

Quenches near Ising quantum criticality as a challenge for artificial neural networks

Stefanie Czischek, Martin Gärttner, and Thomas Gasenzer

Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany

(Dated: March 23, 2018)

The near-critical unitary dynamics of quantum Ising spin chains in transversal and longitudinal magnetic fields is studied using an artificial neural network representation of the wave function. A focus is set on strong spatial correlations which build up in the system following a quench into the vicinity of the quantum critical point. We compare correlations observed following reinforcement learning of the network states with analytical solutions in integrable cases and tDMRG simulations, as well as with predictions from a semi-classical discrete Truncated Wigner analysis. While the semi-classical approach excels mainly at short times and for small transverse fields, the neural-network representation provides accurate results for a much wider range of parameters. Where long-range spin-spin correlations build up in the long-time dynamics we find qualitative agreement with exact results while quantitative deviations are of similar size as for the semi-classically predicted correlations, and slow convergence is observed when increasing the number of hidden neurons.

Introduction. Simulating the dynamics of interacting quantum many-body systems out of equilibrium is, in general, a hard problem for classical computers due to the exponential scaling of the Hilbert space dimension with the number of particles. Only in a limited number of cases, the quantum dynamics can be solved analytically. Such exact solutions exist mostly in one spatial dimension, such as for the transverse-field Ising model (TFIM) [1–3]. For one-dimensional (1D) systems, also matrix-product-state (MPS) representations of quantum states have proven most useful, including the time-dependent density-matrix renormalization-group (tDMRG) and related methods [4–9]. The tDMRG approach makes use of the fact that, for short-range interactions, an initially unentangled state develops entanglement only gradually and can thus be represented in an efficient way. However, for long times or spatial dimensions larger than one ($d > 1$), efficient and widely applicable numerical methods for calculating the dynamics are essentially absent.

A key to devising such methods is to know how to efficiently represent the quantum states under consideration. A novel idea is to make use of the capabilities of machine learning algorithms and artificial neural network (ANN) representations to efficiently exploit the structure of quantum states in a way similar to pattern recognition in image processing [10, 11]. Such approaches have been applied successfully in various areas of science such as computer vision, natural language processing, as well as in the natural sciences. Examples include quantum control [12, 13], phase classification and recognition in statistical physics [14, 15], gravitational wave analysis [16] and black hole detection [17] in astronomy, and error correction [18, 19] in quantum information science. Recently, it has been proposed to represent quantum states by means of neural networks, providing new ways to compute ground states as well as unitary time evolution [20–22], see also [23–30], directly related to certain types of tensor network states [31–34].

Here we provide a study of the regimes of validity of the ANN representation of a 1D quantum Ising chain in transversal and longitudinal fields and their relation to

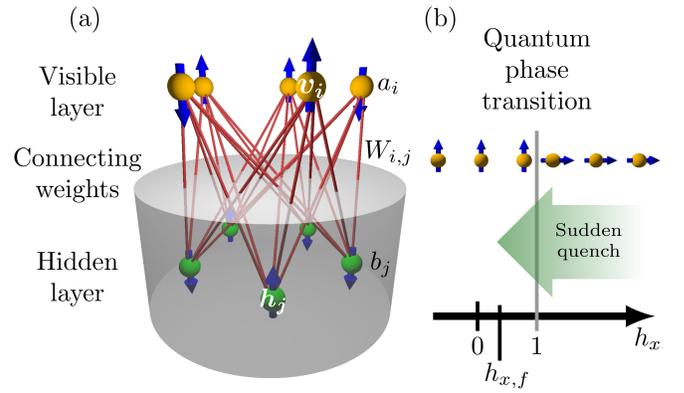


FIG. 1. (a) Setup of the ANN consisting of one visible and one hidden layer connected via weights $W_{i,j}$ between each pair of variables from different layers and bias a_i (visible) or b_j (hidden) for each variable. (b) Quench protocol in the TFIM for sudden quenches from the deep paramagnetic regime (large $h_{x,i}$) to different distances from the QCP within the paramagnetic and into the ferromagnetic phase, where the ground state configurations are depicted by the spins at the top.

the extent of spatial correlations and many-body entanglement, focusing on time evolution after a quench into the vicinity of a quantum critical point. We furthermore compare capabilities of the ANN method to the results obtained with a discrete truncated Wigner approximation (dTWA) approach recently proposed for quantum spin systems. In order to assess the power and weaknesses of these approaches, we make use of analytical solutions for the TFIM, and of numerical simulations by means of exact diagonalization and tDMRG in the case where a longitudinal field breaks the integrability.

