Wave-Function Network Description and Kolmogorov Complexity of Quantum Many-Body Systems

T. Mendes-Santos,[1,2] M. Schmitt,[3,4] A. Angelone,[5,6] A. Rodriguez,[7,8] P. Scholl,[9,10] H. J. Williams,[11] D. Barredo,[9,12] T. Lahaye,[9] A. Browaeys,[9] M. Heyl,[1] and M. Dalmonte[7,13]

[1]*Theoretical Physics III, Center for Electronic Correlations and Magnetism, Institute of Physics, University of Augsburg, 86135 Augsburg, Germany*
[2]*PASQAL SAS, 7 rue L´eonard de Vinci - 91300 Massy, Paris, France*
[3]*Forschungszentrum Jülich GmbH, Peter Grünberg Institute, Quantum Control (PGI-8), 52425 Jülich, Germany*
[4]*University of Regensburg, 93053 Regensburg, Germany*
[5]*Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée, LPTMC, F-75005 Paris, France*
[6]*eXact lab s.r.l., Via Francesco Crispi 56—34126 Trieste, Italy*
[7]*The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34151 Trieste, Italy*
[8]*Dipartimento di Matematica e Geoscienze, Università degli Studi di Trieste, via Alfonso Valerio 12/1, 34127, Trieste, Italy*
[9]*Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127 Palaiseau Cedex, France*
[10]*California Institute of Technology, Pasadena, California 91125, USA*
[11]*Department of Physics, Durham University, South Road, Durham DH1 3LE, United Kingdom*
[12]*Nanomaterials and Nanotechnology Research Center (CINN-CSIC), Universidad de Oviedo (UO), Principado de Asturias, 33940 El Entrego, Spain*
[13]*SISSA-International School of Advanced Studies, via Bonomea 265, 34136 Trieste, Italy*

(Received 18 February 2023; revised 13 April 2024; accepted 17 April 2024; published 21 May 2024)

Programmable quantum devices are now able to probe wave functions at unprecedented levels. This is based on the ability to project the many-body state of atom and qubit arrays onto a measurement basis which produces snapshots of the system wave function. Extracting and processing information from such observations remains, however, an open quest. One often resorts to analyzing low-order correlation functions—that is, discarding most of the available information content. Here, we introduce wave-function networks—a mathematical framework to describe wave-function snapshots based on network theory. For many-body systems, these networks can become scale-free—a mathematical structure that has found tremendous success and applications in a broad set of fields, ranging from biology to epidemics to Internet science. We demonstrate the potential of applying these techniques to quantum science by introducing protocols to extract the Kolmogorov complexity corresponding to the output of a quantum simulator and implementing tools for fully scalable cross-platform certification based on similarity tests between networks. We demonstrate the emergence of scale-free networks analyzing experimental data obtained with a Rydberg quantum simulator manipulating up to 100 atoms. Our approach illustrates how, upon crossing a phase transition, the simulator complexity decreases while correlation length increases—a direct signature of buildup of universal behavior in data space. Comparing experiments with numerical simulations, we achieve cross-certification at the wave-function level up to timescales of 4 μs with a confidence level of 90% and determine experimental calibration intervals with unprecedented accuracy. Our framework is generically applicable to the output of quantum computers and simulators with *in situ* access to the system wave function and requires probing accuracy and repetition rates accessible to most currently available platforms.

DOI: 10.1103/PhysRevX.14.021029

Subject Areas: Quantum Physics, Quantum Information, Statistical Physics

I. INTRODUCTION

Harnessing and probing many-body systems at the single particle (qubit) level are hallmark features of present-day quantum simulators and computers [1–4]. One of the most drastic demonstrations of these tools is the possibility of taking a large number of "photos" of a many-body system, obtained via projective measurements of the full many-body wave function. While this flood of available observations could be seen as a blessing, it immediately encounters practical as well as conceptual challenges: How can this large amount of data be processed, without *a priori* discarding information (in the data science language, before performing a dimensional reduction)? What can one learn, that is, e.g., not available by utilizing low-order correlation functions? Answering these questions requires a structured, mathematical understanding of the experimental wave function snapshots that addresses the *information limbo* between traditional many-body theory based on few-points correlation functions [5] and full-fledged—but experimentally limited to few-particle systems—tomographic methods [6].

Here, we develop a theoretical framework to characterize and classify experimentally accessible collections of wave function snapshots utilizing network theory that is scalable and allows one to retain all available information. The backbone of our method is a mapping between collections of wave function snapshots and a wave function network (WFN), schematically depicted in Fig. 1, that is applicable to spin, bosonic, and fermionic systems. Utilizing well-established tools in network theory we unravel several key characteristics of the underlying quantum wave function that are inaccessible by conventional means.

The pivotal finding is that the resulting quantum wave function networks can become scale-free—a mathematical structure that has found widespread application in several fields, ranging from power distribution and internet networks to epidemics [7–9]. We demonstrate this property using experimental snapshots obtained on a Rydberg quantum simulator operating with more than 100 atoms [2,10] and with large-scale numerical simulations using neural quantum states (NQSs) [11,12]. We then argue about its generic applicability to state preparation protocols and discuss how other types of networks—Erdős-Rényi (ER) [13]—can instead emerge if the resulting dynamics describes uncorrelated states. In terms of observables, required resources and applicability regimes, our approach is complementary to other methods aimed at fully characterizing quantum states via snapshots such as those based on classical shadows [14], randomized measurements [15–17], and chaotic dynamics [18,19]. Its main distinctive features, that we elaborate upon below, are direct interpretability and straightforward scalability for strongly correlated, low-temperature states.

The correspondence between quantum simulator outputs and conventional network theory immediately enables a transfer of methods and concepts from previously disconnected fields. We leverage this connection to address two challenges in the field of quantum simulation. Firstly, we show that we are able to characterize the complexity of the quantum simulator output by determining its Kolmogorov complexity (KC)—the accepted absolute measure of information content of finite objects [20,21] that quantifies the (in)compressibility of the quantum wave function information as contained in the snapshots. This allows us to demonstrate the emergence of critical behavior at the level of information complexity, directly probing at the wave function level the emergent simplicity dictated by renormalization group theory.

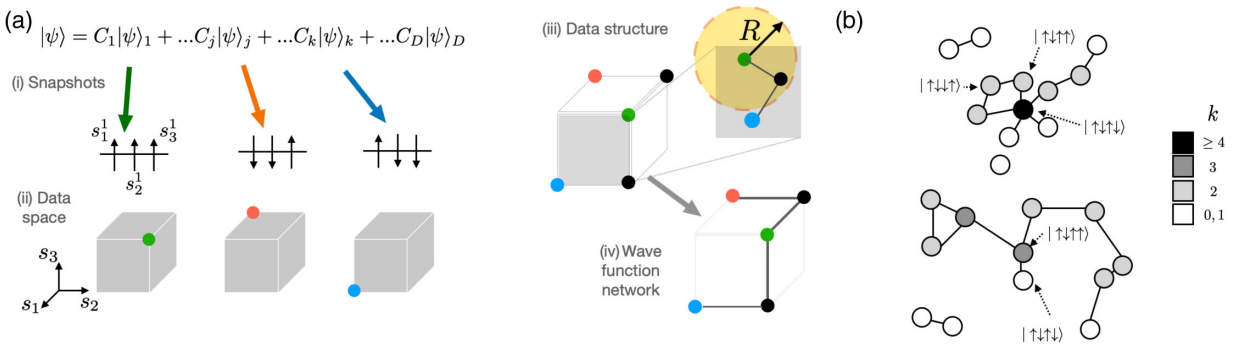

FIG. 1. Network description of many-body wave function snapshots. (a) Construction of the network. First, samples of a wave function are collected (i) and individually mapped onto the target data space (ii). All data are then merged into a single data structure (iii) that defines a set of points in the configuration data space. This data structure is then mapped onto the corresponding wave function network (iv) by drawing links in the network according to a cutoff distance R that is determined by the data structure and the choice of metric (see text). (b) Physical interpretation of the network structure. Within the network, the number of neighbors of given points follows a specific distribution. Points with a large number of links k (i.e., large number of points within R) are hubs, and are indicated in darker colors. As an example, taking snapshots of a classical antiferromagnet below its critical temperature will feature the antiferromagnetic state as a hub (top illustration), while doing so well above its critical temperature will lead to a graph with no hubs and random connections (bottom illustration).

Secondly, we introduce a method to perform cross-platform verification of quantum simulators [22,23]. The method is based on the full network information without the need to perform an exponentially increasing number of measurements for increasing system size, which is the case for generic cross-verification based on the density matrix [22,23]. By means of the Epps-Singleton (ES) test [24] we identify, with statistical significance, a time-scale beyond which cross-verification falters due to experimental imperfections not covered by our theoretical description. In addition, we provide statistically rigorous bounds for previously observed time-delay effects that demonstrate the capability of our methods to identify systematic effects that are invisible to low-order correlation functions. Beyond these two demonstrative tools, the quantum wave function networks introduced in this work provide a new generically applicable framework to probe and characterize the quantum many-body wave function accessible in a variety of atomic and solid-state quantum hardware, solely requiring *in situ* imaging of the many-body wave function.

II. WAVE FUNCTION NETWORKS: THEORETICAL FRAMEWORK

In this section, we describe how datasets generated by a collection of wave function snapshots can be represented by a network structure with nodes and links. For the sake of simplicity, we consider a many-body system composed of spin-1/2 degrees of freedom defined on a two-dimensional lattice: The approach can be straightforwardly generalized to continuum theories, as well as to different types of local Hilbert spaces.

Snapshot dataset.—Each wave function snapshot, labeled by an index j, takes the form

$$X_j[w] = (s_1^j, s_2^j, \ldots, s_N^j), \tag{1}$$

where s_m^j is the measured value of the spin at position m. N is the total number of sites in the system, while w are the external parameters related to the snapshot—in our case the Hamiltonian couplings. Each of these configurations corresponds to a single data point embedded in a data space whose embedding dimension is N. This is depicted in Fig. 1(a) with the three examples of green, orange, and blue dots.

The dataset we are interested in is formed by the collection of all available snapshots:

$$\boldsymbol{X}[w] = \{X_j\} = \{X_1, X_2, \ldots, X_{N_r}\}, \tag{2}$$

where N_r is the number of available snapshots, that is, the number of realizations. The dataset might, in principle, include repetitions—e.g., $X_l = X_f$ for some $l \neq f$—in particular, at very small volumes. It is possible to take care of them, as we detail in Ref. [25]. However, to simplify the remainder of the discussion here, we assume no repetitions are present.

From datasets to wave function networks.—We now discuss how to translate the wave function snapshot datasets into a network structure. There are two key choices that have to be made: (i) the selection of a proper metric in the embedding space, that allows one to compute distances between data points, and (ii) a criterion to activate links between data points, based solely on their distances.

The choice of a proper metric is an important aspect of the approach. Taking inspiration from recent results in the context of classical and quantum statistical mechanics models, we use the Hamming distance [26,27]. Given two configurations X_i, X_j, such distance counts the number of spins that are aligned differently and reads

$$d(X_i, X_j) = \sum_{p=1}^{N} |s_p^i - s_p^j|. \tag{3}$$

The statistics of Hamming distances are related to arbitrary rank correlation functions between local degrees of freedom (i.e., s_k) [26]. Hence, they are sensitive to short-range and long-range correlations alike, which justifies their use as a similarity measure to define links between nodes. Specifically, we define a (geometric) network from our datasets by adopting the following procedure.

(1) Each point X_i in the dataset represents a node.
(2) If two nodes are at distance $d < R$, we draw a link.
(3) The distance R is chosen in a way that is dependent on the number of samples taken and reflects the typical value of distances for a given set of external parameters w. In particular, we define R as

$$R = \langle r_c \rangle = \frac{1}{N_r} \sum_{i=1}^{N_r} r_c(i), \tag{4}$$

where $r_c(i)$ is the distance between point X_i and its c-nearest neighbor.

We define the network formed by the collection of snapshots by the adjacency matrix: $A_{i,j} = \Theta(R - d_{i,j})$ [where Θ is the Heaviside step function and $d_{i,j} = d(X_i, X_j)$], which we dub wave function network.

It is worth mentioning that the choice of R is not arbitrary. One crucial aspect to consider is that the choice of R is bounded by the minimum and maximum distances (let us call such distances $d_{\min}$ and $d_{\max}$, respectively: $R < d_{\min}$ would generate a network with all the nodes isolated, while $R > d_{\max}$ would generate a featureless, fully connected network). The choice of R introduced in Eq. (4) can be considered as a natural one, as it naturally detects correlations at the average scale of the operator fluctuations (in our case, magnetization). In particular, we show that our main conclusions are independent of the choice of R for a certain range of values of c in Eq. (4). We discuss this in detail in Appendix C.

A. Network representation and correlations

At a naive level, one could expect that such WFNs simply reflect the intrinsic randomness of the wave function sampling; that is, they are ultimately generated by a Poissonian process. It turns out that this intuition is fundamentally incorrect.

In order to underpin the relation between network representation and correlations, we start by schematically illustrating the above procedure in Figs. 1(a)(i)–1(a)(iv). A graphical example of a network with spin-1/2 systems and cutoff radius equal to 1 (that is, only configurations differing by a single spin flip are connected) is depicted in Fig. 1(b): there, the black circle represents the Néel state that is connected to several other states by a single spin flip—and, thus, minimal distance—while it is not connected to any other states. This example allows us to intuitively connect physical properties to network properties: A wave function network that carries correlations will feature "hubs," that is, few states with many connections, and a lot of states with few connections. Conversely, a random "infinite-temperature" state will likely feature a majority of states with an intermediate number of neighbors, and will not feature either hubs or states with very few links.

