Learn. Res.* **4**, 1271 (2003).
- [110] P. Comon, Independent component analysis, A new concept?, *Signal Process.* **36**, 287 (1994).
- [111] S. Mannor, D. Peleg, and R. Rubinstein, The cross entropy method for classification, in *Proc. ICML '05* (2005).
- [112] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method* (Springer New York, 2004).
- [113] H. Hino and N. Murata, A Conditional Entropy Minimization Criterion for Dimensionality Reduction and Multiple Kernel Learning, *Neural Computat.* **22**, 2887 (2010).
- [114] H. Hino and N. Murata, Information estimators for weighted observations, *Neural Netw.* **46**, 260 (2013).
- [115] P. Czyż, F. Grabowski, J. E. Vogt, N. Beerenwinkel, and A. Marx, Beyond Normal: On the Evaluation of Mutual Information Estimators, [arXiv:2306.11078](https://arxiv.org/abs/2306.11078) (2023).
- [116] Z. Ao and J. Li, Entropy Estimation via Normalizing Flow, *Proc. AAAI* **36**, 9990 (2022).
- [117] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, Mutual Information Neural Estimation, in *Proceedings of the 35th International Conference on Machine Learning*, Proc. MLR, Vol. 80, edited by J. Dy and A. Krause (2018) pp. 531–540.
- [118] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition* (John Wiley and Sons, 2006).
- [119] H. Strobil, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè, A. Smerzi, and M. K. Oberthaler, Fisher information and entanglement of non-Gaussian spin states, *Science* **345**, 424 (2014).
- [120] J. S. Centre, JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing, JLSRF (2019).