Quantum many-body dynamics in two dimensions with artificial neural networks

Markus Schmitt, Markus Heyl

Angaben zur Veröffentlichung / Publication details:

Schmitt, Markus, and Markus Heyl. 2020. "Quantum many-body dynamics in two dimensions with artificial neural networks." *Physical Review Letters* 125 (10): 100503. https://doi.org/10.1103/physrevlett.125.100503.

Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks

Markus Schmitt[1] and Markus Heyl[2]
[1]Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA
[2]Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Straße 38, 01187 Dresden, Germany

(Received 7 January 2020; revised 19 June 2020; accepted 7 August 2020; published 2 September 2020)

The efficient numerical simulation of nonequilibrium real-time evolution in isolated quantum matter constitutes a key challenge for current computational methods. This holds in particular in the regime of two spatial dimensions, whose experimental exploration is currently pursued with strong efforts in quantum simulators. In this work we present a versatile and efficient machine learning inspired approach based on a recently introduced artificial neural network encoding of quantum many-body wave functions. We identify and resolve key challenges for the simulation of time evolution, which previously imposed significant limitations on the accurate description of large systems and long-time dynamics. As a concrete example, we study the dynamics of the paradigmatic two-dimensional transverse-field Ising model, as recently also realized experimentally in systems of Rydberg atoms. Calculating the nonequilibrium real-time evolution across a broad range of parameters, we, for instance, observe collapse and revival oscillations of ferromagnetic order and demonstrate that the reached timescales are comparable to or exceed the capabilities of state-of-the-art tensor network methods.

DOI: 10.1103/PhysRevLett.125.100503

Introduction.—In the past two decades the field of nonequilibrium quantum many-body systems has seen a rapid development driven, in particular, by the remarkable progress in experiments [1–14]. Today, quantum simulators provide access to dynamics in quantum matter with an unprecedented control, which has led to the observation of genuinely nonequilibrium phenomena such as many-body localization [5,15,16], discrete time crystals [11,17], dynamical quantum phase transitions [7,18,19], or quantum many-body scars [8]. A particular frontier pushed forward by experiments recently is toward the nonequilibrium dynamics in two-dimensional (2D) quantum many-body systems [13,14,20,21]. The theoretical description of such unitary time evolution, yet, faces severe limitations. For instance, rapid entanglement growth or the exponential cost of contraction imposes strong constraints on tensor network approaches. Nevertheless, considerable progress has been reported to capture transient dynamics [22–30]. Recently, it has been proposed that machine learning techniques might overcome these difficulties by encoding quantum many-body states in artificial neural networks (ANNs) [31]. Subsequent efforts, however, raised doubts that this approach can enable the investigation of otherwise inaccessible regimes of nonequilibrium quantum dynamics [32].

In this work we overcome hitherto opaque and ultimately prohibitive numerical instabilities of the real-time ANN approach and we thereby expand state-of-the-art capabilities for the simulation of quantum many-body dynamics. Most importantly, we introduce a novel scheme to obtain a stable solution although only noisy estimates of the variational manifold and its relation to the physical system are known. Moreover, we target specific properties of the ANN itself, for instance, by utilizing deep architectures, i.e., convolutional neural networks. They naturally embody the fundamental physical principles of locality and causality, which can enhance the encoding efficiency. We apply our approach to the paradigmatic transverse-field Ising model on a square lattice, whose nonequilibrium dynamics has recently been shown to be accessible in systems of Rydberg atoms [14,20,21]. With our resulting algorithm we obtain numerically exact results up to timescales comparable to or exceeding the capabilities of current tensor network algorithms, demonstrated by comparison to recent data from infinite projected entangled pair states (IPEPS) [26]. Computing the dynamics for a wide range of parameters, we observe, e.g., collapse and revival oscillations of the ferromagnetic order when strongly quenched by a transverse field; see Fig. 1. Importantly, we find that at this point the expressivity of the ANN is not the limiting factor and the achieved timescales could be extended at mild polynomial expense.