The simple picture defined above is, *per se*, not particularly informative; however, it crucially sheds light on the classes of networks we can expect depending on how correlated the system is. This description of correlated states is reminiscent of what happens in several classes of scale-free networks: typically described by a scaling relation of the probability distribution P_k associated to the number of connections k of each node (or, as is more commonly called, the degree distribution), that follows a power-law distribution,

$$P_k \propto k^{-\alpha}. \tag{5}$$

Such a function monotonically decreases with k, and allows us to distinguish between the majority of nodes that have few links and the minority of those that have many links [see Fig. 1(b)]. While the prominence of hubs seems to be mostly relevant to ordered states, it is in fact a property that is even more robust in the presence of very strong correlations, such as, e.g., those emerging at quantum critical points. Conversely, networks representing random states will not be scale-free—and can be construed as ER networks—where the probability P_k of a node having k neighbors is approximately given by a Poisson distribution [13].

We emphasize that in the many-body regime, the number of snapshots N_r available from an experiment is typically insufficient to tomographically reconstruct the wave function, i.e., $N_r \ll 2^N$. The WFN construction aims at a characterization of a state that focuses solely on the most important (yet unknown) degrees of freedom in the system, and not its entire data structure. So, our method is conceptually different from tomographic methods, including those based on specific *Ansätze*.

B. Illustrative example: Quantum Ising model at equilibrium

Before discussing the experimental relevance of WFNs, we illustrate the emergence of scale-free networks in many-body systems by utilizing an example borrowed from equilibrium statistical mechanics. In Fig. 2, we show the degree distribution P_k obtained via sampling the partition-function snapshots of the 2D quantum Ising model on a square lattice, with samples in the z basis. The Hamiltonian reads

$$H = -\sum_{i,j}\sigma_i^z\sigma_j^z - g\sum_j \sigma_j^x. \tag{6}$$

It features a quantum phase transition at $g_c \approx 3.04$ separating a noncorrelated disordered phase (for $g > g_c$) from a ferromagnetically ordered state. The corresponding P_k is

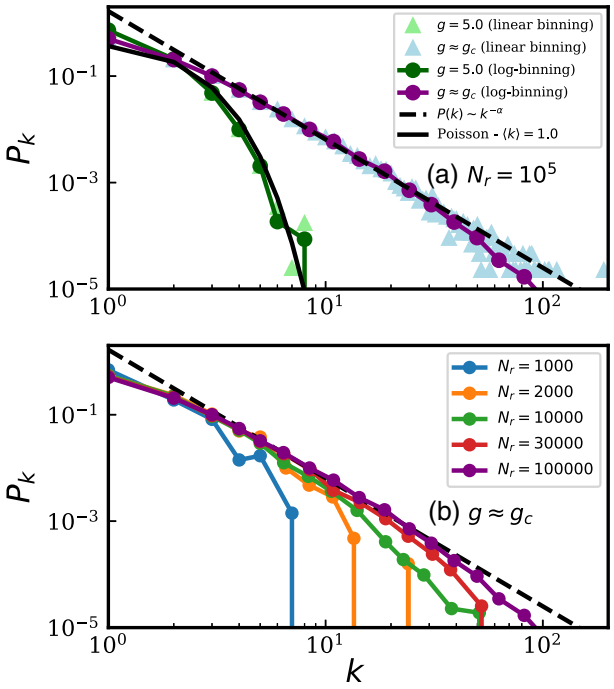

FIG. 2. Degree distribution P_k for the WFN of the ground-state quantum Ising model. Panel (a) shows P_k of the WFN with $N_r = 10^5$ nodes for $g = 5.0$ and $g = 3.04 \approx g_c$. In the paramagnetic region, the resulting network is compatible with a Poisson distribution (solid line, i.e., Erdős-Rényi network) with $\langle k \rangle = 1$. As expected, in the vicinity of the critical point, the WFN becomes scale-free, with $\alpha \simeq 2.4$ (dashed line). For comparison we compute P_k using both linear (triangles) and logarithmic (circles) histograms. Panel (b) shows P_k for different values of N_r for $g \approx g_c$ using logarithmic histograms, again with the scale-free distribution shown as a dashed line. In all cases we build the network using a cutoff $R = \langle r_1 \rangle$ [where $\langle r_1 \rangle$ is defined in Eq. (4)].

obtained by taking snapshots of the partition function, calculated via stochastic series expansion Monte Carlo simulations [28,29] for a system of $N = L \times L = 8 \times 8$ sites, at inverse temperature $\beta = 2L$, which in our calculations was high enough to observe convergence within statistical uncertainty of energy and squared magnetization, i.e., to reach the ground-state regime. Hence, the generated datasets correspond to the ground-state WF snapshots described above.

Figure 2(a) displays the results from $N_r = 10^5$ realizations. Deep in the paramagnetic phase, $g = 5.0$, there are only weak correlations: The corresponding network is very well described by a Poisson distribution with $\langle k \rangle = 1$. In the correlated regime, which is also the most entangled one, close to the phase transition, the network is described by a scale-free structure. We note that such a scale-free structure is unrelated to the absence of scale at criticality (scale-free networks can still be compatible with the presence of real-space finite scales [7]).

Finite-sampling effects.—Once the degree k of neighbors becomes a sizable fraction of the total size of the network, we observe deviations from a scale-free profile, as expected. Above this size, the network properties are influenced by limited sampling [30]. To inspect this, we plot in the lower panel P_k against k for various N_r. We observe that indeed the origin of the bending is due to the finite number of samples, and that the curves for various N_r are all compatible with a single power law, in this case, with exponent $\alpha \simeq 2.4$.

The plot discussed above gives a qualitative viewpoint on the presence of scale-free network here. In Appendix A, we provide a quantitative analysis based on very recently developed statistical tests [30], whose aim is precisely that of diagnosing the presence and strength of scale-free properties under limited sampling and finite volumes. The corresponding results strongly back the scale-freeness of the network structure in the vicinity of the quantum critical point. Additionally, changing the cutoff distance R used to build the WFN does not affect the power-law scaling behavior of P_k for k above a certain threshold k_c; see Appendix C for more details.

C. Origin of scale-freeness: A statistical mechanics viewpoint

Since early works on the World Wide Web [31], the mechanisms at the basis of scale-free network have been the subject of a large body of investigation. A particularly successful approach has been that of generative model, which has been able to shed considerable light on network structures and their properties. Unfortunately, we believe finding (and, importantly, fully understanding) the generative models at the basis of the quantum dynamics we are interested in is in general very challenging: The reason is that understanding this in the general case would be equivalent to having full control on the state probability distribution of arbitrary quantum many-body wave functions.

Nevertheless, it is possible to formulate an argument on why such scale-free structures shall appear in the presence of strong physical correlations, based on generic statistical mechanics considerations. The main idea, that we elaborate in detail in Sec. VI C 2, is that features of networks built with distance (or similarity) measures are inherently related to physical properties that are characterized by nontrivial (including scale-free) distributions. In our cases, such physical properties are related to cumulants of the variables used to create the dataset, namely, the magnetization.

Concretely, our consideration is drawn by explicitly exploiting the dependence of distance measures on physical variables, and, then, the dependence of WFN properties on the latter. For instance, as we shown in Sec. VI C 2, the sum of distances between a node i in the network and all other nodes is

$$d_i = \sum_{j=1}^{N_r} d(X_i, X_j) = \frac{N_r N}{2}\left(1 - \frac{1}{N}\sum_{p=1}^{N} s_p^i \langle s_p \rangle\right), \quad (7)$$

where $\langle s_p \rangle$ represent the physical estimator of the local spin at lattice site p. For a system with translation symmetry and in the limit $N_r \to \infty$, d_i is a function of the magnetization of the corresponding snapshot:

$$m_i = \frac{1}{N}\sum_{p=1}^{N} s_p^i. \quad (8)$$

Hence, the probability distribution associated to distances $P(d)$ is directly related to $P(m)$. The fact that the distribution $P(m)$ diverges the predictions based on the central limit theorem in the vicinity of critical points [32,33] allows then can be used as an indication that the networks built from the statistics of distances is incompatible with the Erdős-Rényi one found in random phases.

In Appendix B, we elaborate further such argument to take into account properties of the WFNs. In such a context, the key reason why probability distributions associated to WFNs can exhibit nontrivial behavior (including scale-freeness) can also be traced back to the way certain network features are connected to strongly correlated physical variables. More concretely, we describe how the strength distribution $P(g)$ is related with the distributions of sums of the physical variables s_p.

Based on these arguments, at equilibrium, we expect the same dichotomic (scale-free versus Erdős-Rényi) structure to appear generically in models that feature both weak- and strong-correlated regions. The key question we address below is whether such structures are purely theoretical constructions or if they can indeed be representative of the intricate dynamics taking place in quantum simulators that are (i) off equilibrium, open, and—a key difference from

simulations—(ii) inherently probed with very high but not 100% fidelity.

III. EXPERIMENTAL OBSERVATION OF ERDŐS-RÉNYI AND SCALE-FREE WAVE FUNCTION NETWORKS

A. Experimental data and analysis of the network

We now discuss the network structure of quantum simulation experiments. We analyze a recent experiment that focuses on the quasiadiabatic state preparation of a large antiferromagnetic state using a Rydberg quantum simulator [10]. This protocol plays a fundamental importance in quantum simulation and computing, and is very widely employed in atomic physics platforms. In addition, it typically features both regimes of no correlations (short times) and of strong correlations, enabling us to test predictions based on both Erdős-Rényi and scale-free networks. Below, we summarize the main features of the experiment that have been reported in Ref. [10].

The experiment consists of arrays of laser-cooled Rb atoms, individually trapped into optical tweezers separated by a distance a. Each atom can be considered as a pseudospin, the ground state being $|\downarrow\rangle$ and a Rydberg state state being $|\uparrow\rangle$. Initially all the atoms are prepared in $|\downarrow\rangle$. The atoms are then laser excited to Rydberg states via a two-photon transition, so that the effective time-dependent Hamiltonian describing the dynamics reads:

$$H(t) = \hbar\delta(t)\sum_i n_i + \frac{\Omega(t)}{2}\sum_i \sigma_i^x + \sum_{ij} J_{ij} n_i n_j, \quad (9)$$

with $n_i = (\sigma_i^z + 1)/2$, and σ_i^α Pauli matrices at the site i. Here, we have that $J_{ij} = C_6/r_{ij}^6$ as the atoms interact via the Van der Waals interaction. This quantum spin model exhibits both paramagnetic and antiferromagnetic phases in its ground state; for a schematic phase diagram, see Fig. 3(a).

In the experiment a dynamical process has been implemented which, upon varying slowly $\Omega(t)$ and $\delta(t)$ over time, transforms an initial paramagnetic state into an antiferromagnetic one, as depicted in Fig. 3(a). The adiabatic theorem guarantees that such a transformation is possible for ground states of systems with a nonzero gap whenever the parameter variations are sufficiently slow. Close to a continuous quantum phase transition, however, the gap closes for a thermodynamically large system and excitations are generated unavoidably. Importantly, the celebrated quantum Kibble-Zurek (QKZ) mechanism predicts that this defect generation, and on a more general level the dynamical

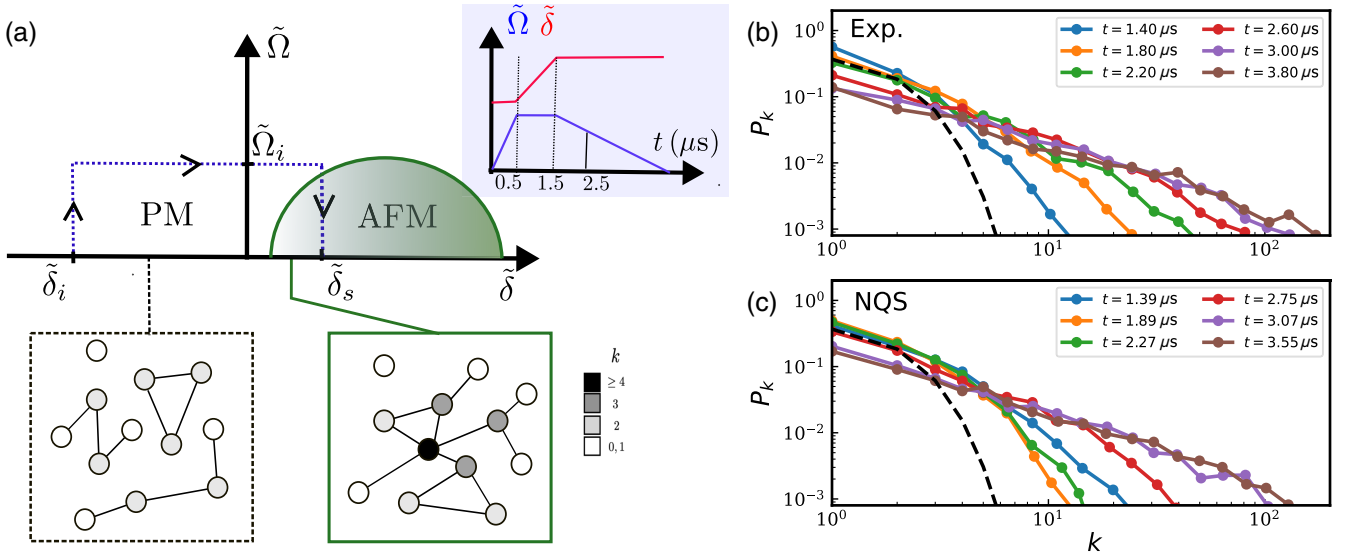

FIG. 3. Observation of scale-free wave function networks in Rydberg quantum simulators. Panel (a) shows a schematic ground-state phase diagram and the quasiadiabatic state preparation scheme. The inset shows the sweep shape, and the corresponding trajectory is represented by the dashed lines in the phase diagram. In the paramagnetic (PM) regime, one expects a network description compatible with a Rényi-Erdős network, while in the vicinity to the antiferromagnetic (AFM) region, that contains the Kibble-Zurek regime, a scale-free network structure is expected with power-law degree distribution P_k, as illustrated by the network structures. Panel (b) presents the P_k versus k of the experimentally observed wave function networks for a square lattice with $L = 8$. At short times, i.e., before crossing the phase transition, the distribution decreases exponentially (similar to Erdős-Rényi degree distribution with $\langle k\rangle \simeq 1$, represented by the dashed lines in the graphs). At later times ($t > 3$ μs), we observe a power-law decay over 2 orders of magnitude, limited only by a bending that is due to a finite value of N_r. Panel (c) shows NQS simulations of this quasiadiabatic protocol for the same square lattice. The scale-free behavior of P_k is again observed until one is sensitive to the effects of finite sampling. We note that the value of the decay exponent is $\alpha < 2$, signifying very stable wave function network properties, that will be discussed later in the presence of defects. For all cases we considered WFNs with $N_r = 2500$ nodes.

property itself of crossing such a transition, displays universal behavior controlled by the underlying quantum phase transition [34–36]. In the context of two-dimensional systems, this has recently been described at the theoretical level [37], and signatures have been observed in Rydberg experiments [38]. For a finite-size system, such as the ones we deal with here, the gap remains always finite. Because of this, a crossover from a QKZ regime toward an adiabatic regime emerges upon lowering the velocity of the ramp [37]. In the experiment an antiferromagnetic ordering pattern has been achieved with a correlation length of the order of the system diameter, so that it is to be expected that the system resides in the crossover regime between QKZ and adiabaticity.