The TFIM features a quantum phase transition where volume-law growth of entanglement is expected at long times, making it inaccessible to MPS based methods. We find that, while the semi-classical method fails to accurately capture the unitary dynamics except at very short times and near zero transverse field, the ANN approach reproduces the exact results for a wide range of param-

ters. We show that in cases where strong long-range spin-spin correlations build up, when the system is quenched into the vicinity of a quantum critical point, the two-layer ANN representation requires a strongly increased number of network parameters. Interestingly, the ANN approach yields qualitatively correct results even where the entanglement entropy increases indefinitely, only limited by the finite system size. Quantitative deviations, however, are of similar size as in semi-classical predictions and highlight the challenge for the employed ANN approach.

1D Ising model in longitudinal and transverse fields.

We consider the dynamics of a periodic chain of N spins governed by the quantum Ising model in a longitudinal field h_z and transverse field h_x , with Hamiltonian

$$\mathcal{H} = - \sum_{i=1}^N \sigma_i^z \sigma_{(i+1) \bmod N}^z - h_x \sum_{i=1}^N \sigma_i^x - h_z \sum_{i=1}^N \sigma_i^z, \quad (1)$$

defined in terms of the Pauli matrices σ_i^α . For $h_z = 0$ this model reduces to the transverse-field Ising model (TFIM). The TFIM is integrable as it can be mapped to free fermions and thus allows for comparisons with exact analytical solutions both, for ground states and unitary time evolution after a parameter quench [1–3, 35]. The spin system undergoes a quantum phase transition at $h_{x,c} = \pm 1$, from a ferromagnetic ($0 < |h_x| < 1$) to a paramagnetic ($|h_x| > 1$) phase, as depicted in Fig. 1(b). For $h_z \neq 0$ the model is no longer integrable and does not show a quantum phase transition as the Z_2 symmetry is broken explicitly [36].

Artificial-Neural-Network approach. The quantum state of N Ising spins can be expressed in terms of the 2^N complex coefficients $c_{\mathbf{v}}$, specifying the amplitudes with respect to the basis states $|\mathbf{v}\rangle = |v_1, \dots, v_N\rangle$ ($v_i \in \{-1, +1\}$), i.e. $|\Psi\rangle = \sum_{\mathbf{v}} c_{\mathbf{v}} |\mathbf{v}\rangle$. In Ref. [20], the representation

$$c_{\mathbf{v}} = \sum_{\{\mathbf{h}\}} e^{\sum_{i,j} v_i W_{i,j} h_j + \sum_i a_i v_i + \sum_j b_j h_j} \quad (2)$$

of these coefficients in terms of a set of complex parameters $W_{i,j}$, a_i , b_j , with $i = 1, \dots, N$, $j = 1, \dots, M$ has been proposed, which is reminiscent of the ANN structure of a restricted Boltzmann machine [37]. This involves an exponential bilinear form of the N visible, i.e., physical spins or neurons v_i and of the M hidden or auxiliary classical spin variables (neurons) $h_j \in \{-1, +1\}$ [not to be confused with the magnetic fields h_x , h_z appearing in Eq. (1)]. The exponential is summed over the hidden spin configurations. In the following we choose $M = \alpha N$ with integer α . In the two-layer neural network, only connections between the visible and the hidden but not within the layers are allowed, cf. Fig. 1(a).