Neural network wave functions.—Considering a system of N spin-1/2 degrees of freedom, the quantum many-body wave function can be represented in the basis of spin configurations, $s = (s_1, s_2, \ldots, s_N)$, $s_j = \uparrow, \downarrow$, as

$$|\psi\rangle = \sum_s \psi(s)|s\rangle. \tag{1}$$

Because of the exponentially large Hilbert space, wave function based numerical methods aiming at large systems

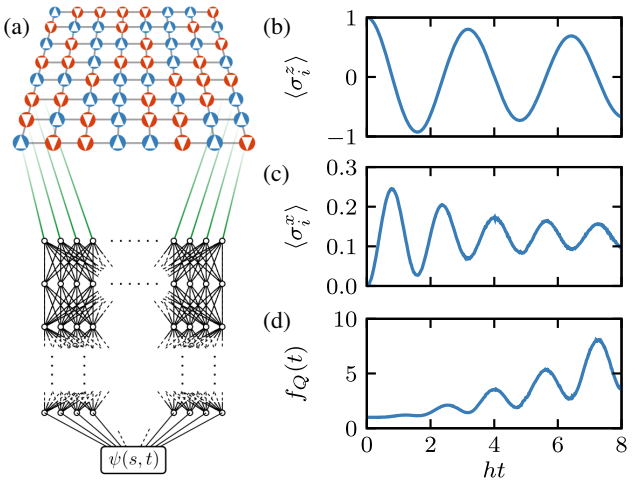

FIG. 1. (a) Schematic illustration of the artificial neural network (ANN) encoding of many-body wave functions in 2D quantum spin systems. A given spin configuration s, blue and red referring to the spin $\uparrow$ and $\downarrow$ state, respectively, functions as the input to an ANN whose output at the end is the corresponding wave function amplitude ψ_s. (b) Collapse and revival of the ferromagnetic order in a quantum Ising model of 8×8 spins on a square lattice after quenching the transverse field from $h = 0$ to $h = 2.63h_c$. (c) Dynamics of the transverse magnetization $\langle\sigma_i^x(t)\rangle$. The quantum Fisher information density $f_Q(t)$ in (d) reveals that genuine multipartite entanglement is generated by the unitary evolution.

need a strategy to avoid storing the individual amplitudes $\psi(s)$ in memory. In this work we construct a general-purpose variational wave function $\psi_\eta(s)$, parametrized by $\eta = (\eta_1, \ldots, \eta_M)$, which constitutes an efficient representation of $|\psi\rangle$ if M is much smaller than the Hilbert space size. Being able to provide a good approximation of the amplitudes on the fly [$\psi(s) \approx \psi_\eta(s)$], the variational wave function serves as a generative model, from which we can sample using conventional Monte Carlo techniques. Concretely, the expectation value of any observable $\hat{O}$ can be obtained as

$$\langle\psi_\eta|\hat{O}|\psi_\eta\rangle = \sum_s |\psi_\eta(s)|^2 O_\eta(s), \tag{2}$$

with $O_\eta(s) = \sum_{s'}\langle s|\hat{O}|s'\rangle\psi_\eta(s')/\psi_\eta(s)$. Since $\langle s|\hat{O}|s'\rangle$ is sparse for few-body observables, the expectation value can be computed efficiently by Monte Carlo sampling the probability $p_\eta(s) = |\psi_\eta(s)|^2$; importantly, there is no sign problem associated with this procedure.

Clearly, it might appear difficult to construct a general-purpose generative machine. However, simple but powerful versions have already been constructed recently for tailored problems [33,34]. Aiming for a more versatile approach we now follow the proposal to employ artificial neural networks [31]. ANNs have the crucial advantage that they are universal function approximators [35–37]. As a consequence, any quantum many-body wave function can, in principle, be represented by ANNs provided the network is sufficiently large. Consequently, the network size acts as a control parameter for our simulations that can be used to check convergence of the results. Moreover, the celebrated gradient backpropagation algorithm [38–40] enables the efficient numerical treatment of this class of variational wave functions.

As one of the key improvements we propose two modifications of the ANN structure compared to previous works. First, we explore deep architectures by means of convolutional neural networks, which naturally respect the fundamental principles of locality and causality. While we provide a detailed description of the convolutional neural networks (CNN) wave function in the Supplemental Material [41], let us point out that CNNs include the restricted Boltzmann machines (RBMs), which have been used in previous works for quantum dynamics [31,32,42,43], as the special case of a fully connected single-layer CNN with a fixed activation function. By contrast, CNNs are typically constructed as deep networks with sparse connectivity and arbitrary activation functions. For ground-state searches, CNN architectures have already been explored previously [44] with a polynomially enhanced efficiency in encoding entanglement as compared to the RBM [45]. The hidden unit density α, which specifies the size of a RBM [31], corresponds to the number of channels in terms of a CNN architecture, where the filter diameter d_F, that defines the connectivity, equals the linear extent of the system. Accordingly, we will denote the size of a CNN with L layers by a tuple $\alpha = (\alpha_1, \ldots, \alpha_L; d_F)$, with α_k the number of channels in the kth layer [41].