In what follows we will support the experimental data with numerically exact theory calculations, which will be key in a later stage in the cross-certification of the quantum simulator output. For that purpose, we will use NQSs, which have been recently introduced as novel class of variational wave functions for the quantum many-body problem [11]. Most importantly for the purpose of this work, recent paramount advances have pointed out a route to numerically calculate quantum many-body dynamics in interacting two-dimensional quantum matter beyond what is achievable with other state-of-the-art methods [12,37]. For details on the utilized numerical method, see Refs. [12,37] and Appendix D.

Contrarily to the work in Ref. [10], we consider for our network analysis here two types of datasets. In the first one, we use postselected data without any defects in the array; i.e., each trap contains exactly one atom. In the second one, we instead consider datasets including a mean number of defects of ∼3%, coming from an imperfect assembly of the atomic array [39]. The purpose of this second choice is that it will allow us to make quantitative statements on the resilience of scale-free structures and, most importantly, on their significance in terms of information—and, thus, complexity—content.

Scale-free and Erdős-Rényi networks.—In Fig. 3(b), we plot the distribution P_k for defect-free experimental data for square lattices of size 8×8 and $N_r = 2500$ at different times. We identify two regimes.

(A) At short times $t = 1.52$ μs, P_k decays exponentially with k, and its distribution resembles the one of a random ER network with $\langle k \rangle \simeq 1$. This indicates that only limited correlations in the z-basis measurements are present in the system.

(B) Upon approaching the quantum phase transition ($t \sim 2.6$ μs) and at later times, the distribution changes drastically. In particular, we observe the emergence of a stable power-law profile with $\alpha < 2$ over almost 2 orders of magnitude, until at large k finite sampling with $N_r < \infty$ introduces an inevitable cutoff in the form of an exponential decay. This phenomenology is characteristic of scale-free networks.

In Fig. 3(c) we include as a comparison numerically exact theoretical results for P_k by means of NQS simulations. We utilize the same system parameters and number of samples as for the experimental data. The simulations capture the exact same qualitative pattern described by the experiment, already indicating that, for the depicted timescale of the experiments, the effect of dissipation on the full many-body wave functions is likely to be negligible, and validating the microscopic modeling at a quantitative level.

As depicted in Fig. 3, at large k, deviations from a power-law scaling become appreciable. More precisely, for a finite N_r, deviations of such behavior are expected to appear above an upper degree $k_{\max}$ [30]. To properly take into account such size effects, we consider a network finite-size scaling (FSS) analysis of WFNs obtained for different values of N_r and t; see Appendix A for more details. The FSS analysis clearly shows that WFNs built at initial times are not scale-free. However, as we approach the time associated with diverging correlations (the critical region), the analysis supports the conclusion that the WFNs do correspond to scale-free networks. Interestingly, in the late time regime where the system features correlations whose lengths exceed the system size, our analysis does not conclusively support the classification of the corresponding WFNs as scale-free, despite their heavy-tailed P_k

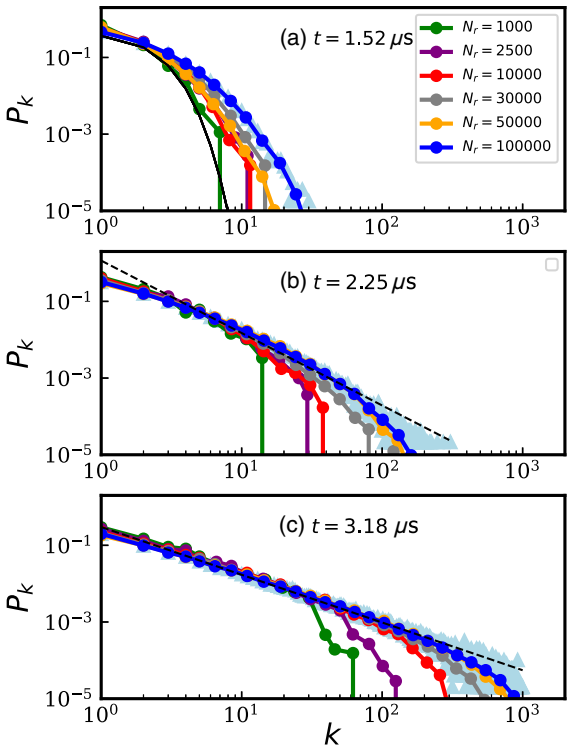

FIG. 4. Dependence of the degree distribution P_k with the total number of nodes in the WFNs N_r at the different values of t. In the scale-free regime the maximum size of the WFN at $k_{\max}$ exhibit a strong dependence with N_r. The WFNs are obtained with datasets generated by NQS simulations of Rydberg experiments.

distribution. This might be signaling the onset of dissipation, or that further statistics is needed for a conclusive test. As an example of the aforementioned behaviors, we show in Fig. 4 the distribution P_k for three reference cases of times t by means of data obtained using NQS. Both qualitatively and quantitatively, P_k exhibits consistent features in all regimes: For ER graphs [Fig. 4(a)], increasing the number of nodes N_r yields essentially the same structure of the network (keeping $\langle k \rangle \simeq 1$). For the case of scale-free networks, see Fig. 4(b), increasing the number of samples has the effect of enlarging the regime in k of power-law behavior shifting the eventual bending, i.e., the deviation from the scale-free structure at large k, to larger and larger k. For long times, a similar behavior is observed (this might point to the fact that the graph is not unambiguously recognized as scale-free only because of limited statistics).

Robustness of quantum simulator outputs.—We observe that at late times $t > 3$ μs, the exponent α of the power-law tail in P_k satisfies the condition $\alpha < 2$ [see Fig. 4(c)]. As is known from network theory, scale-free networks with such an exponent α exhibit very robust information content with respect to perturbations. We identify such a robustness also in the experimental data. Specifically, as can be seen in Fig. 5, the experimental datasets with defects in the atomic array capture the same scaling behavior as without defects. In analogy to network theory, this analysis provides an interesting tool to characterize the robustness of quantum simulators based solely on their outputs, whenever they are described by scale-free or ER networks. An important comment is in order: Such small values of the power-law exponents are typically characteristic of finite networks: this is compatible with our theory, since we know that, in the infinite sampling limit $N_r \to \infty$, our network becomes infinitely large and it will be unavoidable to generate repetitions of the same snapshot. Such repetitions, however, we have excluded from the beginning and would require an adaption of our approach by means of certain weighted networks.

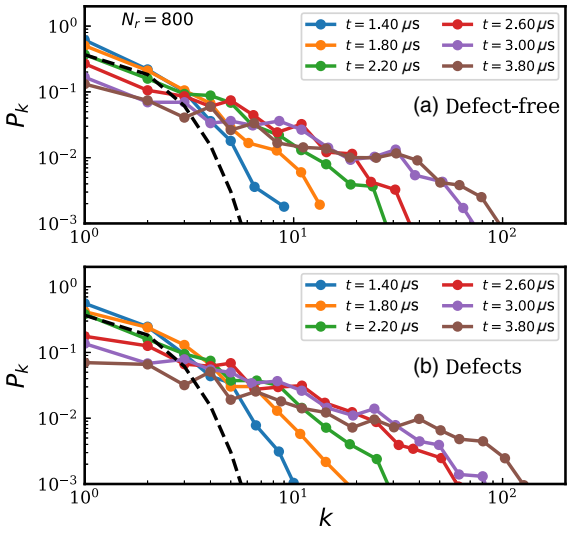

FIG. 5. Robustness of quantum simulator outputs. Comparison of the degree distribution P_k of experimental WFNs generated (a) without defects, and (b) with a mean density of defects of ∼3%. We consider $N_r = 800$ in both cases. The results are qualitatively equivalent.

B. Theory of wave function networks evolution over quasiadiabatic state preparation

The scale-free and ER WFN phenomenologies we observe in both experiment and numerical simulations are not tied to the specific problem we explore here, but, as we argue in this section, are generic features of quasiadiabatic state preparation protocols. Starting from an uncorrelated product state, it is natural to expect that at short times one may typically find random networks of wave function snapshots, i.e., networks with ER-type structures. An example of such an instance is the case covered in this work, where we start from a product state with spins aligned in the z direction. At short times the unitary dynamics will generate some weak but noticeable superposition of other configurations with a few flipped spins, which we expect to appear similar to the presence of local fluctuations such as those caused by dissipation or thermal fluctuations. These are inherently random, and should therefore yield an ER network, with Poisson-like degree distribution. For $N_r \ll 2^N$, such a process is expected to generate a very sparse network with $\langle k \rangle \simeq 1$, due to the fact that the average distance between configurations is roughly a constant.

Upon approaching the quantum phase transition, we observe the emergence of a scale-free network structure. The basic mechanism behind this can be understood upon inspection of the introduced metric in Eq. (3), which is used to impose the fundamental underlying structure on our datasets. The network structure, which we probe through P_k, is generated by correlations in distances between different snapshots. Only in the case where such distances are correlated is it possible to find a power-law distribution P_k of nodes having a connectivity k. As we discuss in the following, these correlations in distances between nodes in the network might be linked to the real-space correlation length in the system. Upon entering the quantum phase transition regime the system develops a large correlation length of the order of the system diameter due to the almost adiabatic dynamics generated through the experimental protocol.

From previous works on data analysis of snapshot measurements, it is expected that such large real-space correlation lengths yield Pareto distributions, i.e., power-law distributions, of distance measures in the dataset [26,27]. In this light our observations of a scale-free network structure in P_k appears natural in particular because P_k quantifies correlations between distances of network nodes. We note that the scale-free property of a scale-free network solely concerns P_k—indeed, other network properties may carry

information that reflects the presence of a finite correlation length.

Once the quantum phase transition regime is reached, the system is effectively described by a large real-space correlation length. Let us note that the dynamical behavior of the considered quantum spin model is expected to be potentially much richer as compared to the mostly studied case of dimension $D = 1$. Upon entering the broken-symmetry phase, the system will eventually thermalize, implying an infinite correlation length. In analogy to classical systems, the temporal process of generating a long-range ordered state is typically associated by coarsening and phase-ordering kinetics [40], which also comes along with universal power-law behavior. In turn, this means that upon crossing the quantum phase transition it is expected that the correlation length grows further in time also deep in the broken-symmetry phase. When linking large real-space correlation lengths with scale-free network structures, this would imply that the scale-free network structure might survive also beyond the quantum phase transition region. This underlines the universal character of the data structure dynamics observed in the experiments. The reasoning above applies also to first-order phase transitions, as long as the correlation length at the critical point is larger than the system diameter (so that, in fact, the correlation functions in the system are not able to discern differences with respect to a continuous transition).

IV. APPLICATION 1: KOLMOGOROV COMPLEXITY OF WAVE FUNCTION SNAPSHOTS

The output of a quantum simulator obtained via wave function snapshots is, *per se*, a classical object. How complex must a classical computer program be in order to reproduce such output? This is quantified by the so-called Kolmogorov complexity [20,21].

For generic strings, computing the KC is an NP-hard problem. The same holds true for generic graphs, where the KC is quantified by the Haussdorf dimension [41]. This implies that computing the Kolmogorov complexity of wave function snapshots is an extremely challenging task that cannot be undertaken in general.

However, as noted in the previous sections, quantum simulators often generate scale-free networks: For these, there exist known nonparametric learning algorithms that allow us to estimate the intrinsic dimension of the data points, and thus, the KC, in a manner that does not depend on scale. In particular, we utilize the 2-NN algorithm [42,43], that has already been applied in the determination of critical properties of both classical and quantum statistical mechanics partition functions [26,27].

The starting point is to consider, for each point X_j in our dataset, the distances to its first and second nearest neighbor, $r_1(X_j)$ and $r_2(X_j)$, respectively. Under the condition that the dataset is locally uniform in the range of second-nearest neighbors, it has been shown in Ref. [42]

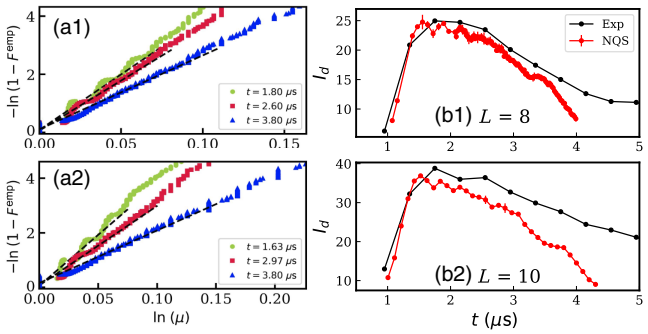

FIG. 6. Complexity scaling in quantum simulators. (a1),(a2) Cumulative distributions against $\mu = r_2/r_1$ for selected times. Panels (a1) and (a2) show results for experimental and NQS simulations datasets, respectively. The quality of a description in terms of Pareto distribution (lines) increases as a function of time, for both simulation and experiment. (b1),(b2) Time dependence of the intrinsic dimension I_d along the quasiadiabatic time evolution for both experimental and simulated datasets. For $t > 2$ μs, the complexity of the WFN is a monotonously decreasing function of time in both experiments and simulations, capturing the emergent simplicity (decrease of degrees of freedom) that is expected from the emergence critical behavior.

that the cumulative distribution function F^{emp} of $\mu = r_2(\boldsymbol{x})/r_1(\boldsymbol{x})$ obeys

$$I_d = -\frac{\ln\left[1 - F^{\mathrm{emp}}(\mu)\right]}{\ln(\mu)}, \tag{10}$$

where I_d is the intrinsic dimension of the dataset. The intrinsic dimension quantifies the number of degrees of freedom required to capture the information content of the dataset. While this is in principle a length-scale-dependent property, our estimator directly focuses on the physically relevant distance that is determined by the sampling of the many-body wave functions.