To find representations of the form (2) for ground and unitarily evolving states of the model (1) a variational determination of the complex synaptic weights $W_{i,j}$ and biases a_i and b_j has been proposed in [20]. This can

be interpreted as reinforcement learning of the ANN and is accomplished by means of a stochastic reconfiguration procedure. This can be achieved in a numerically efficient way as the sum in the representation (2) can be performed analytically and the configurations $\{\mathbf{v}\}$ of the visible spins can be sampled using Markov-chain Monte Carlo methods. For further details of the ANN approach see App. B and Ref. [20].

Note that the number of network parameters $M + N + MN$ scales linearly in the system size N and reduces to $1 + \alpha + M$ if translationally invariant solutions are considered [20, 38]. Hence, the representation scales polynomially in the size of the system. In this respect the method is efficient and similar in spirit to variational Monte Carlo and MPS-based methods [5, 39, 40]. It has been discussed in the context of complexity theory [28] and exact ANN representations of specific classes of states have been found, including topological cluster and 2D toric-code states [24, 26, 28], as well as tensor network and chiral states [31–34, 41].

Discrete truncated Wigner approximation. The dTWA is a semi-classical simulation method for the dynamics of systems defined on a discrete phase space, such as the Ising model defined in Eq. (1) [42–45]. It allows calculating the dynamics by sampling initial states from a positive definite discrete Wigner function evolved by means of classical equations of motion for the spins. By averaging the resulting observables at a given evolution time over a large set of initial-state samples, one finds a semi-classical approximation to the exact unitary evolution (see App. A for further details).

Quenches in the TFIM. In our analysis we assume the system to be initially in the ground state deep in the paramagnetic phase, $h_{x,i} = 100$ and $h_{z,i} = 0$, and quench to a range of points $(h_{x,f}, h_{z,f})$ in the parameter space of (1), cf. Fig. 1(b). At a given evolution time after the quench we evaluate the spin-spin correlation function

$$C_d^{zz}(t) = \langle \sigma_0^z \sigma_d^z \rangle. \quad (3)$$

The absolute value of the correlation function typically shows an exponential decay for short distances d between the spins. We extract a correlation length ξ by fitting $|C_d^{zz}(t)|$ with an exponential function $\exp(-d/\xi)$ at small $d < 3$.

We begin by choosing a zero longitudinal field $h_{z,f}$ and compare the results of the ANN and dTWA methods to exact analytical results for $C_d^{zz}(t)$. Restricting ourselves first to small systems of $N = 10$ spins, Monte Carlo sampling of \mathbf{v} is not necessary in the ANN weight updating step which ensures that effects arising from a finite sample size do not constrain the performance. For such small systems, we can furthermore increase the number $M = \alpha N$ of hidden spins up to a size where the number of network parameters exceeds the dimension of Hilbert space and thus an exact ANN representation of the state should be possible. In this way we can explore the degree to which the full set of basis states is necessary in representing the system's state.