Second, we find that it is crucial for the description of the unitary dynamics to use analytic activation functions for the complex-valued ANNs. In contrast to ground-state searches, which are resilient to the encountering of poles and branch cuts of typical activation functions due to the projective nature of imaginary time evolution, real-time evolution relies on the differentiability of the wave function at any point of the variational manifold in the full complex plane. In our simulations we use as activation functions a sixth order polynomial in the first layer, which allows us to directly incorporate the $\mathbb{Z}_2$ symmetry, and odd fifth order polynomials in the following layers to avoid the vanishing gradient problem [41,46].

Training and the noisy TDVP.—Training, i.e., optimizing $\psi_\eta(s)$ to represent the dynamical quantum many-body wave function, is performed by demanding for each time step τ that the change of parameters $\dot{\eta}$ minimizes the distance between the time-evolved state $e^{-i\tau H}|\psi_{\eta(t)}\rangle$ and $|\psi_{\eta(t)+\tau\dot{\eta}}\rangle$ as measured by the Fubini-Study metric $\mathcal{D}$ [31]. The corresponding optimization objective is

$$r^2(t) = \frac{\mathcal{D}(|\psi_{\eta(t)+\tau\dot{\eta}}\rangle, e^{-i\tau H}|\psi_{\eta(t)}\rangle)^2}{\mathcal{D}(|\psi_{\eta(t)}\rangle, e^{-i\tau H}|\psi_{\eta(t)}\rangle)^2}, \tag{3}$$

where the constant denominator is introduced as a natural scale for $r^2(t)$. Minimization with respect to $\dot{\eta}$ yields a first order differential equation for the variational parameters $\eta_k(t) \in \mathbb{C}$,

$$S_{k,k'}\dot{\eta}_{k'} = F_k, \tag{4}$$

where $S_{k,k'} = \langle\langle O_k^* O_{k'}\rangle\rangle_c$ and $F_k = -i\langle\langle O_k^* E_{\mathrm{loc}}\rangle\rangle_c$, with $k, k' = 1, \ldots, M$ and $\langle\langle AB\rangle\rangle_c = \langle\langle AB\rangle\rangle - \langle\langle A\rangle\rangle\langle\langle B\rangle\rangle$ a connected correlation function. Here, we introduced the variational derivatives $O_k(s) = \{[\partial \ln \psi_\eta(s)]/\partial\eta_k\}$ and the local energy $E_{\mathrm{loc}}(s) = \sum_{s'}\langle s|H|s'\rangle\{[\psi_\eta(s')]/[\psi_\eta(s)]\}$. The brackets $\langle\langle\cdot\rangle\rangle$ denote expectation values with respect to the normalized probability distribution obtained from $|\psi_\eta(s)|^2$. Note that Eq. (4) is the well-known time-dependent variational principle (TDVP) equation, which for holomorphic $\psi_\eta(s)$ equivalently follows from an action principle [31,47–49]. The Fubini-Study distance (3) additionally provides us with a practical figure of merit and in the following we will regard the integrated residual, $R^2(t) = \int_0^t dt' r^2(t')$, as a measure of the accuracy of our simulations. For completeness, we include a derivation of Eq. (4) and the explicit form of the residual (3) in Ref. [41].

While solving the exact TDVP equation will yield the optimal parameter update given the variational ansatz, it is important to realize that in practice we will have incomplete knowledge of the equation itself, because both $S_{k,k'}$ and F_k can only be estimated by Monte Carlo sampling. In previous works a pseudoinverse was used to regularize the inversion of the typically ill-conditioned S matrix to avoid contributions from small eigenvalues that can lead to a numerical instability [31,32,42]. In this work, instead, we follow a different approach by precisely identifying and disregarding the noisy components of Eq. (4). This new regularization scheme is crucial to be able to reach the network sizes and timescales presented in the following.

