Data-driven discovery of statistically relevant information in quantum simulators

R. Verdel [1,*] V. Vitale [1,2,3] R. K. Panda [1,2,4] E. D. Donkor,[1,2] A. Rodriguez [1,5] S. Lannig,[6]
Y. Deller,[6] H. Strobel [6] M. K. Oberthaler [6] and M. Dalmonte [1,2]

[1]*The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34151 Trieste, Italy*
[2]*SISSA—International School of Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy*
[3]*Université Grenoble Alpes, CNRS, Laboratoire de Physique et Modélisation des Milieux Condensés (LPMMC), Grenoble 38000, France*
[4]*INFN Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy*
[5]*Dipartimento di Matematica e Geoscienze, Universitá degli Studi di Trieste, via Alfonso Valerio 12/1, 34127 Trieste, Italy*
[6]*Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*

Quantum simulators offer powerful means to investigate strongly correlated quantum matter. However, interpreting measurement outcomes in such systems poses significant challenges. Here, we present a theoretical framework for information extraction in synthetic quantum matter, illustrated for the case of a quantum quench in a spinor Bose-Einstein condensate experiment. Employing nonparametric unsupervised learning tools that provide different measures of information content, we demonstrate a theory-agnostic approach to identify dominant degrees of freedom. This enables us to rank operators according to their relevance, akin to effective field theory. To characterize the corresponding effective description, we then explore the intrinsic dimension of data sets as a measure of the complexity of the dynamics. This reveals a simplification of the data structure, which correlates with the emergence of time-dependent universal behavior in the studied system. Our assumption-free approach can be immediately applied in a variety of experimental platforms.

I. INTRODUCTION

Recent remarkable advances in highly controlled synthetic quantum devices have revolutionized the study of strongly correlated systems [1–6]. A key element of many of such platforms is their capacity to produce large data sets of many-body snapshots, for example, via generalized projective measurements of the entire wave function [7]. However, the analysis of such outcome poses in general serious challenges, which typically force us to rely on assumptions for certain quantities, disregarding part of the information content of the generated data—in data science language, a dimensional reduction with an uncontrolled loss of information. A particularly important problem is the identification of the most informative observables to describe such quantum many-body systems—a paramount task at the core of quantum field theory [8,9], that is even more daunting for systems driven out of equilibrium. To address this, one needs to develop methods to process the maximum amount of information in quantum simulator output, which are able to identify relevant features—and thus degrees of freedom—emerging from the underlying physical system, without making any assumption nor uncontrolled dimensional reduction.

In this work, we introduce a theoretical framework for data-driven information discovery in quantum simulation, which is schematically illustrated in Fig. 1. We start by considering collections of independent quantum simulator snapshots, which resolve, for example, the dynamics of a many-body system in space and time [Fig. 1(a)]. Such data sets are characterized using nonparametric unsupervised learning methods [Fig. 1(b)]. Finally, from this system-agnostic and unsupervised description of the data, we infer relevant information for the physical system under study [Fig. 1(c)].

This framework is based on three techniques: (i) spectral entropies calculated from a principal component analysis (PCA) of the data, (ii) the information imbalance between a subset and the full set of measured quantities, and (iii) the intrinsic dimension of the concomitant data manifolds. These tools which quantify—from different angles—the information content and correlations in the data, have found several successful applications in various fields, such as chemical and biomolecular science [10–18], ecology [19], stock market dynamics [20–25], and image analysis [26–29].

To demonstrate the capabilities of our approach we apply it to experimental data of a spinor Bose-Einstein condensate (BEC) [30]: We evaluate the full set of experimentally measured densities without knowledge of the post-processing steps which are necessary in order to infer the relevant spin variables from them. Our main results are as follows. (i) PCA spectral entropies and information imbalance allow for a theory-agnostic determination of the most informative measured observables. The predictive power of these methods is demonstrated by showing that they can also unveil combinations of the measured densities, which are key to describe the spin structure of the system [30–33]. (ii) The behavior of the intrinsic dimension as a function of time, displays a rapid

*rverdel@ictp.it

(a) **Quantum simulator data sets**

(b) **Descriptive data analysis**

Non-parametric unsupervised learning

▷ PCA spectral entropy
▷ Information imbalance
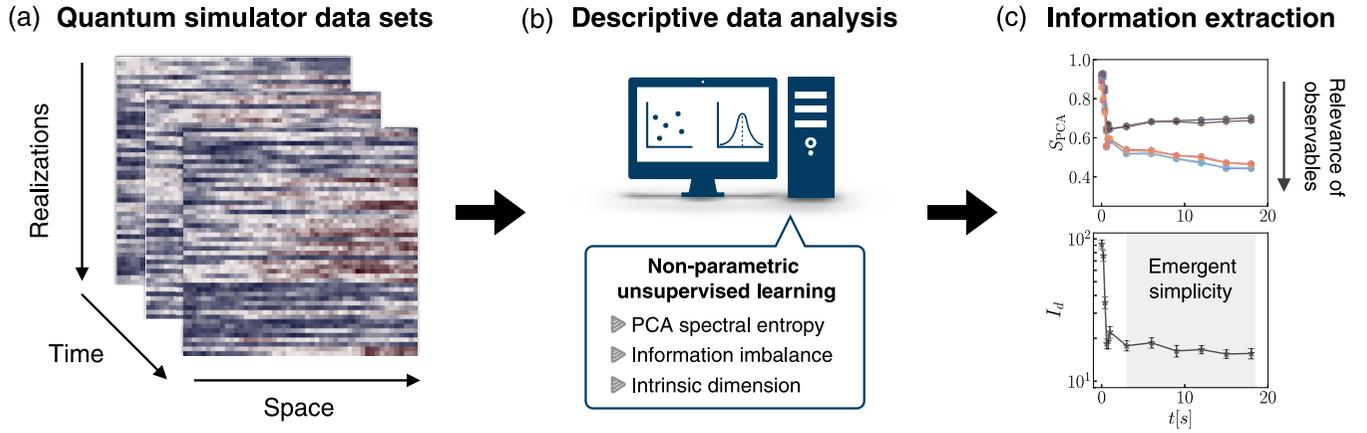▷ Intrinsic dimension

(c) **Information extraction**

FIG. 1. Assumption-free unveiling of relevant information in quantum simulation. (a) We start from snapshots of a many-body system, which are represented as 2D arrays at different times. At a fixed time, each row corresponds to a different realization, while each column is a different data feature, e.g., the atomic density in a given magnetic substate at a given spatial location. (b) Using nonparametric unsupervised learning tools, we perform an exploratory analysis to uncover interesting features of the data, without making any assumptions. (c) From this description, we infer relevant properties of the physical system. (Top) By quantifying the information content and correlations in a data set, the principal component analysis entropy provides a measure of relevance of observables, thence guiding the identification of the most informative degrees of freedom. (Bottom) After a quick fall to relatively small values, the intrinsic dimension of data sets features a long, stable plateau as a function of time (shaded region), providing a lower bound for the timescale after which the dynamics may become simpler and be captured by universal scaling.

decay to significantly smaller values, after which it features very long, stable plateaus. As argued below, this observation is in strong agreement with the formation of spin structure and the emergence of self-similar dynamics [30,31,34–36].

The remaining of this paper is structured as follows. In Sec. II, we introduce the different data science methods used in this work. We describe the spinor BEC experiment and the structure of our data sets in Sec. III. In Sec. IV, we present a theory-agnostic scheme to identify relevant fields, in a systematic and unbiased way, directly from (a limited number of) experimental observations. In Sec. V, we complement our data-driven analysis by characterizing the Kolmogorov complexity of the studied quantum dynamics, and address its capability of recognizing physical information. Conclusions and possible extensions of our work are discussed in Sec. VI.

II. METHODS

Before diving into the central part of our paper, we present an introduction to the data science tools that are employed in this work. We focus on nonparametric methods that are oriented towards extracting information from data, without making (strong) assumptions on the functional form of the probability distribution underlying the data. For the purpose of this exposition, we consider here an abstract data set, structured as a rectangular matrix $\mathbf{X} = \{\vec{X}^i\}_{i=1}^{N_r}$, of dimension $N_r \times p$, where each p-dimensional row vector $\vec{X}^i$ represents a single *realization* (observation) of a set of *p features* (input variables) that are measured in all realizations. This type of structure is ubiquitous to both stochastic Monte Carlo simulations, as well as quantum experiments featuring projective measurements of a large number of degrees of freedom. In the next section, we shall define precisely the concrete data sets that will be analysed with the techniques presented below.

A. Principal component analysis entropy

Principal component analysis is one of the most popular nonparametric methods for unsupervised learning, and has found a vast number of applications, including classical and quantum many-body problems [37–44]. The central idea of PCA is to use an orthogonal transformation to seek for directions along which the data exhibit most variation [45,46]. This is motivated by empirical evidence showing that in many cases such "high-variance" directions capture the relevant information of the data. This problem reduces to diagonalizing the sample covariance matrix $\mathbf{\Sigma} = \mathbf{X}^{\star T}\mathbf{X}^{\star}/(N_r - 1)$, where $\mathbf{X}^{\star}$ is the column-centred data matrix in which the mean value of each column is subtracted from the entries in the column. The solution to the eigenvalue-eigenvector problem $\mathbf{\Sigma}\vec{w}_k = \lambda_k \vec{w}_k$, yields $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_R \geqslant 0$ ($R \leqslant \min\{N_r, p\}$ is the rank of $\mathbf{X}^{\star}$), and the normalized eigenvectors $\vec{w}_k$. The eigenvalues λ_k are the variances of the principal components (PCs), which are determined by the eigenvectors as $\vec{p}_k = \mathbf{X}^{\star}\vec{w}_k$. We note that the procedure above is equivalent to performing a singular value decomposition (SVD) on $\mathbf{X}^{\star}$, and that the semi-definite positiveness of $\mathbf{\Sigma}$ then follows from the fact that the λ_k are proportional to the squared singular values of $\mathbf{X}^{\star}$ [45,46]. A standard measure of importance of a given PC is given by the corresponding normalized eigenvalue $\tilde{\lambda}_k := \lambda_k / \sum_l \lambda_l$, that is, the proportion of total variance that is accounted for by the k-th PC.