In Fig. 6(a), we depict the relation between F^{emp} and μ obtained from [Fig. 6(a1)] experiments and [Fig. 6(a2)] NQS simulations. In both cases, and for all times considered, the distribution is compatible with Pareto (additional oscillations appear at short times, likely due to the very simple structure of the network). These results guarantee the applicability of the 2-NN approach [42].

In Fig. 6(b), we show the time dependence of the KC as measured by the intrinsic dimension across the ramp. Both experimental and simulation data clearly display two regimes. (i) Up to 2 μs, the complexity increases. This effect is trivial: The initial state is very close to a product state along the z direction, so that at short times there is just one single dominant snapshot as the measurement outcome. The unitary evolution will necessarily generate additional correlations afterward, thus increasing complexity. (ii) From 2 μs onward, the complexity becomes a monotonously decreasing function of time. This second regime is a manifestation of the emergence of universal

behavior while crossing a phase transition. Following quasiadiabatic dynamics, the correlation length monotonously increases as a function of time: This implies that, in order to describe network properties, fewer variables are actually required—at equilibrium, these would just be the critical exponents and the amplitudes of correlation functions. The observations above are thus a direct manifestation of the emergent simplicity associated to universality at critical points [27] and represent, to the best of our knowledge, the first experimental demonstration of the link between complexity and quantum critical behavior.

We note that, after some time, NQS simulations predict a faster decrease of complexity with respect to the experimental data. We attribute this to the fact that the simulations can only partly keep track of the time evolution: The neural network structure utilizes a smaller number of effective variables, compatible with a decrease of KC, with respect to those that are describing the time evolution realized in experiment.

V. APPLICATION 2: CROSS-CERTIFICATION BASED ON NETWORK PROPERTIES

One of the key challenges for quantum computers and simulators is to verify their correct functioning or to certify the validity of their outcome. One basic idea in the field is cross-certification, which consists of directly comparing the output of one quantum machine with another—either quantum or classical. Recent protocols based on random unitary circuitry, aiming to compare the full ground-state wave functions, have been experimentally demonstrated to be superior to tomographic methods [16,44]. However, resources still scale exponentially with system size, making the present methods inapplicable to large devices.

Here, we take a complementary angle and focus on a comparison based on wave function snapshots that take into account the maximum amount of extractable information with currently available resources. At the formal level, our goal is to compare two distributions in the limit where $N_r \ll 2^N$, i.e., which is relevant to experiments exploring many-body problems (oppositely, for $N \simeq 10$, it is possible to reach by brute force the regime $N_r \simeq 2^N$ [44]). Clearly, a configuration-by-configuration comparison sampled by two distributions is meaningless unless the states are very close to a product state; for generic states, the probability of sampling the same set of configurations will scale exponentially to zero with system size.

The network representation we use allows us to bypass this limitation. Specifically, we wish to compare two WFNs obtained either by two experiments or by an experiment and a simulation; see Fig. 7(a). For the concrete case considered here, let us point out that the numerical simulation by itself is a formidable challenge, which we target again by means of the NQS approach [11,12,37]; see also Appendix D for details. Finding and quantifying similarities between two networks is a problem largely explored in different applications of network theory and is particularly useful for datasets that cannot be distinguished by direct inspection or low-order correlations [45]. In our case, such comparisons between networks are directly related to the choice of metric used to define the WFN. For scale-free WFN, this is particularly suitable, as we are guaranteed to have chosen a metric distance capturing correlations in the system.

As a simple and efficient way to compare experimental and simulated WFNs, we check the hypothesis that the corresponding degree distributions are equal by employing a nonparametric test, known as the Epps-Singleton test [24].

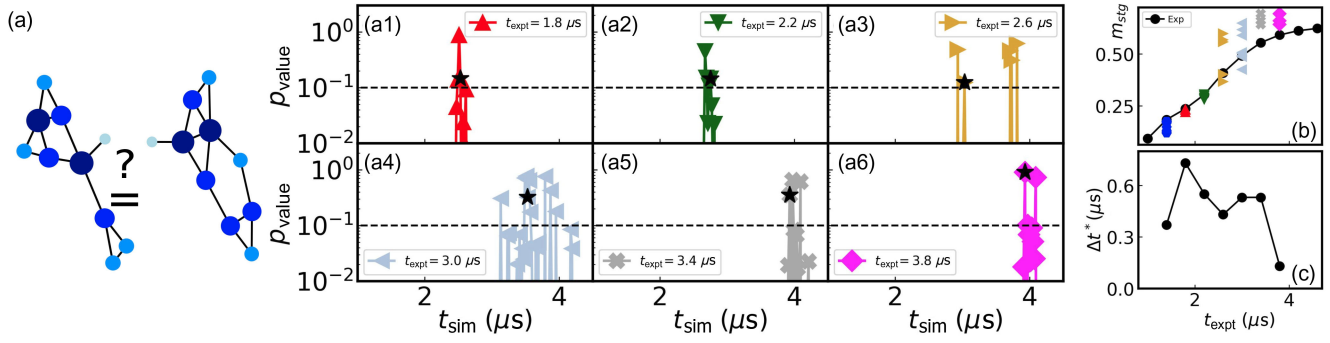

FIG. 7. Comparing the experimental and simulated WFNs, as illustrated in (a). In particular, we consider the Epps-Singleton two-sample test to check the hypothesis that the experimental and simulated degree distribution P_k are equal. For each experimental time t_{expt}, we consider ES tests with the different simulated results at times t_{sim}. Both WFNs have $N_r = 2500$ nodes, and we choose a cutoff distance $R = \langle r_1 \rangle$ to generate them. Panels (a1)–(a6) show the corresponding p_{value} as a function of t_{sim}: results with $p_{\mathrm{value}} > 0.1$ (marked by the dashed lines) are interpreted as statistically significant. To cross-check our analysis, we also consider in (b) the order parameter m_{stg} as a function of t_{expt}. Each simulated result corresponds to the different times t_{sim} for which $p_{\mathrm{value}} > 0.1$. Such an analysis allows us to identify t^*_{sim}, the times where the best agreement exists between simulation and experimental data; the results corresponding to t^*_{sim} are marked as the black star points in (a1)–(a6). Finally, panel (c) shows the corresponding time shift, $\Delta t^* = t^*_{\mathrm{sim}} - t_{\mathrm{expt}}$, between experiments and NQS simulations (see text).

The latter allows us to identify, with statistical significance, when two WFNs are different. In the following, since we will employ this test to establish the identity of two WFNs, we will take as statistically significant proof of our claim cases in which the p value of the test takes values $p_{\text{value}} > 0.1$, i.e., in which both experimental and simulation data are compatible with a common probability distribution.

The results are summarized in Fig. 7. In Figs. 7(a1)–7(a6), we present the corresponding p_{value} of the ES tests obtained by comparing experimental data at a given time t_{expt} with the simulation data over a given time window (i.e., $1 < t_{\text{sim}} < 4.2$ μs). This allows us to identify time windows where the quantum and classical simulators can be cross-certified with statistical significance in terms of the maximum amount of information available from their wave function networks. Interestingly, we note that the cross-certification agreement occurs at time windows that are shifted from the actual experimental times presented in the legends of Figs. 7(a1)–7(a6).

The fact that the cross-certification agreement of quantum systems occurs at a time t_{expt} that is different from times considered in the simulations t_{sim} can be attributed to miscalibrations of the parameters of the Hamiltonian [e.g., $\Omega(t)$ and $\delta(t)$]. Similar observations are found when comparing experiments with simulations of physical observables based on matrix product states [10]. Although such miscalibrations do not affect the actual physics, quantifying the corresponding time shift is essential for the cross-validation of the quantum simulators.

In general, we find that the ES test can provide, for a given t_{expt}, multiple candidate simulation times t_{sim} for which $p_{\text{value}} > 0.1$; see Fig. 7. In order to finally select across these multiple candidates, we perform a second test by computing for each of the candidates an independent second quantity. Here, we consider the staggered magnetization $m_{\text{stg}} = \sum_{i_x,i_y} (-1)^{i_x+i_y} \sigma^z_{i_x;i_y}$, where we included the results for all the candidate simulation times [see Fig. 7(b)]. Finally, we choose the candidate t^*_{sim}, in which the simulated results for the order parameter are closest to the experimental data. As one can see from Fig. 7, up to a time of $t_{\text{expt}} = 3.0$ μs, we find that this procedure is capable of cross-certifying the experimental and theoretical data within the achievable accuracy, which is limited, for instance, by the finite time grid of the theoretical data. For intermediate times $3.0 \lesssim t_{\text{expt}} \lesssim 4.0$ μs, small deviations start to emerge, whereas for times $t_{\text{expt}} \gtrsim 4.0$ μs, the cross-certification fails, which could be caused by dissipation effects in the experiment that are not included in the theory calculation, or by a decreasing accuracy of our variational computation (similarly to what is observed in the complexity scaling).

This scheme further defines an optimal time shift $\Delta t^* = t^*_{\text{sim}} - t_{\text{expt}}$ for the experimental data. Figure 7(c) shows the estimated values of Δt^*. Importantly, in the time interval that the quantum simulator can be cross-certified, we identify a small time dependence of Δt^*, which has not been addressed previously. We note that the procedure does not work well for $t < 1.5$ μs, as expected: There, the network is not scale-free yet, so a direct comparison can only provide some rough qualitative guidance.

VI. CONCLUSIONS AND OUTLOOK

We have introduced a network theory framework to interpret the maximum amount of information extractable from quantum simulators—wave function snapshots. Remarkably, such networks can become scale-free for strongly correlated states of matter, and are of direct experimental relevance, as we demonstrate with data from a large-scale Rydberg atom array experiment. We have illustrated the power of network description with two applications: demonstrating the scaling of complexity across a quantum phase transition during Kibble-Zurek scaling, and cross-certifying the wave function of a quantum and classical simulator up to system sizes that have never been attained previously.

Our work opens up a series of research directions based on a transfer of methods and concepts between network and quantum science. At the big picture level, it would be important to determine to what extent Erdős-Rényi and scale-free networks are able to characterize quantum simulators and computers. While our framework provides strong evidence that it works at and close to equilibrium, the structure of wave function networks in genuinely out-of-equilibrium situations is presently completely unknown. Understanding network properties corresponding to such dynamics might provide qualitative insights into how equilibrium is established at the wave function level, complementing current efforts focusing on observables, and providing direct links between dynamics and Kolmogorov complexity. Going beyond the case of unitary dynamics, understanding the role of dissipation might help characterize the stability of quantum dynamics to noise, which will ultimately always kick in and—very likely—imprint an Erdős-Rényi structure onto the system wave function.

In addition to conceptual insights, our framework is ideally suited to developing scalable quantum information tools. Examples range from improving cross-certification methods for large-scale experiments, where computing direct wave function overlap is impractical, to applying it to datasets obtained through generalized quantum measurements (e.g., positive operator-valued measure), from which information about basis-independent properties can be extracted. In the latter case, it will be particularly interesting to investigate the prospects of accessing genuine quantum properties. On a broader level, we believe that the parallelism between two very active, but so far disconnected, fields could be an ideal playground for developing new insights into how information is associated to many-body phenomena.

ACKNOWLEDGMENTS

We thank G. Bianconi, J. Grilli, M. Marsili, R. Panda, R. Verdel, V. Vitale, and P. Zoller for insightful discussions. The work of M. D. and A. A. was partially supported by the ERC under Grant No. 758329 (AGEnTh), and by the MIUR Programme FARE (MEPH). M. D. further acknowledges funding within the QuantERA II Programme that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101017733. D. B. acknowledges support from MCIN/AEI/10.13039/501100011033 (RYC2018-025348-I and NextGenerationEU PRTR-C17.I1). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 853443). Moreover, the authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC) [46]. This work is supported by the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 817482 (PASQuanS), the Agence Nationale de la Recherche (ANR, project RYBOTIN), and the European Research Council (Advanced Grant No. 101018511-ATARAXIA). M. S. has received support by Helmholtz Initiative and Networking Fund, Grant No. VH-NG-1711.

APPENDIX A: ANALYSES OF THE EMPIRICAL DEGREE DISTRIBUTIONS: FINITE-SIZE SCALING

Recently, the emergence of power laws in empirical networks has been questioned. One argument is that alternative (non-scale-free) distributions can better describe the degree distribution of networks. Rigorous statistical tests indeed confirm that this is the case for several networks [47] (previously taken as scale-free), although essential aspects (finite-size effects of the network [30] or independence of the empirical observations [48]) are typically ignored in such analyses. To line up our conclusions with such recent observations, in this appendix we discuss in detail the analyses we employ to prove that in the vicinity of quantum critical points WFNs become scale-free.