For our analysis we consider the TDVP equation (4) in the eigenbasis of S,

$$\sigma_k^2\dot{\tilde{\eta}}_k = \langle\langle Q_k^* E_{\mathrm{loc}}\rangle\rangle_c \equiv \rho_k, \tag{5}$$

where $S_{k,k'} = V_{k,l}\sigma_l^2(V^\dagger)_{l,k'}$, $Q_k = (V^\dagger)_{k,k'}O_k$, and $\dot{\tilde{\eta}}_k = (V^\dagger)_{k,k'}\dot{\eta}_{k'}$. Our key observation is the fact that the signal-to-noise ratio of σ_k, $\mathrm{SNR}(\sigma_k)$, is independent of k, while $\mathrm{SNR}(\rho_k)$ shows a clear k dependence [41]. This numerical observation is consistent with the behavior of signal-to-noise ratios derived analytically by assuming that the joint distribution of Q_k and E_{loc} is Gaussian. In this case, $\mathrm{SNR}(\sigma_k) = \sqrt{N_{\mathrm{MC}}/2}$ is completely determined by the number of Monte Carlo samples N_{MC}, whereas

$$\mathrm{SNR}(\rho_k) = \sqrt{\frac{N_{\mathrm{MC}}}{1 + \frac{\sigma_k^2}{\rho_k^2}\mathrm{Var}(H)}} \tag{6}$$

depends on k, and additionally the physical energy variance $\mathrm{Var}(H)$. We find that Eq. (6) agrees also quantitatively well with the empirically estimated SNR [41].

For our regularization scheme we compute $\mathrm{SNR}(\rho_k)$ and discard all components of ρ_k, which fall below a fixed threshold. Thereby, we ignore contributions to the TDVP equation of which we have insufficient knowledge due to finite N_{MC}. Remarkably, the SNR in the Gaussian approximation is directly related to the resulting TDVP residual (3): Disregarding the component k increases the residual by $\Delta_k = \{[|\rho_k|^2]/[\sigma_k^2\mathrm{Var}(\hat{H})]\} = \mathrm{SNR}(\rho_k)^2/N_{\mathrm{MC}}$. Hence, increasing N_{MC} at a fixed cutoff will systematically reduce the bias introduced by the regularization; the most important contributions will, however, be accounted for already with small N_{MC}, as this result indicates that they have the largest SNR. Further details of this approach and supporting numerical data are included in the Supplemental Material [41].

Further technical aspects.—To propagate the ANN wave function in time we use a second order consistent adaptive integrator. In our implementation we exploit the fact that $S_{k,k'}$ is the metric tensor of the variational manifold [50], which induces a meaningful measure for the quantification of the integration accuracy [41].

In some cases, when using small network sizes, we found that the details of the resulting dynamics can depend on the initialization of the network. Similar observations have recently been reported in a more general context of training neural networks and it has been proposed to address this issue by ensemble averaging over a number of independently initialized and trained networks [51]. We adapted this idea and found that the ensemble average shows good agreement with results from larger networks and the IPEPS reference data [41].

The Monte Carlo sampling from the CNN wave function amplitudes can become computationally very intense, calling for an efficient parallel implementation. Sampling is straightforwardly parallelizable over many processors of a distributed memory machine using a message passing scheme. The network evaluation allows for a shared memory parallelization using the individual cores of a processor or attached accelerators like graphics processing units (GPUs). Thereby, this machine learning approach allows us to make full use of the computational resources of cutting-edge supercomputers for the simulation of quantum many-body dynamics. Further details are contained in the Supplemental Material [41].

Transverse-field Ising model.—As a paradigmatic example of a quantum many-body system we consider the transverse-field Ising model on a 2D square lattice, defined by the Hamiltonian

$$H = -J\sum_{\langle i,j\rangle}\sigma_i^z\sigma_j^z - h\sum_j \sigma_j^x. \tag{7}$$

Here, the $\sigma_i^{x/z}$ denote the Pauli x and z matrices and $\langle i,j\rangle$ is the set of all neighboring sites in the lattice. The model exhibits a quantum phase transition at the critical transverse field $h_c/J = 3.04438(2)$ [52] separating a ferromagnetic phase at $h < h_c$ from a paramagnetic phase. This model has recently developed a particular practical relevance, as it is now naturally realized in Rydberg atom quantum simulators [14,20,21]. Different aspects of its dynamics in 2D have been addressed previously in Refs. [25,26,29,33,53,54].