The PCA algorithm described above forms the basis of a dimensional reduction scheme, for situations in which the first *few* PCs capture most of the variation present in *all* of the original variables [45,46]. However, determining in a systematic way how many PCs can be disregarded without significant information loss is, in general, a difficult task. Instead of dealing with this aspect of PCA, our goal here is to leverage *all* the information contained in the PC decomposition. To this

end, we introduce an information theoretic-inspired quantity built from the full (normalized) PCA spectrum as follows. Since $\tilde{\lambda}_k \geqslant 0$ for all k, and by construction $\sum_{k=1}^{R} \tilde{\lambda}_k = 1$, we can regard the set of all normalized eigenvalues $\{\tilde{\lambda}_k\}$, as a probability distribution and define the *PCA entropy* in analogy to Shannon's entropy [47], namely,

$$S_{\text{PCA}}(\{\tilde{\lambda}_k\}) := -\frac{1}{\ln(R)} \sum_{k=1}^{R} \tilde{\lambda}_k \ln(\tilde{\lambda}_k), \quad (1)$$

where for convenience we have normalized this entropy by its maximum possible value, $\ln(R)$.

The PCA entropy in Eq. (1), provides a tool to explore the *informational* aspect of the PC decomposition. Indeed, S_{PCA} measures how spread is the information on the principal axes: When no principal direction represents a preferential direction of information accumulation (i.e., $\tilde{\lambda}_k \approx 1/R$, for all k), then $S_{\text{PCA}} \approx 1$, whereas, if information accumulates around some of the principal directions, then $S_{\text{PCA}} < 1$. In fact, in the extreme case in which a single PC explains almost the full variation of the data (i.e., $\tilde{\lambda}_1 \approx 1$), then $S_{\text{PCA}} \approx 0$. Importantly, $S_{\text{PCA}} < 1$ implies the presence of correlations among the original variables. We note, however, that the opposite is in general not true. Consider, for example, a set of two-dimensional data points randomly distributed around a circle: None of the principal directions explains more variation of the data than the other and therefore $S_{\text{PCA}} \approx 1$, in spite of the input variables exhibiting a clear correlation between them.

As mentioned in the introduction, the concept of PCA entropy (and the closely related 'SVD entropy') has found several applications, ranging from unsupervised feature selection methods in bioinformatics [11,12], to schemes to characterize complexity in ecological networks [19] and financial signals [20–25]. However, very little is known about its predictive power in the context of the many-body problem.

B. Information imbalance

Another recently introduced method to quantify information content goes under the name of *information imbalance* [13]. More specifically, this method quantifies the relative information retained when using two distance measures, built with different subsets of data features. In a physical context, this technique can therefore provide an ideal tool to systematically compare—in a fully data-driven manner—different *observables* (subsets of features) and determine which of those can describe better the full space of measured quantities.

The information imbalance method is briefly explained in the following. (The reader is referred to Ref. [13] for a more detailed explanation.) Given two distance measures $D_A(\vec{X}^i, \vec{X}^j)$ and $D_B(\vec{X}^i, \vec{X}^j)$, defined on the same data space, we can rank the neighbors of a point $\vec{X}^i$, by sorting, from smallest to largest, the pairwise distances between such a point and the rest of points using the two distance measures. These rankings are encoded in the so-called rank matrices $R_{A/B}^{ij}$. Hence, $R_A^{ij} = 1$, means that $\vec{X}^j$ is the 1st nearest neighbor of $\vec{X}^i$ in space A, and so on. Here, we restrict ourselves to the case in which the two considered distance measures refer to the Euclidean distance computed with two subsets of features. For example, for data points in two dimensions

with components (x^i, y^i), two possible choices of the distance measures are $D_A(\vec{X}^i, \vec{X}^j) = |x^i - x^j|$ and $D_B(\vec{X}^i, \vec{X}^j) = \sqrt{(x^i - x^j)^2 + (y^i - y^j)^2}$.

The key insight of the information imbalance method is the fact that the full correlation structure between the two metrics under study is essentially captured by the conditional rank distribution $p(R_B|R_A = 1)$, that is, the probability distribution of the ranks R_B^{ij} in space B restricted to pairs of points $(\vec{X}^i, \vec{X}^j)$, such that $R_A^{ij} = 1$ (i.e., nearest-neighbor points according to A). Then, the closer this distribution is to a delta function peaked at 1, the more information about space B is contained in space A. The deviation of $p(R_B|R_A = 1)$ from such a delta function is quantified by the conditional expectation $\langle R_B|R_A = 1 \rangle$ [13], which is used to define the information imbalance from space A to space B, namely,

$$\Delta(A \rightarrow B) = \frac{2}{N_r} \langle R_B|R_A = 1 \rangle \approx \frac{2}{N_r^2} \sum_{i,j:R_A^{ij}=1} R_B^{ij}. \quad (2)$$

In the limit case in which nearest neighbors in A are exactly the same as those in B, we have that $\sum_{j:R_A^{ij}=1} R_B^{ij} = 1$ (for a given i), and hence $\sum_{i,j:R_A^{ij}=1} R_B^{ij} = N_r$. Therefore, the information imbalance in Eq. (2), vanishes as $1/N_r$. A vanishing information imbalance thus indicates that A can fully predict B, in the sense specified above. In the extremely opposite case in which the distance ranks estimated with the two metrics are completely uncorrelated, we have that $\sum_{j:R_A^{ij}=1} R_B^{ij} = \frac{1}{N_r-1} \frac{1}{2} N_r(N_r - 1) = \frac{N_r}{2}$ and hence $\sum_{i,j:R_A^{ij}=1} R_B^{ij} = \frac{N_r^2}{2}$. Therefore, in this case $\Delta(A \rightarrow B) = 1$, and we say that "$A$ is not informative of B."

A scheme for feature selection can then be carried out by measuring the information imbalance from a space of a subset of features to the space of all features. A similar approach has been used, for instance, in Ref. [14] to compare the information that is captured by different atomic descriptors with respect to standard order parameters (and vice versa) in molecular systems.

C. Intrinsic dimension

To complement our data analysis we consider a key concept in the sub-field of manifold learning, namely, the *intrinsic dimension* (I_d). The I_d quantifies the least number of functionally independent parameters needed to describe the data [48–50]. This quantity has a deep connection with information theory, for it serves as a proxy of the Kolmogorov or algorithmic complexity[1] [51–54]. Beyond this informational aspect, the notion of intrinsic dimension plays an important role in unsupervised machine learning, as exemplified in various applications ranging from biomolecular science [15–18], to image analysis [26–29]. Only recently has this concept been employed in physics, more specifically, in the study of critical behavior—in and out of equilibrium—in classical and quantum statistical mechanics systems [44,54–56].

---

[1]Intuitively, the Kolmogorov complexity measures the complexity of a string as the length of the shortest computer program (in a predefined programming language) that outputs the string.

Estimating I_d is, in general, a far-from-trivial task and, in fact, an active field of research [49,50]. Here we use a distance-based I_d estimator that leverages information of only local neighborhoods, namely, the TWO-NN algorithm [15], which we briefly describe in the following. (The reader is referred to Ref. [15] for a more in-depth discussion.) For each point $\vec{X}^i$ in a generic data set, we compute the distance to its first and second nearest neighbors, denoted by r_1^i, r_2^i. Next, we define the ratio $\mu^i := r_2^i/r_1^i$. For data that are locally uniformly distributed on a I_d-dimensional hypersphere, the probability distribution function of μ is given by $f(\mu) = I_d \mu^{-I_d-1}$. The cumulative distribution function $F(\mu)$, obtained upon integration, then satisfies

$$-\ln[1 - F(\mu)] = I_d \ln(\mu), \qquad (3)$$

which is used to estimate I_d through a linear fit of the points $\{(\ln(\mu), -\ln[1 - F_{\text{emp}}(\mu)])\}$, where $F_{\text{emp}}(\mu)$ is the empirical cumulate. In practice, verifying a linear relation between $\ln(\mu)$ and $-\ln[1 - F_{\text{emp}}(\mu)]$), serves also as a good consistency check of the (mild) assumption of local uniformity of the data set.

III. QUANTUM SIMULATION ON A SPINOR BEC AND EXPERIMENTAL DATA SETS

We consider the dynamics realized by a BEC of ^{87}Rb in the $F = 1$ hyperfine spin ground state manifold confined in a quasi-one-dimensional elongated harmonic dipole trap (the data evaluated here are taken from Ref. [30]; see this publication for further details on the experiment). The system is initialized with all atoms in the magnetic substate $m_F = 0$. By instantaneously changing a control parameter we tune spin-changing collision processes into resonance. This procedure implements a quench across a quantum phase transition which brings the system far from equilibrium.