In a network containing N_r nodes, the degree can be at maximum equal to $N_r - 1$. Hence, while for an infinite network ($N_r \to \infty$) one might observe a power-law behavior in the scaling regime $k \geq k_{\min}$, for a finite network, deviations of such behavior are expected to appear above an upper degree $k_{\max}$. In this way, a more precise definition of a scale-free network takes into account the upper-degree cutoff $k_{\max}$ in the form of the finite-size scaling hypothesis,

$$P(k, N_r) = k^{-\alpha} f(k/k_{\max}), \tag{A1}$$

where N_r represents the number of nodes of the network, α is the power-law exponent, and the upper-degree cutoff $k_{\max}$ is related to N_r: $k_{\max} \sim N_r^{-d}$, d being a negative exponent that defines the finite-size effects of the network. The key aspect of the FSS hypothesis is that $f(x)$ is a universal function of $x = kN_r^d$ that, for $x \to 0$, is $f(x) = \text{const}$, in such a way that the power-law behavior of $P(k, N_r)$ is recovered for an infinite network.

The question of whether a WFN is a scale-free network is then assessed by confirming the validity of the FSS hypothesis. To keep our presentation self-contained, we highlight below the main aspects of the FSS analyses; for a thorough discussion, we refer the reader to Refs. [30,49].

The FSS analyses start with a WFN built with a cutoff R and N_r nodes. We then determine the scaling region where the corresponding degree distribution $P(k, N_r)$ is fitted by a power-law function, i.e., $k \geq k_{\min}$. We use the approach of Clauset and co-workers [47,50], which determines the optimal value of $k_{\min}$ by creating a power-law fit starting from different values of $k_{\min}$ in the $P(k, N_r)$, and then selects the one that results in the minimal Kolmogorov-Smirnov distance between the empirical data and the fit. We implement such a scheme with the code of Ref. [51].

As a second step, we generate a set of subnetworks of different sizes $n < N_r$ by removing n_{rem} nodes randomly among the N_r nodes of the original WFN. For each value of n, we construct 100 subnetworks and analyze the properties of the averaged $P(k, n_r)$ (or the corresponding cumulative distribution). The analysis is performed for nodes with $k \geq k_{\min}$. Hence, the number of nodes n_r entering in Eq. (A1) corresponds to the averaged number of nodes with degree $k \geq k_{\min}$. Furthermore, we consider the cumulative distributions $C(k, n_r) = \int_k^\infty P(q, n_r) dq$ [instead of $P(k, n_r)$], which allow us to mitigate errors caused by the binning used to compute $P(k, n_r)$. For scale-free networks, $C(k, n_r)$ is also described by the FSS hypothesis,

$$C(k, n_r) = k^{-\gamma} \tilde{f}(k n_r^d), \tag{A2}$$

with exponents $\gamma = \alpha - 1$. We now discuss the approaches we use to assess whether $C(k, n_r)$ is described by the FSS hypothesis.

1. Data collapse

We first consider the *data collapse* technique. As is customarily done in statistical physics to analyze the behavior of physical observables near phase transitions, we select fitting parameters γ and d by attempting to perform a data collapse of results of $C(k, n_r) k^\gamma$ versus $k n_r^d$ for different subnetworks. In practice, the basic idea is to try to fit results for different n_r with the same master curve. The fitting quality allows determining the exponents γ and d and establishing how well the FSS hypothesis describes the empirical data. In particular, we follow the steps described in Ref. [52] to define the master curve and the

quantity S measuring the quality of collapse. We implement the data collapse with code described in Ref. [53].

2. Moment ratio tests

We also consider the scaling of the *moment ratios* of $C(k, n_r)$. Such an analysis is based on the fact that the ith moment $\langle k^i \rangle$ of a scale-free network scales as

$$\langle k^i \rangle = \int_{k_{\min}}^{\infty} dk k^{i-1} k^{-\gamma} f(k n_r^d) \propto n_r^{-d(i-\gamma)}, \qquad \text{(A3)}$$

for $i > \gamma$; instead for $i \leq \gamma$, $\langle k^i \rangle$ converges to a constant. As a result, the ratio between two moments scales as

$$r_i = \langle k^i \rangle / \langle k^{i-1} \rangle \propto n_r^{-d}, \qquad \text{(A4)}$$

when $i - 1 > \alpha$. Likewise, for a scale-free network, r_i scales with the number of edges n_E as

$$r_i \propto n_E^{-d_E}, \qquad \text{(A5)}$$

where the exponent d_E is related to d (more on this below). Let us stress that n_E is the number of edges in the scaling regime, i.e., the total number of edges that are connected with nodes with $k \geq k_{\min}$. The log-log plot of r_i with n_r and n_E for different i is then used to probe whether a network is scale-free [30].

3. Results

a. Two-dimensional quantum Ising model

We start considering WFNs generated by the ground state of the 2D quantum Ising model in the vicinity of the quantum critical regime; we consider $R = \langle r_1 \rangle$ (and $R = \langle r_2 \rangle$). The power-law fitting [47,50] selects $k_{\min} = 3$ ($k_{\min} = 9$) and in both cases $\alpha > 2.0$, i.e., $\alpha \approx 2.35$ ($\alpha \approx 2.54$). As shown in Figs. 8(a) and 8(b), the cumulative distributions for different values of n_r result in a high-quality collapse, and the moment ratios r_i are well described by parallel lines for different values of i. More quantitatively, the main outcomes of the analysis are the following.

(1) We obtain $S < 1.0$ in both cases: $S = 0.92$ ($S = 0.77$) for $d = -0.50 \pm 0.04$ and $\gamma = 1.4 \pm 0.1$ ($d = -0.42 \pm 0.09$ and $\gamma = 1.48 \pm 0.03$); such results are consistent with our estimation of α obtained with the fitting of the degree distribution,

(2) The scaling exponents d and d_E are compatible (within error bars), which is consistent with the fact that the scaling of $\langle k \rangle \sim n_r^{d/d_E - 1}$ is independent of n_r for a scale-free network with $\alpha > 2$.

Moreover, we observe that the exponents d and γ (which dictates how the uppper-degree cutoff scale with network size, $k_{\max} \sim n_r^{-d}$) are well described by the empirical relation $d \approx -(\gamma + 1)^{-1}$, which is consistent with a scale-free network [30,49]. For those cases, thus, the finite-size analysis shows clear evidence that the WFNs correspond to a scale-free network: According to the classification proposed in Ref. [30], those are *strong* scale-free networks.

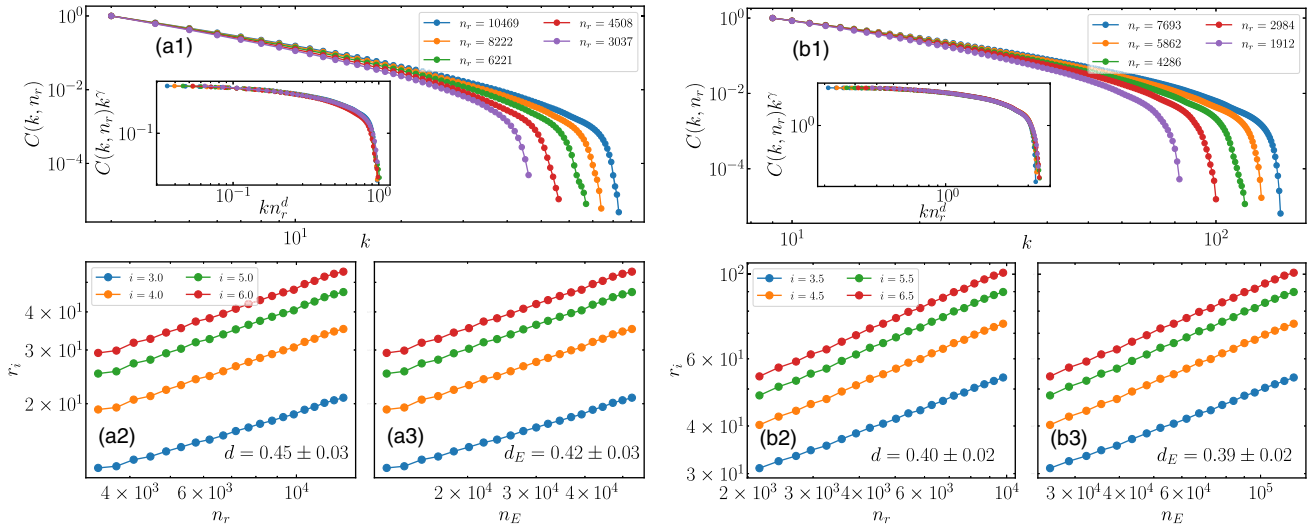

FIG. 8. Panels (a) and (b) show the FSS analyses for the WFNs generated at the critical point of the quantum Ising model for $R = \langle r_1 \rangle$ and $\langle r_2 \rangle$, respectively. Panel (a1) [(b1)] shows the cumulative distribution, while the inset shows the data collapse. The best collapse occurs for $d = -0.50 \pm 0.04$ and $\gamma = 1.4 \pm 0.1$ ($d = -0.42 \pm 0.09$ and $\gamma = 1.48 \pm 0.03$). Panels (a2) and (a3) [(b2) and (b3)] show the scaling of the moment ratio r_i with the number of nodes n_r and the number of n_E, respectively, for different values of i. The exponent d (and d_E) corresponds to the slope of the fitting; the values presented in the panels are the average of the results for the different values of i.

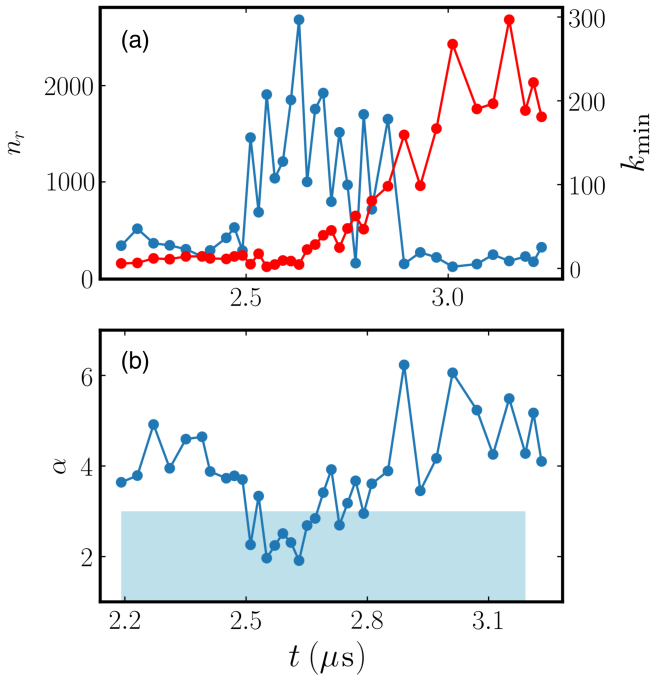

FIG. 9. Results of the power-law fitting for different times during the state preparation protocol. Panel (a) shows the estimated $k_{\min}$ (red points) and the number of nodes n_r with $k \geq k_{\min}$ (blue points). Panel (b) shows the power-law exponent α; the blue region in the graph corresponds to the region in which $1 \leq \alpha \leq 3$. The results, for each time, are an average of a set of 50 WFNs with $N_r = 9000$ nodes.

b. Time evolution of Rydberg atom systems

We now consider FSS analyses of WFNs generated at different times during the quasiadiabatic state preparation protocol; we focus on datasets obtained with simulations with NQSs. As described in the main text, during the dynamical process, the system is characterized by a regime of no correlations, a critical regime (for $t \approx 2.6$ μs) where it is close to a continuous quantum phase transition, and a regime where correlations are characterized by a length exceeding the system size (for $t > 3.0$ μs).

Figure 9 shows results of the power-law fitting for different times: the estimated $k_{\min}$, α, and the number of nodes in the scaling regime n_r. Within the critical regime, $2.5 < t < 2.7$ μs, the power-law exponent is $1 < \alpha < 3$. Now, we scrutinize the network structure over the different time windows.

(i) In the regime of no correlations, we observe that $\alpha > 3.0$ and $k_{\min}$ is $k_{\min} \approx O(k_{\max})$. Furthermore, the number of nodes for $k \geq k_{\min}$, which is $n_r < \ln(N_r)$ for some generated subnetworks, indicates that such WFNs do not correspond to a scale-free network [30], in full agreement with our theoretical arguments. The FSS analysis for $t = 1.52$ μs is depicted in Fig. 10, and confirms that the WFN is not scale-free.

(ii) Approaching the transition time leads to diverging correlations. As depicted in Fig. 10(b) for $t = 2.56$ μs, the WFN in this regime is consistent with a scale-free network; we obtain $S = 1.75$ for

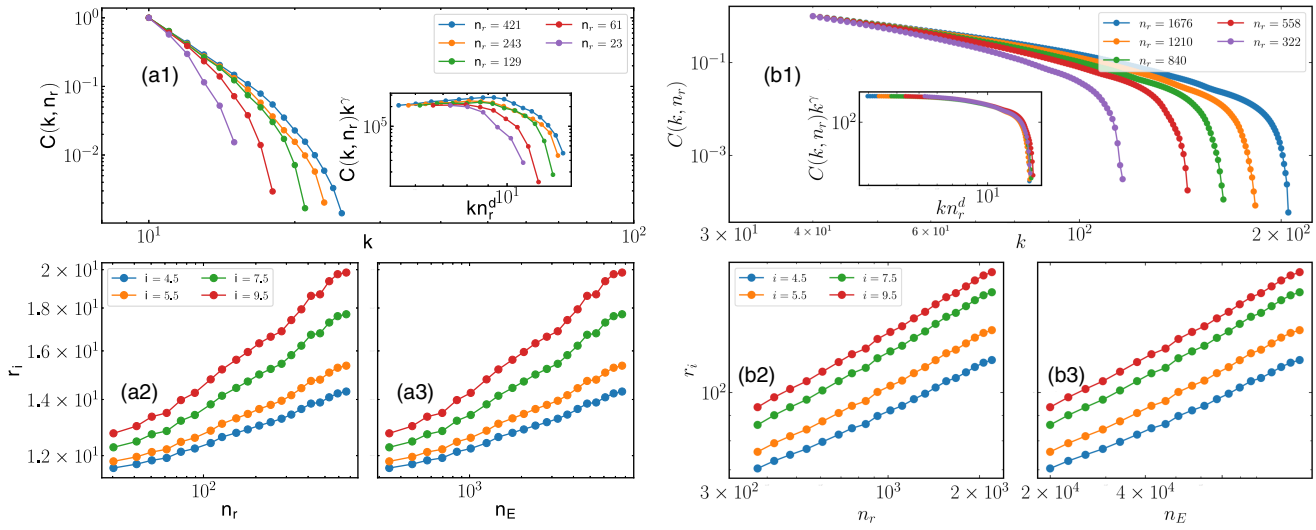

FIG. 10. Panels (a) and (b) show the FSS analyses for the WFNs generated at times $t = 1.52$ μs and $t = 2.56$ μs of the state preparation protocol, respectively. Panel (a1) [(b1)] shows the cumulative distribution, while the inset shows the data collapse. We observe large deviations from the FSS hypothesis in (a1). In panel (b1) the best collapse occurs for $d = -0.35 \pm 0.02$ and $\gamma = 1.80 \pm 0.06$, and $S = 1.75$. Panels (a2) and (a3) [(b2) and (b3)] show the scaling of the moment ratio r_i with the number of nodes n_r and the number of n_E, respectively, for different values of i. The exponent d (and d_E) corresponds to the slope of the fitting; the values presented in the panels are the average of the results for the different values of i. In both cases, $R = \langle r_1 \rangle$.