In the following we demonstrate that the far from equilibrium dynamics induced by quantum quenches can be efficiently simulated using neural network wave functions, independent of the considered parameter regimes. We choose typical initial conditions of quantum simulators, namely uncorrelated product states $|\psi_0\rangle$. After preparation the dynamics generated by the Hamiltonian H yields the formal solution $|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle$.

Collapse and revival oscillations.—We start with a quench from a ferromagnetically polarized state $|\psi_0\rangle = |\uparrow\rangle = \prod_l |\uparrow\rangle_l$ into the paramagnetic phase at $h = 2.63h_c$. The resulting dynamics is shown in Figs. 1(b)–1(d). The order parameter $\langle\sigma_l^z\rangle$ exhibits collapse and revival dynamics with decaying amplitude, which is a consequence of relaxation due to interactions. This is accompanied by the oscillatory buildup of a transverse magnetization. Notably, significant entanglement is also generated; see Fig. 1(d), where we show the quantum Fisher information density $f_Q(t) = (1/N)\sum_{i,j}\langle\sigma_i^z\sigma_j^z\rangle_c$. After two oscillations of the order parameter, $f_Q(t) > 8$, implying that genuine multipartite entanglement has been developed of at least 9 spins [55,56]. We checked the accuracy upon increasing the network size and found that a single-layer fully connected CNN with $\alpha = 5$ is sufficient for convergence [41].

Quench from a paramagnetic initial condition.—Next, we consider quenches starting from a paramagnetic initial state $|\psi_0\rangle = |\rightarrow\rangle$. In this case we can compare our results to data obtained recently with an IPEPS algorithm [26].

In Fig. 2 we show results for quenches to weak and strong fields as well as to the quantum critical point, which has previously been identified to constitute a particularly challenging regime for the neural network approach [31,32]. For large fields $h_x = 2h_c$, we can observe relaxation of the transverse magnetization $\langle\sigma^x(t)\rangle$ to a steady state value with remaining temporal fluctuations due to the finite system size. In this regime quantum correlations only develop dominantly for nearest-neighboring spins. For the critical transverse field $h = h_c$ the magnetization decays to a much smaller value and significant quantum correlations spread in a light-cone fashion also to larger distances indicating a strongly correlated state. At weak transverse fields the dynamics appears more local than in the case of strong fields with quantum correlations emerging almost exclusively between nearest neighbors on the shown timescales.

Importantly, we find excellent agreement with the dynamics computed using IPEPS for all cases up to the maximally reached times in IPEPS, which are included in Fig. 2 as dashed lines for comparison. While IPEPS directly operates in the thermodynamic limit, the utilized machine learning approach enables us to reach significantly larger times for system sizes up to $N = 10 \times 10$. The direct comparison shows that the system size we reach is sufficient to exclude finite-size effects in local observables up to the timescales reached with IPEPS.

To independently assess the accuracy, we perform our simulations with varying network sizes and architectures. While fully connected single-layer CNNs are sufficient to