For different times t after the quench we simultaneously infer the two orthogonal spin projections F_x and F_y from the observed densities with spatial resolution along the longitudinal trap direction [7] via

$$F_x = (n_{2,+2} - n_{2,-2})/(n_{2,+2} + n_{2,0} + n_{2,-2}),$$
$$F_y = (n_{1,+1} - n_{1,-1})/(n_{1,+1} + n_{1,0} + n_{1,-1}), \qquad (4)$$

where n_{F,m_F} is the density in the state with hyperfine manifold F and magnetic sublevel m_F. This is achieved by first performing a $\pi/2$ spin rotation around the y axis to map the F_x-projection to the z axis, which allows its detection via density differences. Then, by transferring half of the population of each m_F level from $F = 1$ to $F = 2$, the F_x projection is stored in the populations of $F = 2$. Finally, another $\pi/2$ spin rotation around the x axis, which exclusively addresses the $F = 1$ manifold, maps the F_y projection to a detectable population difference in $F = 1$. All 6 density distributions are read out with spatial resolution along the longitudinal trap direction via Stern-Gerlach separation and absorption imaging.

At the final parameters of the quench, which places the system into the regime of the easy-plane ferromagnetic phase [32,33], these define the transverse spin field $F_\perp = F_x + iF_y$ [57]. Here, the interplay between energy offsets and spin interactions favor a finite transverse magnetization. During

the dynamics the transverse spin field approaches its ground state distribution, which manifests itself in the formation of a ring in the transverse spin histogram after approximately 1–3 s, as shown in Fig. 2(b). Nevertheless, in this regime, the system is still highly excited and transverse spin phase excitations evolve dynamically in a self-similar fashion [31]. Such relaxation dynamics is quite rich but complex, making a controlled microscopic characterization extremely challenging. In fact, the interpretation above is motivated by heuristic arguments. The key point we are interested in here is to obtain such description based solely on experimental observations, analyzed in a blind-folded manner. That is, we wish to extract essential descriptive elements (most important operators and complexity of the dynamics) without relying on any assumption.

We note that the structure of the experimental setup described above is that of a continuous-variable quantum simulator (see, for instance, Refs. [58–60]). In particular, the present experiment can be regarded as an analog quantum simulation of the out-of-equilibrium dynamics in a quantum field theory associated to the underlying physical system. Furthermore, since the spinor Bose gas under consideration features universal dynamics, this specific setup can be used to probe universal dynamics in a wide range of systems that share the same universal features.

Let us now describe the structure of the collected experimental data sets analysed in this work. At each evolution time, each density is sampled linearly at $p = 184$ spatial locations along the longitudinal trap direction. Such measurements are repeated so as to gather $N_r = 225$ independent realizations. We denote a single realization of a spatial density profile by a p-dimensional vector $\vec{n}_\alpha^i(t)$, where the considered internal state is succinctly labeled by $\alpha \equiv (F, m_F)$. Thus, for each observable (density), and at each evolution time t, we obtain a data set $\mathbf{M}_\alpha(t) = \{\vec{n}_\alpha^1(t), \vec{n}_\alpha^2(t), \ldots, \vec{n}_\alpha^{N_r}(t)\}$, which can be represented as a $(N_r \times p)$ rectangular data matrix. Using the terminology of the previous section, the features of the experimental data at hand are therefore the measured densities at selected positions. Examples of single realizations at different evolution times are shown in Appendix A. Further, we also consider *joint* data sets formed by concatenating horizontally data sets of the measured densities for different α at a given time. More specifically, each row in a joint data matrix is formed by appending, one after the other, single realizations of the observables of choice. Thus, for example, a joint data set of two observables will have twice as many features as the data set of one individual observable, but the same number of rows N_r. The particular order in which we concatenate the combined observables is irrelevant for our methods. When needed, we will simply specify joint data sets by using the symbol of the corresponding observables joined by "$\|$."

IV. IDENTIFICATION OF RELEVANT OBSERVABLES

We now perform a descriptive analysis of the data sets above, with the goal of identifying *relevant* observables, that is, those observables that capture dominant spatial correlations across the evolution of the system. This is, in fact, a crucial task in order to determine good degrees of freedom emerging from the underlying physical system.
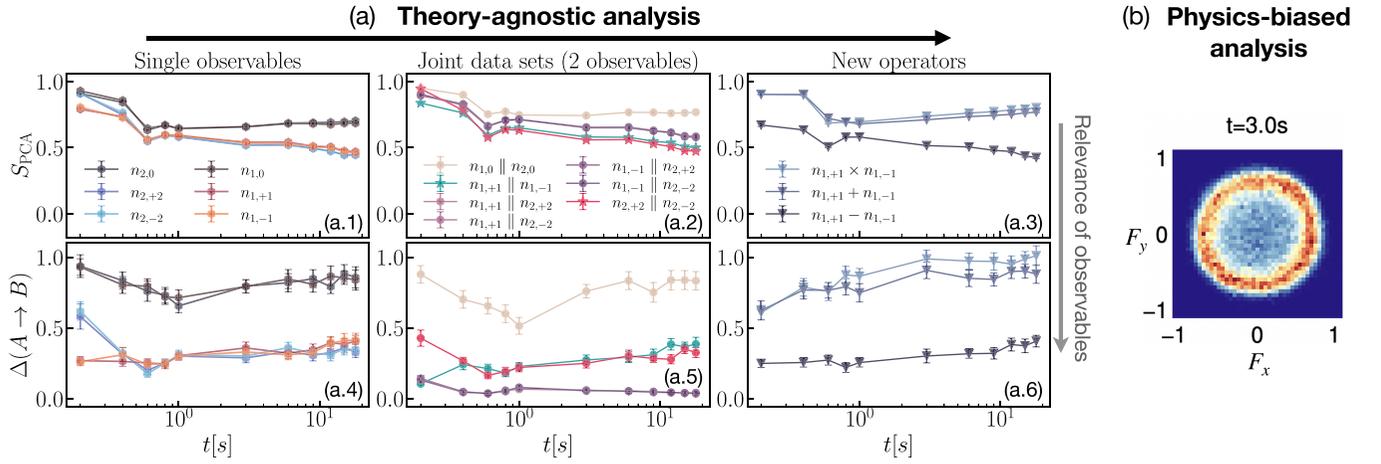
FIG. 2. Assumption-free identification of relevant observables. (a) PCA entropy, S_{PCA}, and information imbalance, $\Delta(A \rightarrow B)$, as relevance metrics of physical observables: Lower values of S_{PCA} signal stronger correlations between the features of an observable, while lower values of $\Delta(A \rightarrow B)$ indicate that the features of a given observable (space A) are more informative of the full set of measured features (space B). Both metrics clearly show that $n_{1,\pm1}$ and $n_{2,\pm2}$ are more relevant (in the sense above) over the full evolution [(a1) and (a4)]. Identification of relevant groups is also possible by analyzing joint data sets: For pairs of observables, $n_{1,+1} \parallel n_{1,-1}$ and $n_{2,+2} \parallel n_{2,-2}$, have the lowest S_{PCA} [⋆ markers in (a2)] (see ranking of all possible pairs in Appendix B). Features from both "relevant" pairs are in fact required to better describe space B [points with $\Delta(A \rightarrow B) \approx 0$ in (a5), for which A is defined by the set of features of $n_{1,\pm1} \parallel n_{2,\pm2}$. Note that these points are almost on top of each other.] Relevant new operators defined from the measured observables can also be identified, as illustrated here for a few combinations of $n_{1,+1}$ and $n_{1,-1}$, with $n_{1,+1} - n_{1,-1}$ being the most relevant [(a3) and (a6)]. (b) Histogram of the transverse spin variable in the $F_x - F_y$ plane at $t = 3$ s, featuring a ringlike structure. Based on physical arguments [30–33], this variable is the relevant field to describe the quenched system. Our theory-agnostic approach identifies the relevant observables from which this variable is inferred [see Eq. (4)], hence cross-validating the latter analysis.

The framework presented here builds on the complementary tools discussed in Secs. II A and II B. We first compute the PCA entropy of the measured spatial density profiles and their combinations, and use it as a direct probe of the spatial correlations captured by those observables at each evolution time. Next, we use information imbalance as a way to determine which observables retain more information from the full space of observations, thereby providing a complementary metric for observable relevance. Let us note that, as discussed in Sec. II A, S_{PCA} cannot reveal the presence of correlations if the embedding manifold of the data is curved. On the other hand, the estimator of the information imbalance in Eq. (2) depends only on the local neighborhood of each data point, and hence is well-suited to deal with arbitrarily nonlinear manifolds [13]. In this sense, ranking the relevance of observables with both techniques provides a way to cross-verify the validity of our results.

Our main results are shown in Fig. 2(a). Let us first analyze the results for the PCA entropy [panels (a1)–(a3)]. A clear separation between two groups of observables is noted as the system evolves [panel (a1)], with $n_{1,\pm1}$ and $n_{2,\pm2}$ having lower values of S_{PCA}. We conclude that these observables capture stronger spatial correlations and are hence more relevant in the sense specified above. Next, we consider joint data sets of two observables [panel (a2)]. The most relevant pairs according to this analysis are $\{n_{1,+1}, n_{1,-1}\}$ and $\{n_{2,+2}, n_{2,-2}\}$. The latter result is in excellent agreement with the physics-motivated analysis, in which such observables play a key role in the definition of the transverse spin [see Eq. (4) and Fig. 2(b)]. Going one step further, in panel (a3), we explore concrete functional combinations of the pair of observables $\{n_{1,+1}, n_{1,-1}\}$

(similar results are found for $\{n_{2,+2}, n_{2,-2}\}$), which define new operators. We find that $n_{1,+1} - n_{1,-1}$ has the lowest S_{PCA}, once again in agreement with the physics-motivated *ansatz* [Eq. (4)].