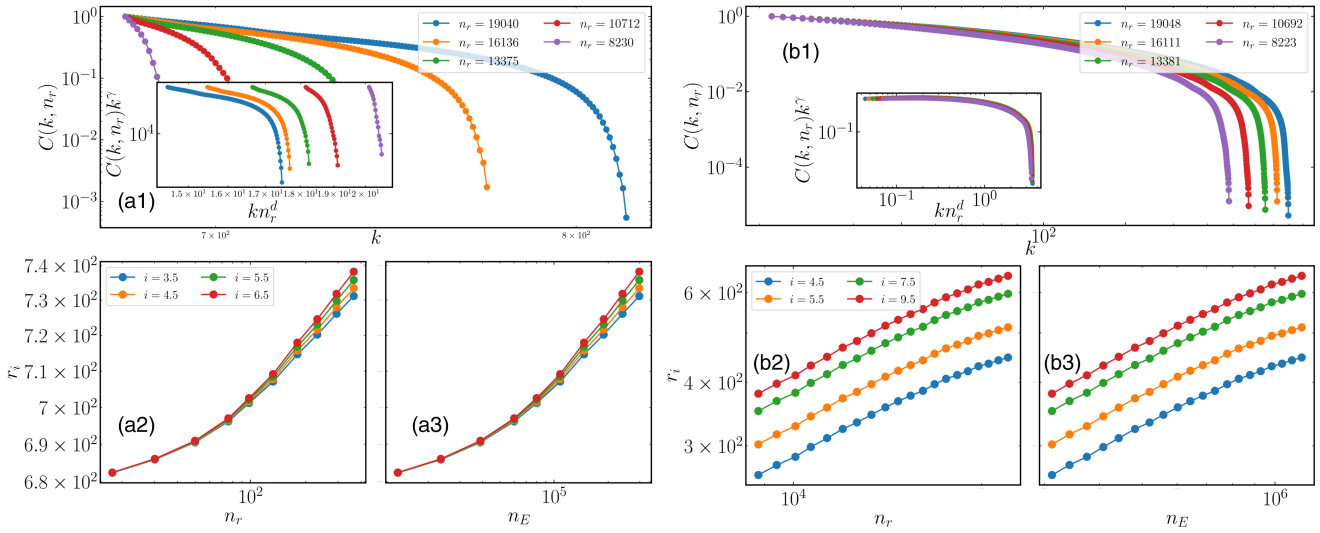

FIG. 11. Same results as Fig. 10 for $t = 3.18$ μs. In the panels on the left (a1, a2, a3) we select $k_{\min}$ with the power-law fitting, while in the panels on the right (b1, b2, b3) we impose $k_{\min} = 10$; see text. In both cases, $\langle R \rangle = R_1$.

$d = -0.35 \pm 0.02$ and $\gamma = 1.80 \pm 0.06$ with the data collapse analysis.

(iii) At later times (for $t > 3$ μs), the analysis is more challenging to interpret due to the fact that the estimated $k_{\min}$ is $k_{\min} \approx O(k_{\max})$. The FSS analyses (performed with the selected $k_{\min}$) show large deviations of a power-law distribution; see Fig. 11(a). We note, however, that in this case, $P(k, n_r)$ [or $C(k, n_r)$] is characterized by a heavy-tailed distribution. For instance, Fig. 11(b) shows the heavy-tailed $C(k, n_r)$ obtained by imposing an arbitrary $k_{\min}$ (i.e., we consider $k_{\min} = 10$), which might strongly suggest a scale-free behavior. However, here, while the collapse analysis is in agreement with the absence of any scale [see inset of Fig. 11(a1)], the log-log moment ratio analysis shows deviations of a straight line. We believe that a possible explanation here is that the network is either scale-free or features a very large scale, but that a precise determination of this fact is not feasible given the very large, self-consistently determined value of $k_{\min}$.

In conclusion, the finite-size scaling analysis we perform strongly confirms that quantum critical Ising networks are scale-free, and that the same phenomenology is observed while crossing the phase transition during quasiadiabatic ramps. The long-time system dynamics might also be scale-free, but at present, signatures are not unambiguous.

APPENDIX B: WHY NETWORK BUILT FROM DISTANCE (OR SIMILARITY) MEASURES CAN BE SCALE-FREE

We now present an argument that provides reasons why features of networks constructed from distance (or similarity [54]) measures between wave function snapshots can be characterized by heavy-tailed (and in some cases, power-law) distributions. The key element of our argument hinges on the demonstration that certain properties of distances between physical configurations and of WFNs (e.g., degree or strength distribution) are linked to physical observables and, more importantly, to probability distribution functions (PDFs) known to exhibit heavy-tailed behavior.

1. Distance (or similarity) measures and physical observables

The definition of the Hamming distance [Eq. (3)] can be recast as

$$d(X_i, X_j) \equiv d_{i,j} = \sum_{p=1}^{N} \frac{1}{2}(1 - s_p^i s_p^j), \tag{B1}$$

assuming that each site of the lattice can be in the states $s_p^i = +1$ (spin up) or $s_p^i = -1$ (spin down).

We start with the following observation: Quantities related to sums over all the nodes of the network can be linked to estimators of physical observables. For example, the estimator of the local magnetization is

$$\langle s_p \rangle = \frac{1}{N_r} \sum_{j=1}^{N_r} s_p^j. \tag{B2}$$

The estimator become equal to the exact value of the local magnetization $M_p = \sum_{j=1}^{2^N} P(X_j) s_p^j$ in the limit of $N_r \to \infty$. For the cases considered here, $P(X_j) = |\psi(X_j)|^2$, $\psi(X_j)$ being a normalized wave function. For a finite N_r, the estimator exhibits an error that scales as $1/\sqrt{N_r}$.

Equation (B2) and the definition of the Hamming distance, Eq. (B1), allow us then to show how the total distance $d^T = \sum_{i,j}^{N_r} d_{i,j}$ and the quantity $d_i = \sum_{j=1}^{N_r} d_{i,j}$

(dubbed node distance) are related to physical observables. The total distance is

$$d^T = \sum_{i=1}^{N_r}\sum_{j=1}^{N_r} d_{i,j} = \sum_{p=1}^{N}\frac{1}{2}\left(N_r^2 - \sum_{i=1}^{N_r} s_p^i \sum_{j=1}^{N_r} s_p^j\right) = \frac{N_r^2 N}{2}\left[1 - \frac{1}{N}\sum_{p=1}^{N}\langle s_p\rangle^2\right]. \tag{B3}$$

It is also worth mentioning that similar arguments to those described above are used in the context of spin glass physics [55].

More importantly for our argument, the node distance is equal to

$$d_i = \sum_{j=1}^{N_r} d_{i,j} = \sum_{j=1}^{N_r}\sum_{p=1}^{N}\frac{1}{2}(1 - s_p^i s_p^j) = \frac{N_r N}{2}\left(1 - \frac{1}{N}\sum_{p=1}^{N} s_p^i\langle s_p\rangle\right). \tag{B4}$$

In the limit of $N_r \to \infty$ and considering a system with translation symmetry, we can simplify

$$d_i = \frac{N_r N}{2}(1 - m^i\langle s\rangle), \tag{B5}$$

where $m^i = 1/N\sum_p s_p^i$ is the magnetization of the node i. It is straightforward to extend the aforementioned results for other distances or similarity measures. As an example, we consider the cosine similarity, $c_{i,j} = \sum_{p=1}^{N} s_p^i s_p^j$. In this case, the node similarity is

$$c_i = \sum_{j=1}^{N_r} c_{i,j} = N_r\sum_{p=1}^{N} s_p^i\langle s_p\rangle = N_r N m^i\langle s\rangle. \tag{B6}$$

The above arguments highlight a crucial point for understanding the nontrivial behavior of distributions characterizing distance and similarity measures—namely, the PDF of properties related to nodes [e.g., $P(d)$ and $P(c)$] is directly related to the PDF of the sum of a strongly correlated random variable: the magnetization, $p(m)$.

2. WFNs and physical observables

We now turn back to our network construction to illustrate how properties of WFNs are related to strongly correlated physical variables. For the sake of argument, we consider a weighted network, i.e., a network whose edges between two nodes i and j has a weight given by

$$w_{i,j} = c_{i,j}A_{i,j}, \tag{B7}$$

where $A_{i,j} = \Theta(c_{i,j} - C)$ is the adjacency matrix of the network; $\Theta(x) = 1$ when $x > 0$, and $\Theta(x) = 0$, otherwise. Given the linear relation between the Hamming distance and the cosine similarity, $c_{i,j} = N - 2d_{i,j}$, it is straightforward to show that the connectivity between the nodes of such a weighted network is equal to the ones of the network described in Sec. II. Specifically, their adjacency matrices are equal, $A_{i,j} = \Theta(c_{i,j} - C) = \Theta(R - d_{i,j})$, while the introduced cutoff C is $C = N - 2R$.

In addition to the node degree $k_i = \sum_j A_{i,j}$ defined in the main text, we consider the strength g_i of a node i, given by

$$g_i = \sum_{j=1}^{N_r} w_{i,j} = \sum_{j=1}^{N_r} c_{i,j}A_{i,j}. \tag{B8}$$

For a scale-free weighted network, the PDF $P(g)$ is also described by power-law distributions (and it directly simplifies to connectivity in the case of equal weights). Considering the cutoff C (or R) to build the WFN, the node strength can be written as

$$g_i = \sum_{(J)_i} c_{i,J} = \sum_{p=1}^{N} s_p^i \sum_{(J)_i} s_p^J, \tag{B9}$$

where $\sum_{(J)_i}$ represents a sum over the nodes that are connected to the node i, instead of being a sum over all nodes of the network. We now consider heuristic arguments to establish some generic features of $P(g)$.

The cutoff serves as a filter for establishing edges between nodes, prioritizing connectivity between nodes characterized by large m^i values. In WFNs, the hubs are nodes that have the highest m^i values. Consequently, the primary impact of the cutoff is to emphasize the role of large m in the distribution. It is important to note, however, that m^i couples with other physical properties that are dependent on the cutoff C (or R) and the node i, as shown in Eq. (B9). Although we do not specifically identify these physical variables, this line of reasoning enables us to highlight the crucial aspect for understanding the nontrivial behaviors of $P(g)$. Specifically, the strength g_i is related to sums of strongly correlated random variables. As discussed in Ref. [27], the latter are related to low-order correlation functions: Thus, it appears extremely unlikely that the corresponding network structure would feature a scale, when such correlations are, by definition, scale-free.

A limiting case corresponds to a fully connected WFN, i.e., the adjacency matrix is $A_{i,j} = 1$. In this case, Eq. (B8) is equivalent to Eq. (B6) and $P(g)$ is related to the distribution of magnetization $P(g) = P(N_r N\langle s\rangle m)$, so, $P(g)$ can exhibit nontrivial behavior reminiscent of the PDF of magnetization.

Summary. Summing up our arguments, the primary reason why WFNs can exhibit scale-free behavior can be traced back to the way certain network features are

connected to strongly correlated physical variables. To elucidate this point, we demonstrated how the strength distribution $P(g)$ correlates with the distributions of the sum of the variable s_p. It is important to note that such distributions can deviate from Gaussian distributions, thereby diverging from the predictions of the standard central limit theorem.

Interestingly, this argument can be made even sharper for the universality class that we consider here. Results from the 3D Ising classical model [32] (as well as for the one-dimensional quantum Ising model [33]) indicate that the large-m behavior of the tails of $p(m)$ in critical regimes is characterized by heavy-tailed distributions, which include power-law factors, and definitely go well beyond Gaussian *Ansätze*. These findings, in conjunction with the analytical and heuristic arguments presented in this section, reinforce our empirical observation that WFNs near quantum Ising critical points are scale-free.

APPENDIX C: WAVE NETWORK STRUCTURE AND THE CHOICE OF THE DISTANCE CUTOFF R

As described in Sec. II, the structure of the wave function network is defined by choosing a cutoff distance R in an embedded space defined by the Hamming distances, which allow us to define links between nodes. We now discuss in more detail the influence of R on the observation that WFNs can exhibit a scale-free structure.

Let us consider the list of all pairs of distances $d(X_i, X_j)$ between nodes X_i and X_j. One crucial aspect to consider is that the choice of R is bounded by the minimum and maximum distances on such a list (let us call these $d_{\min}$ and $d_{\max}$, respectively): $R < d_{\min}$ would generate a network with all the nodes isolated, while $R > d_{\max}$ would generate a featureless, fully connected network. The choice of R introduced on Eq. (4) naturally takes into account the typical scale distance in the embedded space, which depends on the N_r or the Hamiltonian parameters.