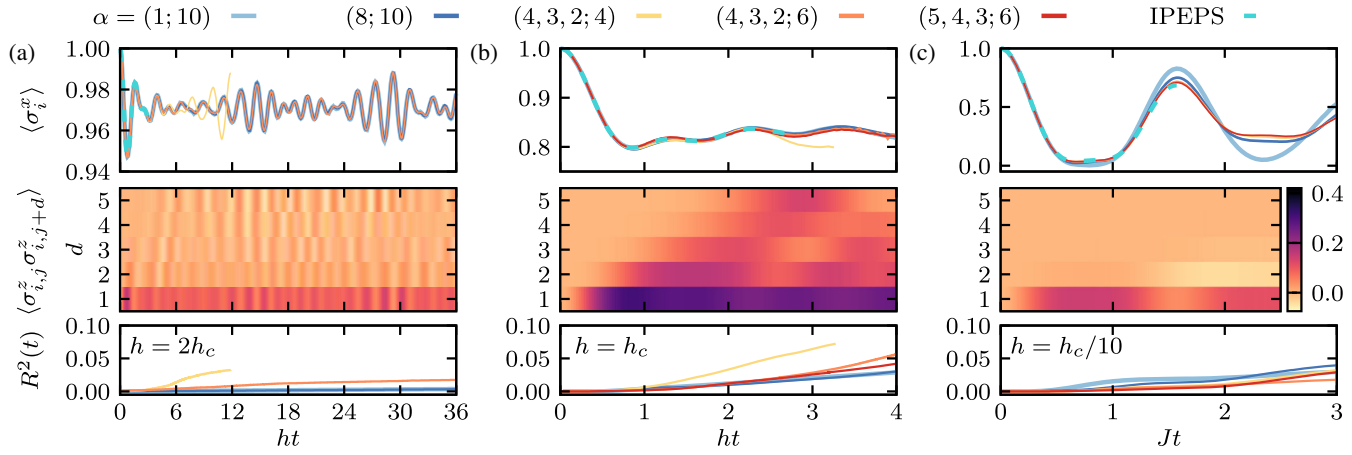

FIG. 2. Time evolution after quenching a transverse-field Ising model of size $N = 10 \times 10$ from the paramagnetically polarized initial state $|\psi_0\rangle = |\rightarrow\rangle$ (a) into the paramagnetic phase at $h = 2h_c$, (b) to the critical point, and (c) into the ferromagnetic phase at $h = h_c/10$. For direct comparison the top row includes data obtained with IPEPS from Ref. [26]. The agreement is very good in all cases for the networks with the smallest error $R^2(t)$ (bottom row). The second row shows space-time plots of correlation functions $\langle\sigma_{i,j}^z\sigma_{i,j+d}^z\rangle$ along the lattice axis from the simulations with minimal error.

reach convergence on timescales similar to or exceeding IPEPS for quenches into the paramagnetic phase or to the critical point, going to a deep CNN with sparse connectivity yields a substantial improvement over the single-layer network for $h = h_c/10$, indicated also by a significant reduction of the TDVP error $R^2(t)$. In that case, the dynamics remains more local, which can be exploited by using CNNs, as we discuss in the Supplemental Material [41]. We expect that this feature of deep CNNs can become relevant more generally when addressing larger system sizes, where correlations will remain constrained to smaller fractions of the system extent for longer times.

Discussion.—We have shown that variational time evolution of artificial neural network states constitutes a controlled and accurate approach to simulate dynamics in 2D quantum matter, which is competitive with current state-of-the-art tensor network algorithms. An alternative tensor network approach besides IPEPS is based on matrix product states and the approximation of 2D systems using cylindrical geometries [25,29]. For our purpose, however, we chose IPEPS as a reference, because it reflects the full C_4 symmetry of the square lattice and we avoid ambiguities caused by boundary effects in the reference data.

The availability of a versatile numerical method for time evolution paves the way to study the nonequilibrium quantum many-body dynamics in 2D and for new benchmarks of quantum simulators against classical simulations. The timescales and system sizes presented in this work can be extended at mild polynomial costs; importantly, we demonstrated that the network expressivity is currently not the limiting factor. These findings raise fundamental questions about our understanding of the complexity of quantum states. Moreover, the approach can, for example, be extended to systems with longer-ranged interactions and without translational invariance, which are challenging to address with tensor network methods.

We thank P. Czarnik, J. Dziarmaga, and P. Corboz for providing the IPEPS data from their work [26]. Moreover, we acknowledge fruitful discussions with J. Budich, M. Bukov, H. Burau, G. Carleo, M. Dupont, P. Karpov, and M. Behr. M. S. was supported through the Leopoldina Fellowship Programme of the German National Academy of Sciences Leopoldina (LPDS 2018-07) with additional support from the Simons Foundation. M. H. acknowledges support by the Deutsche Forschungsgemeinschaft via the Gottfried Wilhelm Leibniz Prize program. Furthermore, this project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No. 853443). Parts of the numerical simulations were performed at the Max Planck Computing and Data Facility in Garching. Moreover, the authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. [57] 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) [58].

Note added.—Recently, we became aware of related work by I. López-Gutiérrez and C. Mendl, which appeared simultaneously [59].