We now turn our attention to the information imbalance analysis [panels (a4)–(a6)], which provides a complementary view on the relevance of observables. Here, A refers to the space of features associated to a given observable, while B is the full space of $(184 \times 6 = 1104)$ measured features. In panel (a4), we can see that the observables with a lower PCA entropy have also a lower information imbalance. In other words, the observables that capture dominant spatial correlations are also more informative of the full space of data features (in the information imbalance sense). Interestingly, in the analysis of pairs of observables [panel (a5)], we note that in order to describe the full space of observations, one needs to consider features from both of the relevant pairs. Indeed, we see that the space A that combines features from the observables $n_{1,\pm1}$ and $n_{2,\pm2}$ have $\Delta(A \rightarrow B) \approx 0$, for the full evolution [note that the points corresponding to the four possible combinations of these observables are almost on top of each other in panel (a5)]. The new operator $n_{1,+1} - n_{1,-1}$ has also a significantly lower information imbalance than the other considered operators in panel (a6).

In Appendix A, we show some instances of single realizations of the measured density profiles. While one can see a certain correlation in the behavior of the identified relevant observables (at the level of single realizations), our analysis rules out, in a systematic and unbiased way, the presence of other important correlations among the measured quantities. Further results for the analysis carried out here are shown in Appendix B.

V. ALGORITHMIC COMPLEXITY OF THE QUANTUM DYNAMICS

Finally, we provide a further characterization of the data sets. Specifically, we study their intrinsic dimension I_d, at the considered evolution times. This allows us to characterize the *algorithmic complexity* of the quantum dynamics probed in our experiment. We note that this notion of complexity is fundamentally different from *computational complexity* – prominent examples of the latter being entanglement [61–63] and quantum circuit complexity [64,65]. In the context of quantum many-body systems, computational complexity deals in general with characterizing the number of classical resources needed to efficiently simulate a quantum state (e.g., the bond dimension of a matrix product state representation or the number of gates needed in a quantum circuit to describe a target state). On a very distinct note, algorithmic complexity quantifies the notion of compression of information in a classical object, which in our particular case refers to the classical encoding of a quantum state given by the output of measurements (or even classical calculations). We note that these two notions of complexity do not necessarily align with each other, and while the computational complexity has been explored intensively in the realm of strongly interacting systems, much less is known about the algorithmic complexity of quantum many-body states.

As explored in recent studies on critical phenomena in many-body systems [44,54–56], the algorithmic complexity does provide a physical picture of the complexity of many-body states, revealing for example an emergent simplification of the data manifold in systems featuring universality, where the physics also becomes parametrically simpler and can be described by a handful of universal exponents and functions. The present work thus extends the study of this kind of complexity to the realm of far-from-equilibrium quantum many-body dynamics.

Shown in Fig. 3(a) is the plot of I_d as a function of time, of the data sets corresponding to the identified relevant operators, $n_{1,+1} - n_{1,-1}$ and $n_{2,+2} - n_{2,-2}$. We observe the same trend in both instances: a quick decay of I_d to considerably smaller values, subsequently displaying long, stable plateaus. The reduction of the I_d signals a simplification of the data structure due to the buildup of correlations among the input variables. The latter is a direct manifestation of the correlations among the elementary constituents of the system. From the physical viewpoint, the postquench correlations are associated with the formation of a ringlike structure, with approximately constant radius, in the transverse component of the collective spin degree of freedom (insets) [30,31]. In turn, spatial correlations of the spin phase excitations exhibit universal scaling dynamics [31]. In the present experiment, the universal scaling regime starts approximately at $t \sim 6$ s [30]. The physical basis for such scaling evolution is a dynamical reduction of the relevant parameters in the system. This is strongly consistent with the observed structural simplification of the data, as also observed in recent studies of critical behavior—in and out of equilibrium—in classical and quantum statistical mechanics systems [44,54–56]. Therefore, in the present case, the observed I_d plateau provides a theory-agnostic lower bound for the timescale after which the dynamics may become simpler, allowing for the emergence of self-similar behavior.
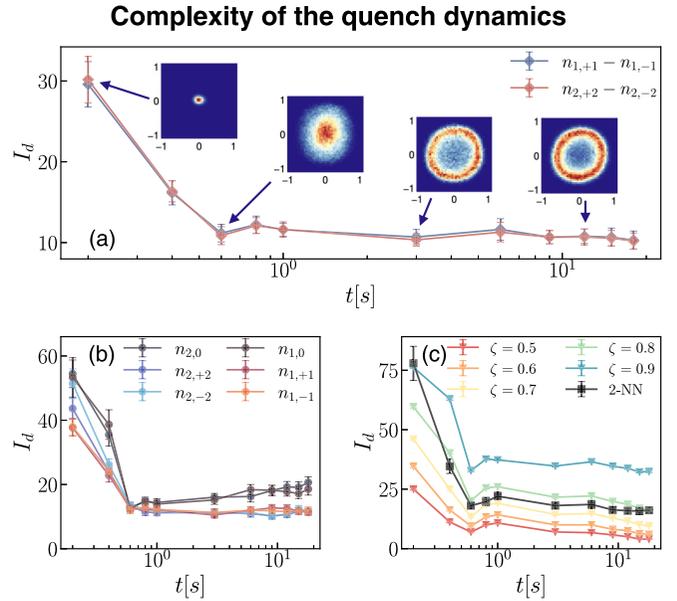

FIG. 3. Intrinsic dimension as a function of time for (a) the relevant observables, $n_{1,+1} - n_{1,-1}$ and $n_{2,+2} - n_{2,-2}$, (b) all measured densities individually, and (c) joint data sets of all six observables together. In all instances, an initially large I_d quickly decays to smaller values (around $t = 0.6$ s), subsequently exhibiting long plateaus. The insets in panel (a) show histograms of the transverse spin in the F_x-F_y plane at selected times. The first drop in the I_d is associated to a grow in the spin length, which remains approximately constant for $t \gtrsim 1$ s. In the latter regime a ringlike structure is then observed (illustrated here at $t = 3$ and 12 s). Spatial correlations of spin phase excitations exhibit self-similar dynamics in a regime that starts around $t \gtrsim 3$ s [30]. The observed structural simplification of the data strongly correlates with such universal dynamics. Hence, the plateaus in the plot of I_d provide a lower bound for the onset of simpler dynamics and universal scaling. Panel (c) also shows the I_d estimate based on PCA for various values of the cutoff ζ (see main text).