Another important aspect is that we deal with distances in a "high-dimensional" embedded space (the embedded dimension is equal to the number of spins N), where the so-called curse of dimensionality is expected to play a fundamental role. For instance, we could expect that the difference between the minimum and the maximum distance (i.e., $d_{\min} - d_{\max}$) would become indiscernible compared to any reasonable choice of R [56] given that the volume of a high-dimensional space increases so fast that the available data become sparse when $N_r \ll 2^N$. If this were the case, we would have observed just a featureless, fully connected network. In the correlated regime, however, we observe nontrivial network structures, which can be attributed to the fact that, in reality, the intrinsic dimension of the WFNs is much lower than the dimension of the embedded space [26,27].

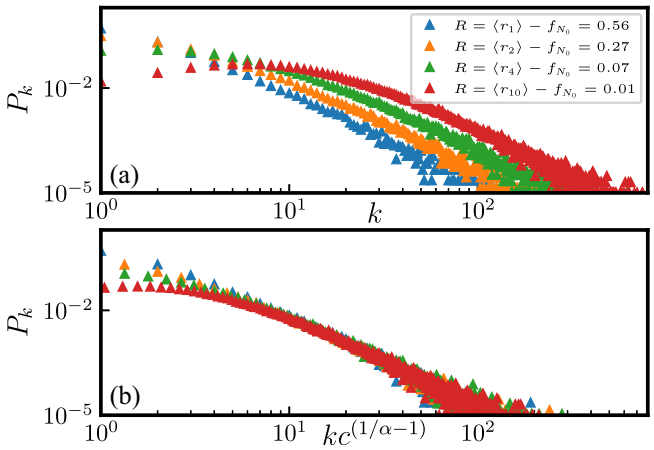

FIG. 12. Degree distribution, P_k, for the WFNs of the ground-state quantum Ising model at the critical point, $g = g_c$. Panel (a) shows P_k of the WFNs with different values of R, where $R = \langle r_c \rangle$ is the average of the c-nearest neighbor distances, as defined in Eq. (4). We also show the values of the fraction of isolated nodes, f_{N_0}, for the different values of R. In panel (b) we consider a phenomenological collapse for the different P_k, where c is the value of the indice related to $\langle r_c \rangle$ and $\alpha = 2.4$.

Let us discuss how changes on R influence the scale-free WFN. Figure 12 shows the degree distribution P_k associated with the WFN generated at the quantum critical point of the quantum Ising model. By increasing R, we observe two main effects. First, the P_k is shifted by larger values of k. Second, the threshold k_c above which P_k starts to behave as power law increases; see Fig. 12(a). Specifically, we observe that our data for different values of R collapse in a same curve when we rescale the x axis; see Fig. 12(b). This result indicates that for the scale-free WFN, the main effect of increasing R is to enlarge the cutoff k_c below which the network is not scale-free. In addition, we show the fraction of isolated nodes f_{N_0} by increasing R. For $R = \langle r_{10} \rangle$ less then 1% of the nodes are isolated; however, we still observe a power-law behavior for almost one decade in k. Overall, we notice that for $g \approx g_c$ we can always observe a scale-free WFN for a wide range of choices of R.

APPENDIX D: SIMULATIONS WITH NEURAL QUANTUM STATES

Neural quantum states have emerged recently as a new versatile class of variational wave functions [11]. The goal is to find an efficient representation of a many-body wave function $|\psi\rangle$ in the form of a parametrized function $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$ that maps a computational basis configuration $\boldsymbol{s} = (s_1, \ldots, s_N)$ to a complex number, such that

$$|\psi_{\boldsymbol{\theta}}\rangle = \sum_{\boldsymbol{s}} \psi_{\boldsymbol{\theta}}(\boldsymbol{s}) |\boldsymbol{s}\rangle. \tag{D1}$$

Here, $|\boldsymbol{s}\rangle = |s_1\rangle \otimes \cdots \otimes |s_N\rangle$ denotes the computational basis states of a system with N degrees of freedom, and for

our purposes $s_i \in \{\uparrow, \downarrow\}$. There are a number of appealing reasons to choose $\psi_{\boldsymbol{\theta}}(\boldsymbol{s})$ in the form of an artificial neural network (ANN) to render the *Ansatz* an NQS. Most importantly, rigorous representation theorems guarantee that any possible wave function can be approximated by an ANN in the limit of large network sizes [57–60]. This means that the approach is numerically exact in the sense that the accuracy of results can be certified self-consistently by convergence checks. While the general function approximation theorems do not tell us whether the representation in the form of an ANN is efficient, it has been shown that NQSs cover some volume law entangled states and correlated states of systems in two spatial dimensions, which are notoriously difficult to capture with established methods [61–66]. Finally, the complexity of the algorithms involved scales gently with system size and number of parameters, and large parts are amenable to large-scale parallelization to take advantage of distributed graphic processing unit clusters [67].

While the variational *Ansatz* with a limited number of parameters solves the problem of efficient representation, the efficient extraction of information from the wave function is achieved by Monte Carlo sampling. For example, the quantum expectation value of an operator $\hat{O}$ can be rewritten as

$$\langle\psi_{\boldsymbol{\theta}}|\hat{O}|\psi_{\boldsymbol{\theta}}\rangle = \sum_{\boldsymbol{s}} \frac{|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2}{\langle\psi_{\boldsymbol{\theta}}\rangle\psi_{\boldsymbol{\theta}}} O_{\mathrm{loc}}(\boldsymbol{s}), \tag{D2}$$

with the local estimator $O_{\mathrm{loc}}(\boldsymbol{s}) = \sum_{\boldsymbol{s}'} O_{\boldsymbol{s},\boldsymbol{s}'}[\psi_{\boldsymbol{\theta}}(\boldsymbol{s}')/\psi_{\boldsymbol{\theta}}(\boldsymbol{s})]$ that can be computed efficiently for local operators with only a polynomial number of nonvanishing matrix elements $O_{\boldsymbol{s},\boldsymbol{s}'} = \langle\boldsymbol{s}|\hat{O}|\boldsymbol{s}'\rangle$. This means that the expectation value can be estimated efficiently by Monte Carlo sampling the Born probability distribution $(|\psi_{\boldsymbol{\theta}}(\boldsymbol{s})|^2/\langle\psi_{\boldsymbol{\theta}}\rangle\psi_{\boldsymbol{\theta}})$ [68], and the same holds for all quantities of interest appearing in NQS algorithms. Note that the only way to access the wave function in quantum simulation experiments is projective measurements, which are likewise a sampling of the Born distribution; this is a very useful parallel when attempting a direct comparison of the obtained data, because obtaining samples from the wave function could turn out to be very costly with alternative numerical approaches [10].

An optimal approximate solution of the Schrödinger equation $i(d/dt)|\psi_{\boldsymbol{\theta}}\rangle = \hat{H}|\psi_{\boldsymbol{\theta}}\rangle$ within the manifold of wave functions $|\psi_{\boldsymbol{\theta}}\rangle$ is obtained via a time-dependent variational principle (TDVP) [11,12,69]. This leads to an ordinary differential equation prescribing the time evolution of the variational parameters,

$$\mathrm{Im}[S_{k,k'}]\dot{\theta}_{k'} = -\mathrm{Im}[iF_k], \tag{D3}$$

with the quantum metric tensor $S_{k,k'} = \langle\partial_{\theta_k}\psi_{\boldsymbol{\theta}}\rangle\times\partial_{\theta_{k'}}\psi_{\boldsymbol{\theta}} - \langle\partial_{\theta_k}\psi_{\boldsymbol{\theta}}\rangle\psi_{\boldsymbol{\theta}}\langle\psi_{\boldsymbol{\theta}}\rangle\partial_{\theta_{k'}}\psi_{\boldsymbol{\theta}}$ and the force vector $F_k = \langle\partial_{\theta_k}\psi_{\boldsymbol{\theta}}|\hat{H}|\psi_{\boldsymbol{\theta}}\rangle - \langle\partial_{\theta_k}\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle\langle\psi_{\boldsymbol{\theta}}|\hat{H}|\psi_{\boldsymbol{\theta}}\rangle$; note that the imaginary part appears on both sides of the equation as we are considering real parameters [67,69]. Hence, the time-evolved wave function starting from a given initial state can be obtained by integrating Eq. (D3). In previous works it was found that careful regularization is crucial to achieve state-of-the-art results in this way [12,37]. For the present work, we developed a new way of phrasing and solving the variational problem, which we call the *conditional* TDVP. The details of this approach will be described in a separate paper [70]. All results presented here were obtained in this way.

The network architecture used in our simulation is a variant of the recurrent neural network (RNN) for two-dimensional systems introduced in Ref. [71]. The structure of this architecture is depicted schematically in Fig. 13. The starting point is a one-hot encoding $\boldsymbol{\sigma}_{i,j}$ of the local spin configurations $s_{i,j}$, i.e., $\boldsymbol{\sigma}_{i,j} = (1,0)$ if $s_{i,j} = \uparrow$ or $\boldsymbol{\sigma}_{i,j} = (0,1)$ if $s_{i,j} = \downarrow$. The neural network is then evaluated by traversing the two-dimensional lattice in a snakelike manner. Let us denote the kth lattice site index along the snake path as (i_k, j_k) and assume that the linear dimension of the lattice is L. At each lattice site, a conditional single qubit state $\psi(s_{i_k,j_k}|s_{1,1},\ldots,s_{i_{k-1},j_{k-1}})$ is generated in the way detailed below. From these conditional states, the coefficient of the many-body wave function is obtained as

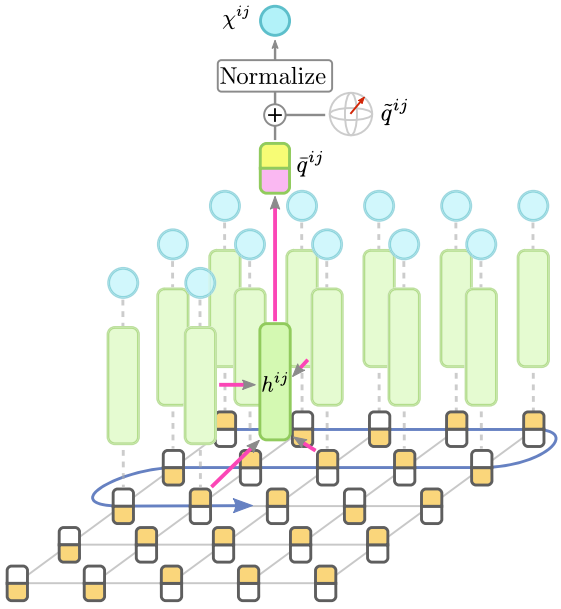

FIG. 13. Schematic depiction of the used neural network architecture. For evaluation the lattice is traversed along the path indicated by the blue arrow. A hidden state h^{ij} is computed at each site using the one-hot encoded local basis configurations and the hidden states of previously visited neighboring sites, as indicated by the pink arrows, which correspond to dense layers. From the hidden state a correlated contribution to the conditional qubit state $\bar{q}^{ij}$ is computed and an additional uncorrelated contribution $\tilde{q}^{ij}$ is added to it to obtain the logarithmic conditional amplitudes χ^{ij} after normalization, as noted in Eq. (D6).

$$\psi(\boldsymbol{s}) = \prod_{k=1}^{L^2} \psi(s_{i_k,j_k}|s_{1,1},\ldots,s_{i_{k-1},j_{k-1}}). \tag{D4}$$

For the conditional states at every lattice site, a local hidden state $\boldsymbol{h}^{(i,j)}$ is computed based on the spin configuration and hidden state of two neighboring sites as

$$h_l^{(i_k,j_k)} = f(W_{lm}^H h_m^{(i_{k-1},j_{k-1})} + W_{lm}^V h_m^{(i_{k-L},j_{k-L})}) + f(W_{lm}^{S_1}\sigma_m^{(i_{k-1},j_{k-1})} + W_{lm}^{S_2}\sigma_m^{(i_{k-L},j_{k-L})}). \tag{D5}$$

Here, f denotes the nonlinear activation function and $W_{lm}^{(\cdot)}$ denote the weights of the dense layers; double indices indicate summation. At the boundaries, where required neighboring sites do not exist, the corresponding input is replaced by zeros. Next, the hidden state is processed by a dense layer with two-dimensional output $(\bar{q}_{\mathrm{R}}^{ij}, \bar{q}_{\mathrm{I}}^{ij})$, corresponding to the real and imaginary parts of a complex number $\bar{q}_{ij}$. This number constitutes the correlated contribution to the logarithmic $\uparrow$ coefficient of the conditional local qubit state, up to normalization and a global phase. In addition, we introduced one complex-valued variational parameter $\tilde{q}^{ij} = \tilde{q}_{\mathrm{R}}^{ij} + i\tilde{q}_{\mathrm{I}}^{ij}$ for each lattice site, which corresponds to a contribution to the conditional qubit state that is uncorrelated. With $q^{ij} = \bar{q}^{ij} + \tilde{q}^{ij}$ we finally produce the logarithmic conditional wave function amplitudes,

$$\chi_{\uparrow}^{ij} = \frac{1}{2}\log\left(\frac{\exp(q_{\mathrm{R}}^{ij})}{1+\exp(q_{\mathrm{R}}^{ij})}\right) + iq_{\mathrm{I}}^{ij},$$
$$\chi_{\downarrow}^{ij} = \frac{1}{2}\log\left(\frac{1}{1+\exp(q_{\mathrm{R}}^{ij})}\right) + iq_{\mathrm{I}}^{ij}, \tag{D6}$$

such that

$$\psi(s_{i_k,j_k}|s_{1,1},\ldots,s_{i_{k-1},j_{k-1}}) = \exp\left(\chi_{s_{i_k,j_k}}^{i_k j_k}\right). \tag{D7}$$

The uncorrelated contribution $\tilde{q}^{ij}$ extends the standard RNN architecture. We introduced it because we found it difficult with the plain RNN to capture the initial part of the control protocol, where only the orientation of the uncorrelated qubit states is rotated and hardly any correlations are produced. In our architecture $\tilde{q}^{ij}$ can fully capture the product state, such that the job of the RNN is just to account for correlations on top of it. Including $\tilde{q}^{ij}$ does not affect the autoregressive property of the *Ansatz* introduced by the decomposition into a product of conditionals (D4). This means that the architecture allows for direct sampling of uncorrelated configurations at the cost of a single network evaluation per sample [71,72].