[1] M. Greiner, O. Mandel, T. W. Hansch, and I. Bloch, Collapse and revival of the matter wave field of a Bose–Einstein condensate, Nature (London) **419**, 51 (2002).
[2] T. Kinoshita, T. Wenger, and D. S. Weiss, A quantum Newton's cradle, Nature (London) **440**, 900 (2006).
[3] I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Rev. Mod. Phys. **86**, 153 (2014).
[4] E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt, Real-time dynamics of lattice gauge theories with a few-qubit quantum computer, Nature (London) **534**, 516 (2016).
[5] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, Exploring the many-body localization transition in two dimensions, Science **352**, 1547 (2016).
[6] C. Gross and I. Bloch, Quantum simulations with ultracold atoms in optical lattices, Science **357**, 995 (2017).
[7] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos, Direct Observation of Dynamical Quantum Phase Transitions in an Interacting Many-Body System, Phys. Rev. Lett. **119**, 080501 (2017).
[8] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Probing many-body dynamics on a 51-atom quantum simulator, Nature (London) **551**, 579 (2017).
[9] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature (London) **551**, 601 (2017).
[10] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet, Nat. Phys. **13**, 781 (2017).
[11] S. Choi, J. Choi, R. Landig, G. Kucsko, H. Zhou, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, V. Khemani, C. von Keyserlingk, N. Y. Yao, E. Demler, and M. D. Lukin, Observation of discrete time-crystalline order in a disordered dipolar many-body system, Nature (London) **543**, 221 (2017).
[12] H. Levine, A. Keesling, A. Omran, H. Bernien, S. Schwartz, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, High-Fidelity Control and Entanglement of Rydberg-Atom Qubits, Phys. Rev. Lett. **121**, 123603 (2018).
[13] S. Hild, T. Fukuhara, P. Schauß, J. Zeiher, M. Knap, E. Demler, I. Bloch, and C. Gross, Far-from-Equilibrium Spin

Transport in Heisenberg Quantum Magnets, Phys. Rev. Lett. **113**, 147205 (2014).
[14] D. Barredo, V. Lienhard, S. de Léséleuc, T. Lahaye, and A. Browaeys, Synthetic three-dimensional atomic structures assembled atom by atom, Nature (London) **561**, 79 (2018).
[15] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Observation of many-body localization of interacting fermions in a quasirandom optical lattice, Science **349**, 842 (2015).
[16] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Many-body localization in a quantum simulator with programmable random disorder, Nat. Phys. **12**, 907 (2016).
[17] J. Zhang, P. W. Hess, A. Kyprianidis, P. Becker, A. Lee, J. Smith, G. Pagano, I. D. Potirniche, A. C. Potter, A. Vishwanath, N. Y. Yao, and C. Monroe, Observation of a discrete time crystal, Nature (London) **543**, 217 (2017).
[18] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z. X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature (London) **551**, 601 (2017).
[19] N. Fläschner, D. Vogel, M. Tarnowski, B. S. Rem, D. S. Lühmann, M. Heyl, J. C. Budich, L. Mathey, K. Sengstock, and C. Weitenberg, Observation of dynamical vortices after quenches in a system with topology, Nat. Phys. **14**, 265 (2018).
[20] H. Labuhn, D. Barredo, S. Ravets, S. de Léséleuc, T. Macrì, T. Lahaye, and A. Browaeys, Tunable two-dimensional arrays of single Rydberg atoms for realizing quantum Ising models, Nature (London) **534**, 667 (2016).
[21] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, Probing the Quench Dynamics of Antiferromagnetic Correlations in a 2D Quantum Ising Spin System, Phys. Rev. X **8**, 021069 (2018).
[22] V. Murg, F. Verstraete, and J. I. Cirac, Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states, Phys. Rev. A **75**, 033605 (2007).
[23] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Time-evolving a matrix product state with long-ranged interactions, Phys. Rev. B **91**, 165112 (2015).
[24] J. Hauschild, F. Pollmann, and F. Heidrich-Meisner, Sudden expansion and domain-wall melting of strongly interacting bosons in two-dimensional optical lattices and on multileg ladders, Phys. Rev. A **92**, 053629 (2015).
[25] T. Hashizume, I. P. McCulloch, and J. C. Halimeh, Dynamical phase transitions in the two-dimensional transverse-field Ising model, arXiv:1811.09275.
[26] P. Czarnik, J. Dziarmaga, and P. Corboz, Time evolution of an infinite projected entangled pair state: An efficient algorithm, Phys. Rev. B **99**, 035115 (2019).
[27] C. Hubig and J. I. Cirac, Time-dependent study of disordered models with infinite projected entangled pair states, SciPost Phys. **6**, 31 (2019).
[28] C. Hubig, A. Bohrdt, M. Knap, F. Grusdt, and J. I. Cirac, Evaluation of time-dependent correlators after a local quench in iPEPS: hole motion in the t-J model, SciPost Phys. **8**, 021 (2020).
[29] T. Hashizume, J. C. Halimeh, and I. P. McCulloch, Hybrid infinite time-evolving block decimation algorithm for long-range multidimensional quantum many-body systems, Phys. Rev. B **102**, 035115 (2020).
[30] A. Kshetrimayum, M. Goihl, and J. Eisert, Time evolution of many-body localized systems in two spatial dimensions, arXiv:1910.11359.
[31] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).
[32] S. Czischek, M. Gärttner, and T. Gasenzer, Quenches near Ising quantum criticality as a challenge for artificial neural networks, Phys. Rev. B **98**, 024311 (2018).
[33] M. Schmitt and M. Heyl, Quantum dynamics in transverse-field Ising models from classical networks, SciPost Phys. **4**, 013 (2018).
[34] G. De Tomasi, F. Pollmann, and M. Heyl, Efficiently solving the dynamics of many-body localized systems at strong disorder, Phys. Rev. B **99**, 241114(R) (2019).
[35] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Netw. **4**, 251 (1991).
[36] T. Kim and T. Adal, Approximation by fully complex multilayer perceptrons, Neural Comput. **15**, 1641 (2003).
[37] N. Le Roux and Y. Bengio, Representational power of restricted Boltzmann machines and deep belief networks, Neural Comput. **20**, 1631 (2008).
[38] S. Dreyfus, The numerical solution of variational problems, J. Math. Anal. Appl. **5**, 30 (1962).
[39] D. E. Rumelhart and D. Zipser, Feature discovery by competitive learning, Cogn. Sci. **9**, 75 (1985).
[40] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to Machine Learning for physicists, Phys. Rep. **810**, 1 (2019).
[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.100503 for details about the convolutional neural network architecture, the signature of locality and causality in the neural network encoding, the time-dependent variational principle and the role of noise, the numerical procedure, additional convergence checks of the simulations, and a discussion of computational complexity and parallel computing performance.
[42] G. Fabiani and J. H. Mentink, Investigating ultrafast quantum magnetism with machine learning, SciPost Phys. **7**, 4 (2019).
[43] Y. Wu, L. M. Duan, and D.-L. Deng, Artificial neural network based computation for out-of-time-ordered correlators, Phys. Rev. B **101**, 214308 (2020).
[44] K. Choo, T. Neupert, and G. Carleo, Two-dimensional frustrated $J_1 - J_2$ model studied with neural network quantum states, Phys. Rev. B **100**, 125124 (2019).
[45] Y. Levine, O. Sharir, N. Cohen, and A. Shashua, Quantum Entanglement in Deep Learning Architectures, Phys. Rev. Lett. **122**, 065301 (2019).
[46] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in *Proceedings*