Importantly, this prediction can be made by directly studying the I_d of data sets of all measured densities, as shown in Fig. 3(b), where we observe an overall similar trend. We note, however, that the "irrelevant" observables $n_{1,0}$ and $n_{2,0}$, have a growing I_d, rather than a plateau. This further confirms the relevance predictions based on PCA entropy and information imbalance. Further, in Fig. 3(c), we plot the I_d of joint data sets of the six measured observables together, showing once again the noted trend. In this plot, we also show an I_d estimation based on PCA, which is defined by choosing an *ad hoc* cutoff parameter ζ, for the integrated spectrum of the covariance matrix [45,55], i.e., $\sum_{k=1}^{I_d} \tilde{\lambda}_k \approx \zeta$. We find that for all considered values of ζ, we recover the same qualitative features as the TWO-NN I_d estimate. A quantitative agreement can also be achieved for a suitable choice of ζ in the range $0.7 \leqslant \zeta \leqslant 0.9$, at the different evolution times. Since the TWO-NN estimator only depends on local neighborhoods and is therefore well-suited to deal with curved manifolds, as opposed to the PCA method, this agreement further confirms the applicability of PCA in our previous analysis and implies that curvature effects in the data manifold are negligible.
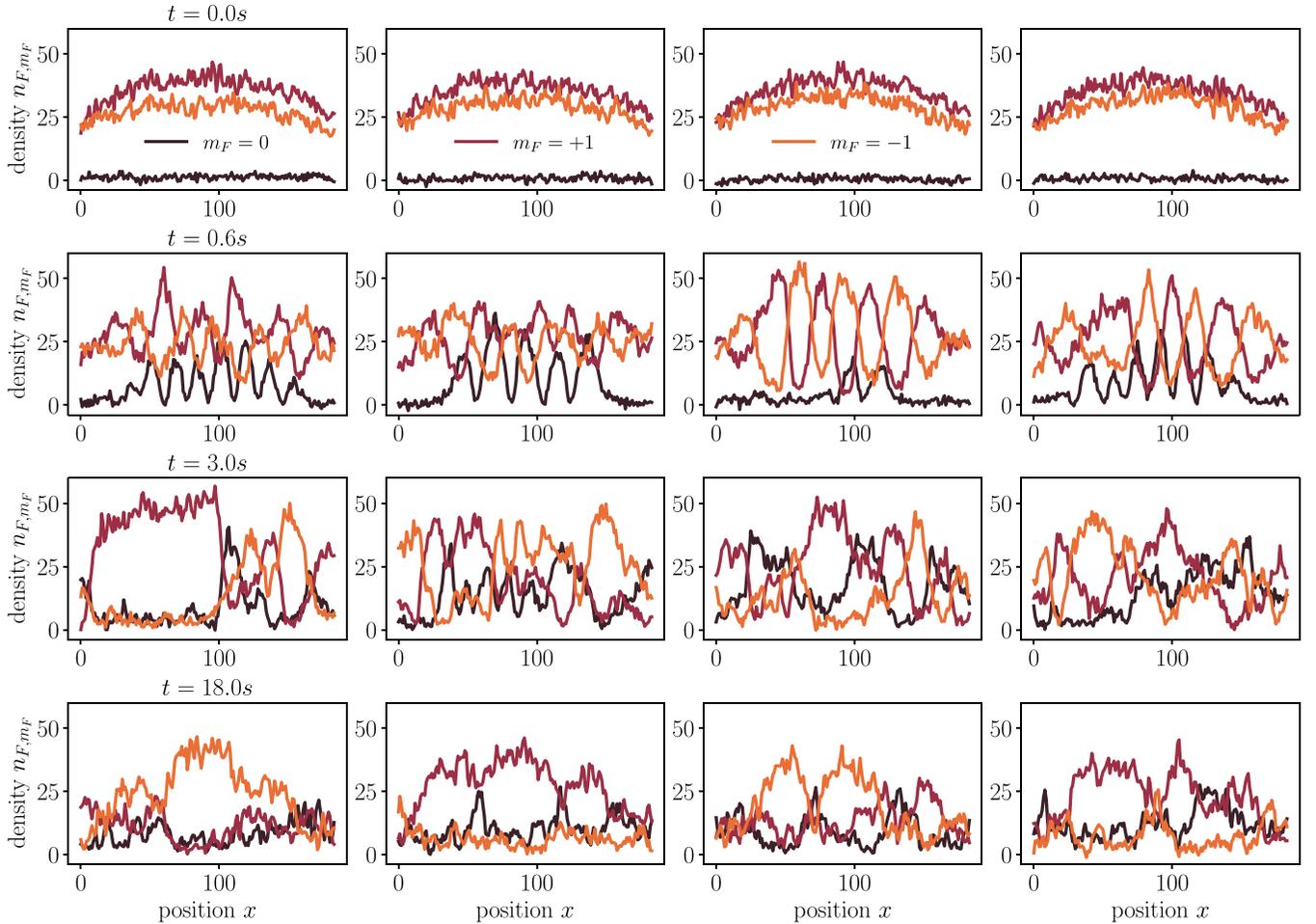
FIG. 4. Experimentally observed density profiles in the hyperfine manifold $F = 1$. Single realizations of the measured densities are shown at selected evolution times in the subsequent rows.

In Appendix C, we show some examples of the linear fits used to estimate I_d as prescribed by the TWO-NN algorithm; see Eq. (3). Statistical error on the results presented in this work were computed using a version of the delete-d Jackknife standard error estimator via a stochastic subsampling algorithm without repetitions [66,67], as detailed in Appendix D.

VI. CONCLUSIONS

We have introduced an assumption-free method to diagnose and rank relevant correlations in the dynamics of out-of-equilibrium quantum systems. The method exploits the full spectrum of principal components, as well as recently developed techniques based on information imbalance. We have successfully identified the most relevant operators describing the dynamics of Bose Einstein condensates, confirming previous heuristic approaches (and thus, validating the physical relevance based solely on experimental observations). Utilizing manifold characterization methods, we have also found stable plateaus of the intrinsic dimension of the data sets corresponding to different times, thus providing bounds on the time frame realizing universal quantum dynamics. Our approach is immediately extended to other classes of

quantum simulators—including fermion gases and lattice spin models—providing a flexible, assumption-free framework to discover physical phenomena, as well as to validate their functioning. Our work complements recent theoretical approaches with similar goals regarding the identification of relevant observables [68,69] and characterizing the complexity of quantum dynamics [54,70,71].
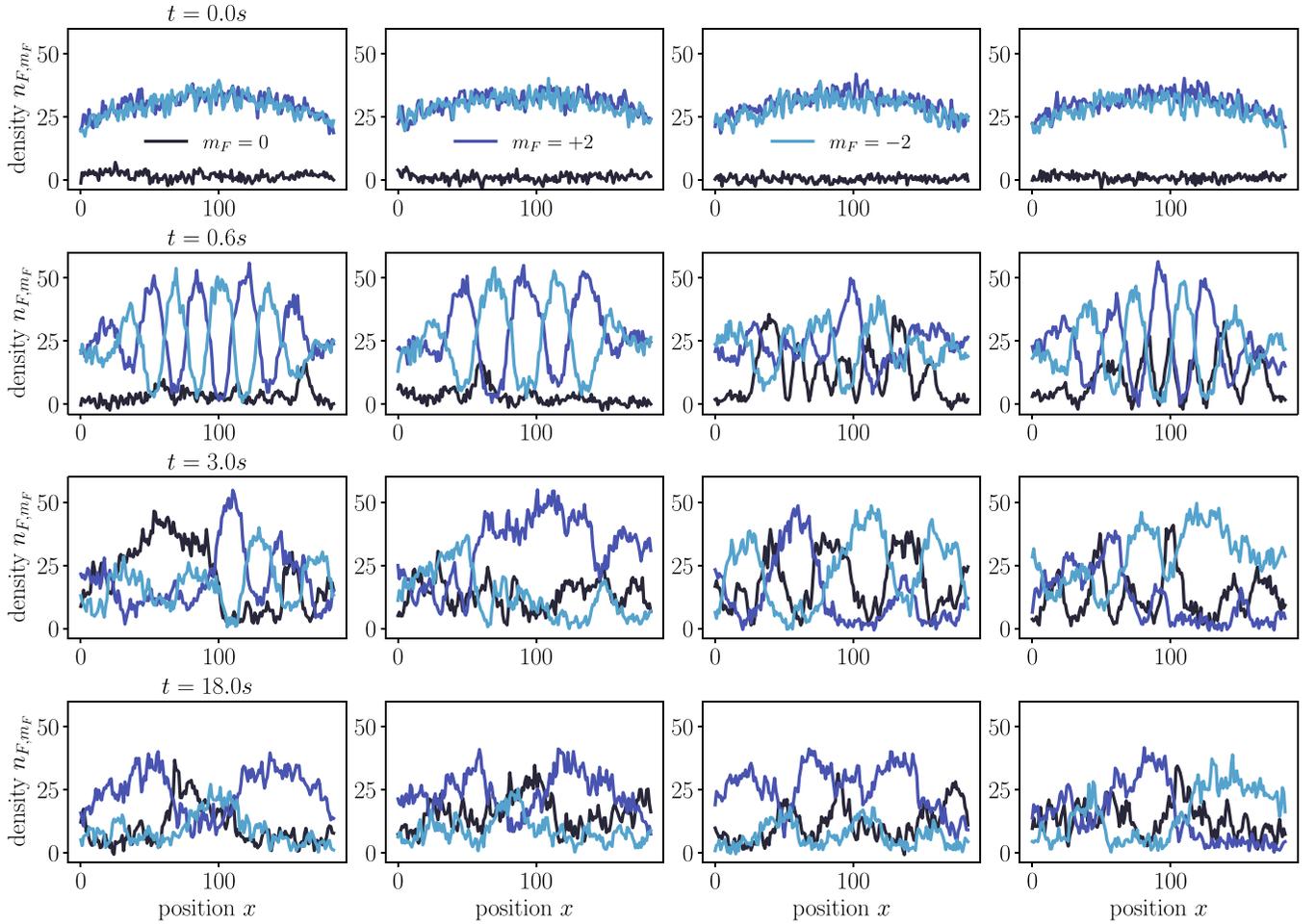
FIG. 5. Experimentally observed density profiles in the hyperfine manifold $F = 2$. Single realizations of the measured densities are shown at selected evolution times in the subsequent rows.

APPENDIX A: EXPERIMENTALLY MEASURED OBSERVABLES

In this Appendix we show, for completeness, examples of single realizations of the measured density profiles at selected evolution times. We refer the reader to Ref. [30], where the experimental data analysed in this work were taken from, for further experimental details. For the sake of clarity, we plot the density profiles corresponding to the hyperfine manifolds $F = 1$ and 2, separately in Figs. 4 and 5, respectively. The

structure of the experimental data is as follows: before measuring the atomic densities a $\pi/2$ rotation around a transverse spin axis is performed. That means that measurements are