For the simulations we incorporate further experimental details, extending the elementary Rydberg atom Hamiltonian given in Eq. (9) of the main text. We include spatial laser intensity profiles that were extracted from the experimental setup such that the considered model Hamiltonian reads

$$H(t) = \hbar\sum_{k,l=0}^{L-1}\delta_k(t)n_{(kl)} + \frac{1}{2}\sum_{k,l=0}^{L-1}\Omega_k(t)\sigma_{(kl)}^x + \sum_{i<j}U_{ij}n_i n_j. \tag{D8}$$

Here, we introduced the notation $(kl) \equiv kL + l$ to map between double and single indices of the lattice sites; accordingly, the lasers shining in along one of the lattice dimensions exhibit an intensity profile perpendicular to that direction. The spatial and temporal form of the control fields during the considered protocol are shown in Figs. 14(a) and 14(b). The coupling is $U_{ij} = U/\Delta r_{ij}^6$ with nearest-neighbor interaction energy $U/h = 1.947$ MHz, where h is Planck's constant, and $\Delta r_{(kl)(mn)} = \sqrt{(k-m)^2+(l-n)^2}$ the Euclidian distance between lattice sites.

At the beginning of the protocol all atoms are prepared in their ground state, meaning that the initial state in the spin language is a polarized state $|\psi(t=0)\rangle = |\downarrow,\ldots,\downarrow\rangle$. The initial part of the protocol mostly consists of a

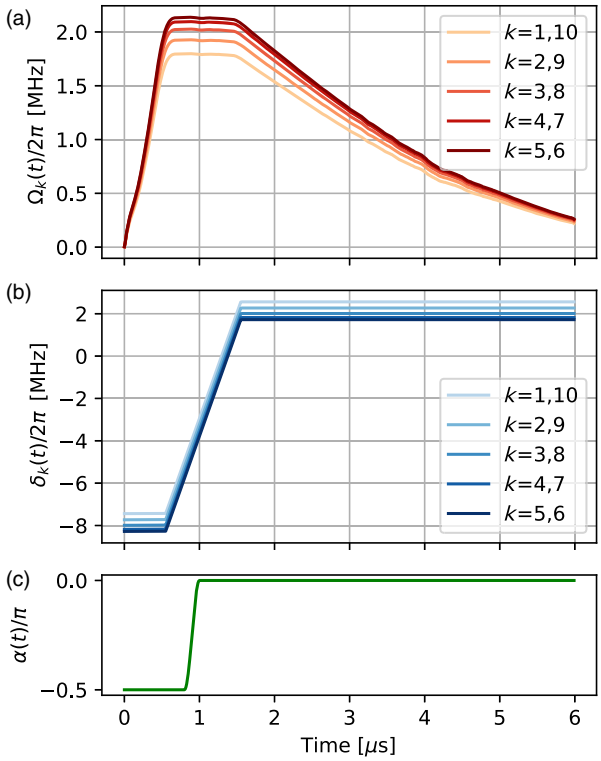

FIG. 14. (a),(b) Control protocols of the external fields $\Omega_k(t)$ and $\delta_k(t)$. (c) Time dependence of $\alpha(t)$, which parametrizes the time-dependent choice of the computational basis as described in the text. Initially, the quantization axis aligns with σ^x, before it is rotated on the time interval between $t_0 = 0.8$ μs and $t_1 = 1$ μs with $\alpha(t) = -(\pi/2)\cos^2[(\pi/2)(t-t_0)/(t_1-t_0)]$ to align with the σ^z quantization axis.

nearly adiabatic rotation of the polarization. This situation is difficult to address with NQS when using a fixed computational basis, because polarized states that align with the computational basis are hard to encode with NQS. Therefore, we implemented our simulation in a time-dependent frame $W(t) = \exp[-i\alpha(t)\sum_i \sigma_i^y]$ with $\alpha(t)$ as shown in Fig. 14(c) such that polarizations that align with the computational basis are avoided throughout the time evolution.

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

[2] A. Browaeys and T. Lahaye, *Many-body physics with individually controlled Rydberg atoms*, Nat. Phys. **16**, 132 (2020).

[3] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, *Circuit quantum electrodynamics*, Rev. Mod. Phys. **93**, 025005 (2021).

[4] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao, *Programmable quantum simulations of spin systems with trapped ions*, Rev. Mod. Phys. **93**, 025001 (2021).

[5] G. D. Mahan, *Many-Particle Physics* (Springer Science & Business Media, Berlin, 2013).

[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2010).

[7] R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks*, Rev. Mod. Phys. **74**, 47 (2002).

[8] S. N. Dorogovtsev, S. N. Dorogovtsev, and J. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, New York, 2003).

[9] M. Posfai and A.-L. Barabasi, *Network Science* (Cambridge University Press, Cambridge, England, 2016).

[10] P. Scholl, M. Schuler, H. J. Williams, A. A. Eberharter, D. Barredo, K.-N. Schymik, V. Lienhard, L.-P. Henry, T. C. Lang, T. Lahaye, A. M. Läuchli, and A. Browaeys, *Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms*, Nature (London) **595**, 233 (2021).

[11] G. Carleo and M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science **355**, 602 (2017).

[12] M. Schmitt and M. Heyl, *Quantum many-body dynamics in two dimensions with artificial neural networks*, Phys. Rev. Lett. **125**, 100503 (2020).

[13] B. Bollobás, *Graph Theory: An Introductory Course* (Springer Science & Business Media, Berlin, 2012), Vol. 63.

[14] H.-Y. Huang, R. Kueng, and J. Preskill, *Predicting many properties of a quantum system from very few measurements*, Nat. Phys. **16**, 1050 (2020).

[15] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, *Mixed-state entanglement from local randomized measurements*, Phys. Rev. Lett. **125**, 200501 (2020).

[16] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, *The randomized measurement toolbox*, Nat. Rev. Phys. **5**, 9 (2022).

[17] 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).

[18] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, *Quantum supremacy using a programmable superconducting processor*, Nature (London) **574**, 505 (2019).

[19] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, R. Finkelstein, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, F. G. S. L. Brandão, S. Choi, and M. Endres, *Preparing random states and benchmarking with many-body quantum chaos*, Nature (London) **613**, 468 (2023).

[20] A. N. Kolmogorov, *On tables of random numbers*, Sankhya A **4**, 369 (1963).

[21] M. Li *et al.*, *An Introduction to Kolmogorov Complexity and Its Applications* (Springer, New York, 2008), Vol. 3.

[22] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, *Quantum certification and benchmarking*, Nat. Rev. Phys. **2**, 382 (2020).

[23] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller, *Theoretical and experimental perspectives of quantum verification*, PRX Quantum **2**, 010102 (2021).

[24] T. Epps and K. J. Singleton, *An omnibus test for the two-sample problem using the empirical characteristic function*, J. Stat. Comput. Simul. **26**, 177 (1986).

[25] This can be done at the data structure level in several manners—either utilizing estimators based on discrete distances or adapting estimators with extra dimensions.

[26] 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).

[27] 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).

[28] A. W. Sandvik and J. Kurkijärvi, *Quantum Monte Carlo simulation method for spin systems*, Phys. Rev. B **43**, 5950 (1991).

[29] A. W. Sandvik, *Stochastic series expansion method for quantum Ising models with arbitrary interactions*, Phys. Rev. E **68**, 056701 (2003).

[30] M. Serafino, G. Cimini, A. Maritan, A. Rinaldo, S. Suweis, J. R. Banavar, and G. Caldarelli, *True scale-free networks hidden by finite size effects*, Proc. Natl. Acad. Sci. U.S.A. **118**, e2013825118 (2021).

[31] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286**, 509 (1999).

[32] M. M. Tsypin and H. W. J. Blöte, *Probability distribution of the order parameter for the three-dimensional Ising-model universality class: A high-precision Monte Carlo study*, Phys. Rev. E **62**, 73 (2000).

[33] A. Lamacraft and P. Fendley, *Order parameter statistics in the critical quantum Ising chain*, Phys. Rev. Lett. **100**, 165706 (2008).

[34] A. Polkovnikov, *Universal adiabatic dynamics in the vicinity of a quantum critical point*, Phys. Rev. B **72**, 161201(R) (2005).

[35] W. H. Zurek, U. Dorner, and P. Zoller, *Dynamics of a quantum phase transition*, Phys. Rev. Lett. **95**, 105701 (2005).

[36] J. Dziarmaga, *Dynamics of a quantum phase transition: Exact solution of the quantum Ising model*, Phys. Rev. Lett. **95**, 245701 (2005).

[37] M. Schmitt, M. M. Rams, J. Dziarmaga, M. Heyl, and W. H. Zurek, *Quantum phase transition dynamics in the two-dimensional transverse-field Ising model*, Sci. Adv. **8**, eabl6850 (2022).

[38] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samajdar, H. Pichler, W. W. Ho, S. Choi, S. Sachdev, M. Greiner, V. Vuletić, and M. D. Lukin, *Quantum phases of matter on a 256-atom programmable quantum simulator*, Nature (London) **595**, 227 (2021).

[39] D. Barredo, S. de Léséleuc, V. Lienhard, T. Lahaye, and A. Browaeys, *An atom-by-atom assembler of defect-free arbitrary two-dimensional atomic arrays*, Science **354**, 1021 (2016).

[40] A. Chandran, A. Erez, S. S. Gubser, and S. L. Sondhi, *Kibble-Zurek problem: Universality and the scaling limit*, Phys. Rev. B **86**, 064304 (2012).

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

[42] 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).

[43] 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).

[44] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, *Cross-platform verification of intermediate scale quantum devices*, Phys. Rev. Lett. **124**, 010504 (2020).

[45] A. L. Barabasi, *Network Science* (Cambridge University Press, Cambridge, England, 2016).

[46] Dorian Krause, *Jülich Supercomputing Centre, JUWELS: Modular tier-0/1 supercomputer at the Jülich Supercomputing Centre*, J. Large-Scale Res. Facil. **5**, A135 (2019).

[47] A. D. Broido and A. Clauset, *Scale-free networks are rare*, Nat. Commun. **10**, 1017 (2019).

[48] M. Gerlach and E. G. Altmann, *Testing statistical laws in complex systems*, Phys. Rev. Lett. **122**, 168301 (2019).

[49] P. L. Krapivsky and S. Redner, *Finiteness and fluctuations in growing networks*, J. Phys. A **35**, 9517 (2002).

[50] A. Clauset, C. R. Shalizi, and M. E. J. Newman, *Power-law distributions in empirical data*, SIAM Rev. **51**, 661 (2009).

[51] J. Alstott, E. Bullmore, and D. Plenz, *powerlaw: A Python package for analysis of heavy-tailed distributions*, PLoS One **9**, e85777 (2014).

[52] J. Houdayer and A. K. Hartmann, *Low-temperature behavior of two-dimensional Gaussian Ising spin glasses*, Phys. Rev. B **70**, 014418 (2004).

[53] O. Melchert, *autoScale.py—A program for automatic finite-size scaling analyses: A user's guide*, arXiv:0910.5403.

[54] L. da Fontoura Costa, *Coincidence complex networks*, J. Phys. Complex. **3**, 015012 (2022).

[55] G. Parisi, *Order parameter for spin-glasses*, Phys. Rev. Lett. **50**, 1946 (1983).

[56] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft, When is, *"nearest neighbor meaningful?"*, in *Proceedings of the 7th International Conference on Database Theory (ICDT'99)* (Springer, Berlin, 1999), pp. 217–235.

[57] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Math. Control Signals Syst. **2**, 303 (1989).

[58] K. Hornik, *Approximation capabilities of multilayer feedforward networks*, Neural Netw. **4**, 251 (1991).

[59] T. Kim and T. Adalı, *Approximation by fully complex multilayer perceptrons*, Neural Comput. **15**, 1641 (2003).

[60] N. L. Roux and Y. Bengio, *Representational power of restricted Boltzmann machines and deep belief networks*, Neural Comput. **20**, 1631 (2008).

[61] D.-L. Deng, X. Li, and S. Das Sarma, *Quantum entanglement in neural network states*, Phys. Rev. X **7**, 021021 (2017).

[62] X. Gao and L.-M. Duan, *Efficient representation of quantum many-body states with deep neural networks*, Nat. Commun. **8**, 662 (2017).

[63] Y. Levine, O. Sharir, N. Cohen, and A. Shashua, *Quantum entanglement in deep learning architectures*, Phys. Rev. Lett. **122**, 065301 (2019).

[64] O. Sharir, A. Shashua, and G. Carleo, *Neural tensor contractions and the expressive power of deep neural quantum states*, Phys. Rev. B **106**, 205136 (2022).

[65] C. Roth, A. Szabó, and A. MacDonald, *High-accuracy variational Monte Carlo for frustrated magnets with deep neural networks*, Phys. Rev. B **108**, 054410 (2023).

[66] M. Reh, M. Schmitt, and M. Gärttner, *Optimizing design choices for neural quantum states*, Phys. Rev. B **107**, 195115 (2023).

[67] M. Schmitt and M. Reh, *jVMC: Versatile and performant variational Monte Carlo leveraging automated differentiation and GPU acceleration*, SciPost Phys. Codebases **2** (2022) 10.21468/SciPostPhysCodeb.2.

[68] M. Van Den Nest, *Simulating quantum computers with probabilistic methods*, Quantum Inf. Comput. **11**, 784 (2011).

[69] J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. Van Leuven, *On the equivalence of time-dependent variational principles*, Chem. Phys. Lett. **149**, 547 (1988).

[70] M. Schmitt *et al.* (to be published).

[71] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, *Recurrent neural network wave functions*, Phys. Rev. Res. **2**, 023358 (2020).

[72] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, *Deep autoregressive models for the efficient variational simulation of many-body quantum systems*, Phys. Rev. Lett. **124**, 020503 (2020).