of the 13th International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 2010, edited by Y. W. Teh and M. Titterington, Vol. 9 pp. 249–256 (JMLR Workshop and Conference Proceedings, Sardinia, 2010).

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

[48] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-Dependent Variational Principle for Quantum Lattices, Phys. Rev. Lett. **107**, 070601 (2011).

[49] G. Carleo, Spectral and dynamical properties of strongly correlated systems, Ph.D. thesis, International School for Advanced Studies (SISSA), 2011.

[50] C.-Y. Park and M. J. Kastoryano, Geometry of learning neural quantum states, Phys. Rev. Research **2**, 023232 (2020).

[51] M. Geiger, S. Spigler, A. Jacot, and M. Wyart, Disentangling feature and lazy training in deep neural networks, arXiv:1906.08034.

[52] H. W. J. Blöte and Y. Deng, Cluster Monte Carlo simulation of the transverse Ising model, Phys. Rev. E **66**, 066110 (2002).

[53] M. Heyl, Scaling and Universality at Dynamical Quantum Phase Transitions, Phys. Rev. Lett. **115**, 140602 (2015).

[54] B. Blaß and H. Rieger, Test of quantum thermalization in the two-dimensional transverse-field Ising model, Sci. Rep. **6**, 38185 (2016).

[55] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, Fisher information and multiparticle entanglement, Phys. Rev. A **85**, 022321 (2012).

[56] G. Tóth, Multipartite entanglement and high-precision metrology, Phys. Rev. A **85**, 022322 (2012).

[57] See www.gauss-centre.eu.

[58] D. Krause, JUST: Large-scale multi-tier storage infrastructure at the Jülich supercomputing centre, J. Large-Scale Res. Facil. **5**, A135 (2019).

[59] I. López-Gutiérrez and C. B. Mendl, Real time evolution with neural-network quantum states, arXiv:1912.08831.