FIG. 7. Information imbalance from the space of features of joint data sets combining two observables to the full space of measured features. This plot shows the results for all joint data sets of two observables [cf. Fig. 2(a.5) of the main text]. Note that the points corresponding to the joint data sets $n_{1,+1} \parallel n_{2,+2}$, $n_{1,+1} \parallel n_{2,-2}$, $n_{1,-1} \parallel n_{2,+2}$, and $n_{1,-1} \parallel n_{2,-2}$, lie basically on top of each other (they are equally informative), with an information imbalance $\Delta(A \to B) \approx 0$. These four data sets combine features of the two relevant pairs $\{n_{1,+1}, n_{1,-1}\}$ and $\{n_{2,+2}, n_{2,-2}\}$.
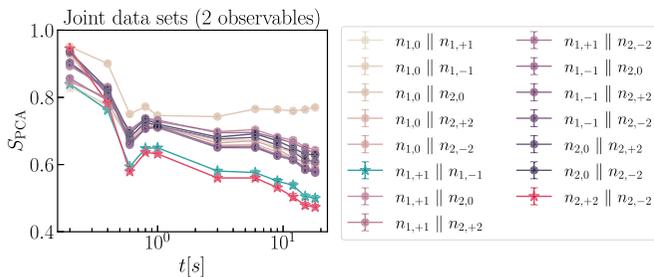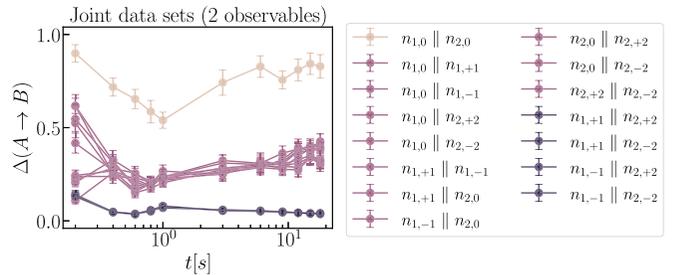

FIG. 6. PCA entropy of joint data sets of two observables. This plot shows the results for all joint data sets of two observables [cf. Fig. 2(a.2) of the main text].
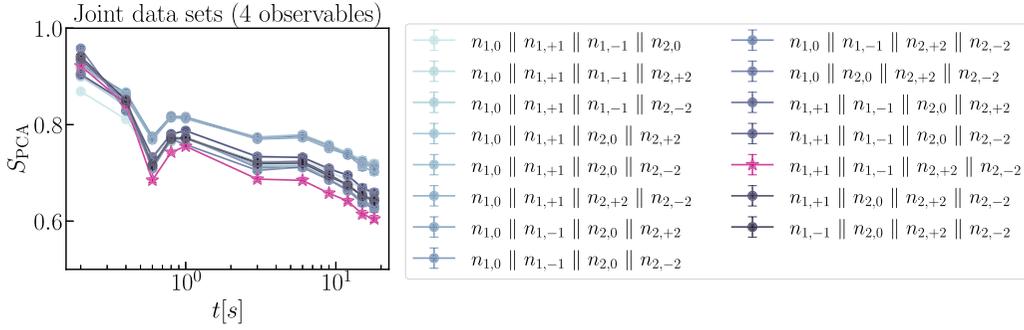
FIG. 8. PCA entropy of joint data sets of four observables. The joint data set with the smallest PCA entropy is the one that combines the two relevant pairs (⋆ markers).

done in the x or y basis. The states at the poles of the spin sphere, therefore, correspond to fully elongated spins along the x or y direction, respectively. At the poles one then measures all atoms in the internal states $(F, m_F) = (1, \pm 1)$ or $(F, m_F) = (2, \pm 2)$ [7].

In the initial state, all atoms are in the magnetic substate $m_F = 0$, as discussed in the main text. However, due to the applied $\pi/2$ pulse before imaging the atoms, the atomic densities get modified and the initial $m_F = 0$ population is split between the substates $m_F = \pm 1$ and $m_F = \pm 2$, for the readouts in $F = 1$ and 2, respectively. This is precisely what we observe in the first row of in Figs. 4 and 5, which corresponds to $t = 0.0$ s. We note however that no spatial correlations are imprinted in such an initial state. This is why both the PCA entropy and the intrinsic dimension of the measured densities are initially large (see Figs. 2 and 3, in the main text). That is, the corresponding data collections at short times exhibit the largest algorithmic complexity due to the statistical independence of the input variables (densities at different spatial positions).

At later evolution times, certain correlations between atomic densities in different substates can be observed at the level of single realizations. These are a consequence of certain symmetries of the system under study. For example, rotational invariance is ultimately responsible for a conserved population imbalance between the substates $m_F = \pm 1$ during the spin-changing collisions that drive the spinor Bose gas out of equilibrium [32,33]. However, excluding the presence of other important correlations by direct inspection becomes very hard and one needs to rely on techniques such as the ones introduced in this work, which allow to identify dominant correlations in a fully systematic and unbiased manner.

APPENDIX B: FURTHER RESULTS ON RELEVANT OBSERVABLES

In this section, we show the ranking of operators beyond the ones discussed in the main text. In particular, in panels (a2) and (a5) of Fig. 2, we only show results for some of the possible combinations of two of the observed densities. In Figs. 6 and 7, we show the results for all possible 15 combinations. In these plots we can better appreciate the separation into three groups of combinations. Focusing first on S_{PCA}, we

can distinguish the two most informative pairs of observables with lower values of S_{PCA}, throughout almost the whole evolution, namely, $\{n_{1,+1}, n_{1,-1}\}$ and $\{n_{2,+2}, n_{2,-2}\}$ (⋆ markers in Fig. 6). These are follow by an set of combinations—each containing at least one of the observables in the identified relevant pairs—with intermediates values of S_{PCA}, and finally the "least" informative combination $\{n_{1,0}, n_{2,0}\}$. The information imbalance in Fig. 7, reveals a similar separation. First, with an almost zero information imbalance, we have the space of features corresponding to the combinations $\{n_{1,\pm 1}, n_{2,\pm 2}\}$ (points with a darker color in Fig. 7). This results, as discussed in the main text, means that in order to describe the full space of measured features we need to combine features from each of the two relevant pairs $\{n_{1,+1}, n_{1,-1}\}$ and $\{n_{2,+2}, n_{2,-2}\}$. The next group of combinations, yielding intermediate values of $\Delta(A \to B)$, contain features within only one of such relevant pairs, e.g., $\{n_{1,0}, n_{1,-1}\}$ or $\{n_{2,+2}, n_{2,-2}\}$. Finally, once again, we identify as the "least" informative combination that of $\{n_{1,0}, n_{2,0}\}$.

Similar observations are obtained if one considers groups of more than two observables. This is illustrated here for groups of four observables (quadruplets) in Figs. 8 and 9. In terms of PCA entropy the most relevant combination is the one that combines the features of the two relevant pairs, namely, the joint data set $n_{1,+1} \parallel n_{1,-1} \parallel n_{2,+2} \parallel n_{2,-2}$ (⋆ markers in Fig. 8). We note however that in this case the relative difference in PCA entropy is not as pronounced as in the case of single or two observables. Regarding information imbalance, we observe once again those joint data sets that combine features from the two relevant pairs can predict almost entirely the full space of features. There are indeed only two combinations that only involve features from a single relevant pair (plus the two "irrelevant" observables $n_{1,0}$ and $n_{2,0}$), namely, $n_{1,0} \parallel n_{1,+1} \parallel n_{1,+1} \parallel n_{2,0}$ and $n_{1,0} \parallel n_{2,0} \parallel n_{2,+2} \parallel n_{2,-2}$ (square markers in Fig. 9), which clearly have a significantly larger information imbalance compared to the rest.

APPENDIX C: LINEAR FIT TO ESTIMATE I_D FROM THE EMPIRICAL CUMULATES IN THE TWO-NN METHOD

In this section, we show examples of the linear fitting procedure used to estimate the value of the intrinsic dimension in the TWO-NN method; see Eq. (3) in the main text. In
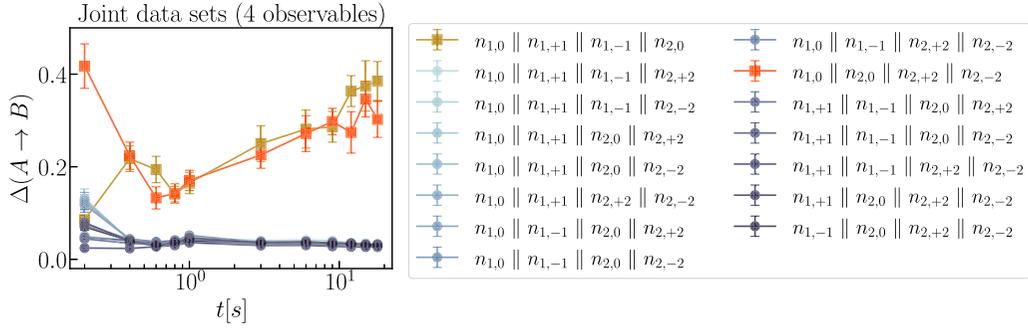
FIG. 9. Information imbalance from the space of features of joint data sets combining four observables to the full space of measured features. All joint data sets that include features from the two relevant pairs $\{n_{1,+1}, n_{1,-1}\}$ and $\{n_{2,+2}, n_{2,-2}\}$ have a very small information imbalance. Instead, the joint data sets that only include features from one of the relevant pairs cannot predict so well the full space of features. The latter are indicated with square markers in this plot.

Fig. 10, we show the empirical cumulative distributions of the ratios $\mu_i = r_{i_2}/r_{i_1}$, sorted in ascending order, for the observable $n_{2,+2} - n_{2,-2}$, at all evolution times. If the condition of constant density in the range of first two nearest neighbors holds, a plot of the resulting points $\{\ln(\mu), -\ln[1 - F_{\text{emp}}(\mu)]\}$ will be a line that passes through the origin and whose slope

gives the estimated value of I_d. Verifying that the empirical cumulates are indeed consistent with a Pareto distribution as described above, is the first step to guarantee the applicability of the TWO-NN method. Besides, on its own, this kind of plot is also very informative about the local structure of complicated data manifolds.
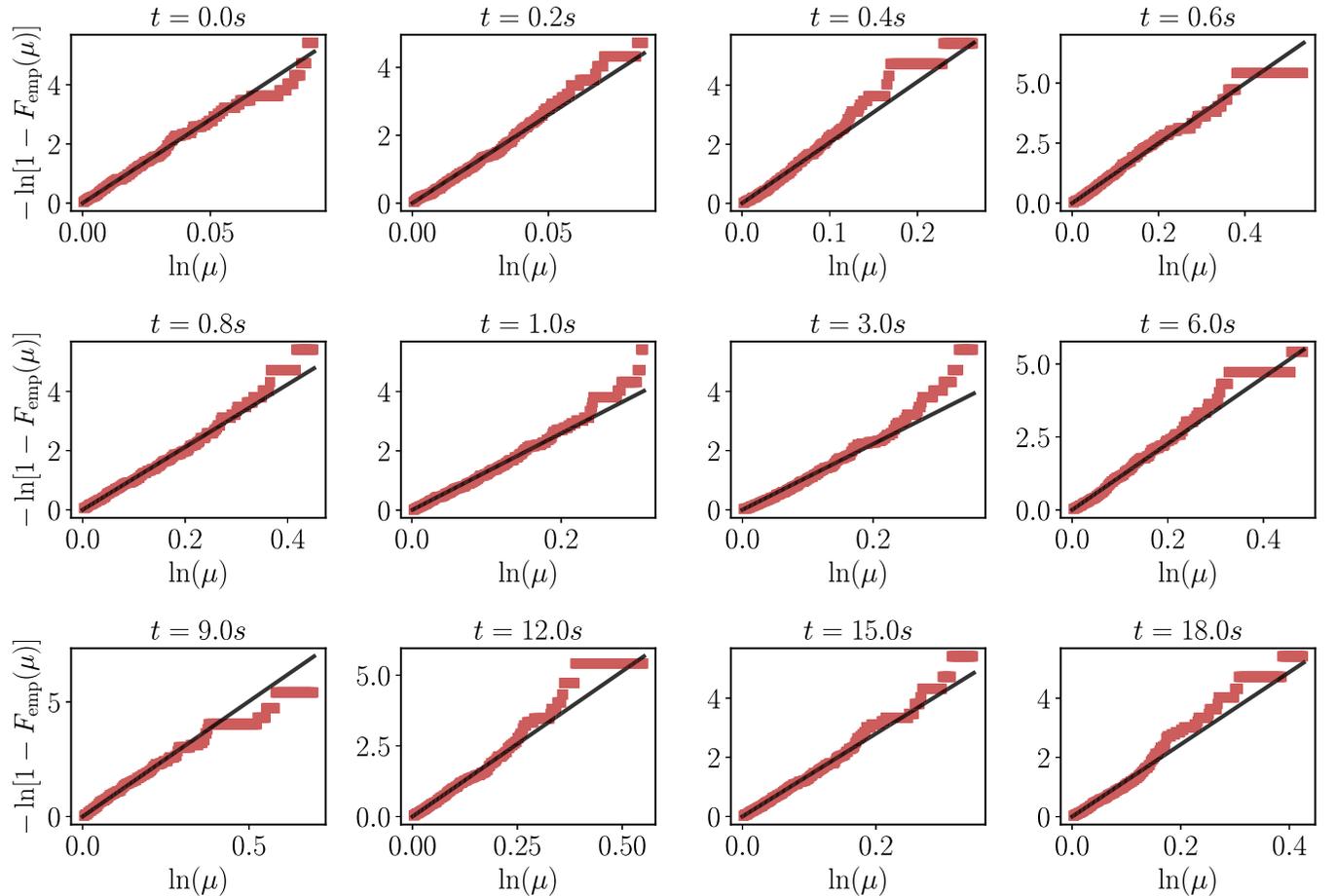

FIG. 10. Empirical cumulative distributions at all evolution times for the data sets corresponding to the relevant field $n_{2,+2} - n_{2,-2}$. The black curve show the linear fit according to Eq. (3) in the main text, whose slope gives the estimated value of I_d. This procedure is valid as long as the empirical cumulative distribution function is consistent with a Pareto distribution, at least over a significant range of values of $\ln(\mu)$, as is clearly the case here.

APPENDIX D: SUBSAMPLING ERROR ESTIMATION

Due to the limited number of realizations used in the present analysis, we opted for using a technique known as subsampling [66,67] to have a sensible estimation of statistical errors. The subsampling algorithm is described below.

At a given time and for a given measured observable, we have $N_r = 225$ independent realizations forming our data set, that is, $\mathbf{M} = \{\vec{n}^i\}_{i=1}^{N_r}$, where for simplicity we have omitted the indices labeling the internal state and the evolution time. Using these data we compute a certain numerical statistic ϑ. Given two preset integers b and $q < N_r$, the subsampling analysis proceeds as follows.

1. Form b random "batches" (subsamples) of data by drawing $q < N_r$ points at random but *without* replacement from the data set $\mathbf{M}$.

2. Estimate the statistic of interest on each subsample, that is, ϑ_j, for $j \in \{1, \ldots, b\}$.

3. Compute the mean of such estimates $\overline{\vartheta} = \frac{1}{b}\sum_{j=1}^{b} \vartheta_j$. The standard error can then be estimated as

follows:

$$\text{SE} \approx \sqrt{\frac{q}{N_r - q}} \cdot \sqrt{\frac{1}{b}\sum_{j=1}^{b}(\vartheta_j - \overline{\vartheta})^2}. \qquad \text{(D1)}$$

This formula is known as the delete-d Jackknife standard error estimator (with stochastic subsampling) [66,67].

While formally this method requires $q/N_r \to 0$, $q \to \infty$, and $b \to \infty$ as $N_r \to \infty$ (so that the distribution of the ϑ_i converges to the sampling distribution of ϑ), in practice the choice of these parameters is problem specific. In our analysis, we did not find significant changes for $b \geqslant 30$. Hence, we fixed $b = 30$. Furthermore, to compute a meaningful statistic on each subsample, we set $q = 100$, satisfying at least $q/N_r < 1$.

We used this method since sampling is performed *without* replacement. This is important as the TWO-NN algorithm used to estimate the I_d works under the assumption of no repetitions among the data points.

[1] I. Bloch, J. Dalibard, and W. Zwerger, Many-body physics with ultracold gases, Rev. Mod. Phys. **80**, 885 (2008).

[2] M. Lewenstein, A. Sanpera, and V. Ahufinger, *Ultracold Atoms in Optical Lattices: Simulating Quantum Many-Body Systems* (Oxford University Press, Oxford, UK, 2012).

[3] I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Rev. Mod. Phys. **86**, 153 (2014).

[4] C. Gross and I. Bloch, Quantum simulations with ultracold atoms in optical lattices, Science **357**, 995 (2017).

[5] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, Nature (London) **607**, 667 (2022).

[6] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

[7] 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).

[8] A. M. Tsvelik, *Quantum Field Theory in Condensed Matter Physics*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2003).

[9] H. Kleinert, *Collective Classical and Quantum Fields* (World Scientific, Singapore, 2018).

[10] O. Alter, P. O. Brown, and D. Botstein, Singular value decomposition for genome-wide expression data processing and modeling, Proc. Natl. Acad. Sci. USA **97**, 10101 (2000).

[11] R. Varshavsky, A. Gottlieb, M. Linial, and D. Horn, Novel unsupervised feature filtering of biological data, Bioinformatics **22**, e507 (2006).

[12] R. Varshavsky, A. Gottlieb, D. Horn, and M. Linial, Unsupervised feature selection under perturbations: Meeting the challenges of biological data, Bioinformatics **23**, 3343 (2007).

[13] A. Glielmo, C. Zeni, B. Cheng, G. Csányi, and A. Laio, Ranking the information content of distance measures, PNAS Nexus **1**, pgac039 (2022).

[14] E. D. Donkor, A. Laio, and A. Hassanali, Do machine-learning atomic descriptors and order parameters tell the same story? The case of liquid water, J. Chem. Theory Comput. **14**, 4596 (2023).

[15] E. Facco, M. d'Errico, A. Rodriguez, and A. Laio, Estimating the intrinsic dimension of datasets by a minimal neighborhood information, Sci. Rep. **7**, 12140 (2017).

[16] E. Facco, A. Pagnani, E. T. Russo, and A. Laio, The intrinsic dimension of protein sequence evolution, PLoS Comput. Biol. **15**, e1006767 (2019).

[17] M. Allegra, E. Facco, F. Denti, A. Laio, and A. Mira, Data segmentation based on the local intrinsic dimension, Sci. Rep. **10**, 16449 (2020).

[18] A. Glielmo, B. E. Husic, A. Rodriguez, C. Clementi, F. Noé, and A. Laio, Unsupervised learning methods for molecular simulation data, Chem. Rev. **121**, 9722 (2021).

[19] T. Strydom, G. V. Dalla Riva, and T. Poisot, SVD entropy reveals the high complexity of ecological networks, Front. Ecol. Evol. **9**, 623141 (2021).

[20] P. Caraiani, The predictive power of singular value decomposition entropy for stock market dynamics, Physica A **393**, 571 (2014).

[21] R. Gu, W. Xiong, and X. Li, Does the singular value decomposition entropy have predictive power for stock market? — evidence from the shenzhen stock market, Physica A **439**, 103 (2015).

[22] R. Gu and Y. Shao, How long the singular value decomposed entropy predicts the stock market? — evidence from the Dow Jones Industrial Average Index, Physica A **453**, 150 (2016).

[23] A. Chakraborti, Hrishidev, K. Sharma, and H. K. Pharasi, Phase separation and scaling in correlation structures of financial markets, J. Phys. Complex. **2**, 015002 (2021).

[24] J. Alvarez-Ramirez and E. Rodriguez, A singular value decomposition entropy approach for testing stock market efficiency, Physica A **583**, 126337 (2021).

[25] G. Espinosa-Paredes, E. Rodriguez, and J. Alvarez-Ramirez, A singular value decomposition entropy approach to assess the impact of Covid-19 on the informational efficiency of the WTI crude oil market, Chaos, Solitons Fractals **160**, 112238 (2022).

[26] E. Levina and P. Bickel, Maximum likelihood estimation of intrinsic dimension, in *Advances in Neural Information Processing Systems*, edited by L. Saul, Y. Weiss, and L. Bottou (MIT Press, Cambridge, Massachusetts, 2004), Vol. 17.

[27] S. Gong, V. Boddeti, and A. K. Jain, On the intrinsic dimensionality of image representations, in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society, Los Alamitos, CA, USA, 2019), pp. 3982–3991.

[28] P. Pope, C. Zhu, A. Abdelkader, M. Goldblum, and T. Goldstein, The intrinsic dimension of images and its impact on learning, in *International Conference on Learning Representations* (ICLR, 2021).

[29] X. Weng, A. Perry, M. Maroun, and L. T. Vuong, Singular value decomposition and entropy dimension of fractals, in *Proceedings of the 2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML)* (IEEE, 2022), pp. 427–431.

[30] M. Prüfer, T. V. Zache, P. Kunkel, S. Lannig, A. Bonnin, H. Strobel, J. Berges, and M. K. Oberthaler, Experimental extraction of the quantum effective action for a non-equilibrium many-body system, Nat. Phys. **16**, 1012 (2020).

[31] M. Prüfer, P. Kunkel, H. Strobel, S. Lannig, D. Linnemann, C.-M. Schmied, J. Berges, T. Gasenzer, and M. K. Oberthaler, Observation of universal dynamics in a spinor Bose gas far from equilibrium, Nature (London) **563**, 217 (2018).

[32] D. M. Stamper-Kurn and M. Ueda, Spinor Bose gases: Symmetries, magnetism, and quantum dynamics, Rev. Mod. Phys. **85**, 1191 (2013).

[33] Y. Kawaguchi and M. Ueda, Spinor Bose–Einstein condensates, Phys. Rep. **520**, 253 (2012).

[34] J. Berges, K. Boguslavski, S. Schlichting, and R. Venugopalan, Universality far from equilibrium: From superfluid Bose gases to heavy-ion collisions, Phys. Rev. Lett. **114**, 061601 (2015).

[35] S. Erne, R. Bücker, T. Gasenzer, J. Berges, and J. Schmiedmayer, Universal dynamics in an isolated one-dimensional Bose gas far from equilibrium, Nature (London) **563**, 225 (2018).

[36] J. A. P. Glidden, C. Eigen, L. H. Dogra, T. A. Hilker, R. P. Smith, and Z. Hadzibabic, Bidirectional dynamic scaling in an isolated Bose gas far from equilibrium, Nat. Phys. **17**, 457 (2021).

[37] L. Wang, Discovering phase transitions with unsupervised learning, Phys. Rev. B **94**, 195105 (2016).

[38] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, Phys. Rev. E **95**, 062122 (2017).

[39] N. C. Costa, W. Hu, Z. J. Bai, R. T. Scalettar, and R. R. P. Singh, Principal component analysis for fermionic critical points, Phys. Rev. B **96**, 195138 (2017).

[40] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, Phys. Rev. E **96**, 022140 (2017).

[41] C. Wang and H. Zhai, Machine learning of frustrated classical spin models. I. Principal component analysis, Phys. Rev. B **96**, 144432 (2017).

[42] K. Ch'ng, N. Vazquez, and E. Khatami, Unsupervised machine learning account of magnetic transitions in the Hubbard model, Phys. Rev. E **97**, 013306 (2018).

[43] E. Khatami, Principal component analysis of the magnetic transition in the three-dimensional Fermi-Hubbard model, J. Phys.: Conf. Ser. **1290**, 012006 (2019).

[44] X. Turkeshi, Measurement-induced criticality as a data-structure transition, Phys. Rev. B **106**, 144313 (2022).

[45] I. Jolliffe, Principal component analysis, in *Encyclopedia of Statistics in Behavioral Science* (John Wiley & Sons, Hoboken, New Jersey, 2005).

[46] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, Phys. Rep. **810**, 1 (2019).

[47] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[48] G.V. Trunk, Statistical estimation of the intrinsic dimensionality of data collections, Inf. Control. **12**, 508 (1968).

[49] P. Campadelli, E. Casiraghi, C. Ceruti, and A. Rozza, Intrinsic dimension estimation: Relevant techniques and a benchmark framework, Math. Probl. Eng. **2015**, 1 (2015).

[50] F. Camastra and A. Staiano, Intrinsic dimension estimation: Advances and open problems, Inf. Sci. **328**, 26 (2016).

[51] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd ed. (Springer, New York, 2009).

[52] L. Staiger, Kolmogorov complexity and Hausdorff dimension, Inf. Comput. **103**, 159 (1993).

[53] V. Vitale, T. Mendes-Santos, A. Rodriguez, and M. Dalmonte, Topological kolmogorov complexity and the Berezinskii-Kosterlitz-Thouless mechanism, arXiv:2305.05396.

[54] T. Mendes-Santos, M. Schmitt, A. Angelone, A. Rodriguez, P. Scholl, H. J. Williams, D. Barredo, T. Lahaye, A. Browaeys, M. Heyl, and M. Dalmonte, Wave function network description and Kolmogorov complexity of quantum many-body systems, arXiv:2301.13216.

[55] T. Mendes-Santos, X. Turkeshi, M. Dalmonte, and A. Rodriguez, Unsupervised learning universal critical behavior via the intrinsic dimension, Phys. Rev. X **11**, 011040 (2021).

[56] T. Mendes-Santos, A. Angelone, A. Rodriguez, R. Fazio, and M. Dalmonte, Intrinsic dimension of path integrals: Data-mining quantum criticality and emergent simplicity, PRX Quantum **2**, 030332 (2021).

[57] L. E. Sadler, J. M. Higbie, S. R. Leslie, M. Vengalattore, and D. M. Stamper-Kurn, Spontaneous symmetry breaking in a quenched ferromagnetic spinor Bose–Einstein condensate, Nature (London) **443**, 312 (2006).

[58] C. Viermann, M. Sparn, N. Liebster, M. Hans, E. Kath, Á. Parra-López, M. Tolosa-Simeón, N. Sánchez-Kuntz, T. Haas, H. Strobel, S. Floerchinger, and M. K. Oberthaler, Quantum field simulator for dynamics in curved spacetime, Nature (London) **611**, 260 (2022).

[59] M. Tajik, M. Gluza, N. Sebe, P. Schüttelkopf, F. Cataldini, J. Sabino, F. Møller, S.-C. Ji, S. Erne, G. Guarnieri,

S. Sotiriadis, J. Eisert, and J. Schmiedmayer, Experimental observation of curved light-cones in a quantum field simulator, Proc. Natl. Acad. Sci. USA **120**, e2301287120 (2023).

[60] 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 (2023).

[61] 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).

[62] Z. He and A. J. Millis, Entanglement entropy and computational complexity of the Anderson impurity model out of equilibrium: Quench dynamics, Phys. Rev. B **96**, 085107 (2017).

[63] Z.-C. Yang, A. Hamma, S. M. Giampaolo, E. R. Mucciolo, and C. Chamon, Entanglement complexity in quantum many-body dynamics, thermalization, and localization, Phys. Rev. B **96**, 020408(R) (2017).

[64] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of quantum complexity growth, PRX Quantum **2**, 030316 (2021).

[65] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert, and N. Y. Halpern, Linear growth of quantum circuit complexity, Nat. Phys. **18**, 528 (2022).

[66] D. N. Politis, J. P. Romano, and M. Wolf, *Subsampling*, 1st ed. (Springer, New York, 1999).

[67] J. Shao and D. Tu, *The Jackknife and Bootstrap*, 1st ed. (Springer, New York, 1995).

[68] C. Miles, A. Bohrdt, R. Wu, C. Chiu, M. Xu, G. Ji, M. Greiner, K. Q. Weinberger, E. Demler, and E.-A. Kim, Correlator convolutional neural networks as an interpretable architecture for image-like quantum matter data, Nat. Commun. **12**, 3905 (2021).

[69] A. Bohrdt, S. Kim, A. Lukin, M. Rispoli, R. Schittko, M. Knap, M. Greiner, and J. Léonard, Analyzing nonequilibrium quantum states through snapshots with artificial neural networks, Phys. Rev. Lett. **127**, 150504 (2021).

[70] M. Schmitt and Z. Lenarčič, From observations to complexity of quantum states via unsupervised learning, Phys. Rev. B **106**, L041110 (2022).

[71] H. P. Casagrande, B. Xing, M. Dalmonte, A. Rodriguez, V. Balachandran, and D. Poletti, Complexity of spin configuration dynamics due to unitary evolution and periodic projective measurements, Phys. Rev. E **108**, 044128 (2023).