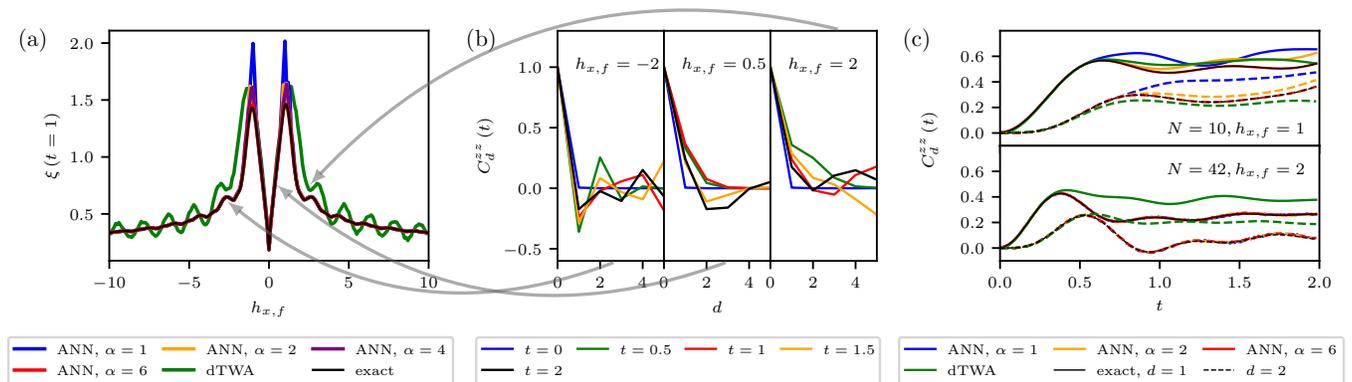


FIG. 2. (a) Correlation length $\xi(t)$ at a fixed time $t = 1$ as a function of the transverse field $h_{x,f}$ after a sudden quench from $(h_{x,i} = 100, h_{z,i} = 0)$ with $h_{z,f} = 0$ within the paramagnetic and into the ferromagnetic phase. The results of the ANN approach for different α are compared with the dTWA calculations and the exact solutions. (b) Time dependence of the correlation function $C_d^{zz}(t)$ after quenches from the deep paramagnetic phase to several $h_{x,f}$. The correlation length in (a) is calculated by fitting an exponential function to the short distance decay of the absolute value of C_d^{zz} . (c) Time evolution of the correlation function $C_d^{zz}(t)$ after a sudden quench to $(h_{x,f} = 1, h_{z,f} = 0)$ for a spin chain with $N = 10$ sites (upper plot) and to $(h_{x,f} = 2, h_{z,f} = 0)$ for a spin chain with $N = 42$ spins (lower plot).

Our results are presented in Fig. 2. Panel (a) shows the correlation length at a fixed time $t = 1$ as a function of $h_{x,f}$. Fig. 2(b) illustrates the d -dependence of $C_d^{zz}(t)$ for a selection of $h_{x,f}$ and times t , and panel (c) shows the correlation function at two different distances d and final fields $h_{x,f}$ as functions of time after the quench. Note that the results for $h_{x,f} = 2$ in Fig. 2(c) are for a larger system of $N = 42$ spins.

As can be seen in Fig. 2(a), for quenches into the vicinity of the quantum critical points (QCPs), $h_{x,f} \simeq \pm 1$, we observe the buildup of long-range correlations, resulting in strongly increased correlation lengths ξ [2, 3]. While the system possesses a quantum critical point at $h_x = \pm 1$, i.e., undergoes a quantum phase transition in the ground state, $\xi(t)$ is not expected to diverge there. The saturation of the dynamically evolving correlation length is closely related to the fact that the one-dimensional system does not dispose of phase transitions at non-zero temperatures [35].

In the vicinity of $h_{x,f} = \pm 1$ the ANN approach, when choosing a small number of hidden spins ($\alpha = 1$), yields correlations which deviate clearly from the exact result shown as a black solid line. Away from these critical values, the obtained results, however, match well with the exact correlations. Increasing the number of hidden spins, we find that the accuracy improves, whereas perfect agreement is only obtained for $\alpha > 6$. Note that then, the number of network parameters is $1 + \alpha + M > 67$ and thus of the order of the Hilbert space dimension after symmetrization, $d_H = 108$. Around $h_{x,f} = 0$, the ANN representation is well controlled as can be shown by means of a perturbative expansion in terms of classical spin networks [46].

The time evolution of $C_d^{zz}(t)$ shown in Figure 2(c), for a quench to $h_{x,f} = 1$ and $d = 1, 2$, corroborates the above findings concerning the dependence on α . At very short

times ($t \lesssim 0.5$), the ANN method gives accurate results already for $\alpha = 1$, even for quenches to criticality. For large spin chains ($N = 42$), $\alpha = 1$ is still sufficient to capture the exact dynamics in regimes of small correlation lengths ($h_{x,f} = 2$ is shown). For such large systems sizes, the weight updating procedure requires Monte Carlo sampling of the visible neurons. This turns out to have no effect other than adding statistical noise to the numerical result, which can be controlled by increasing the size of the Monte Carlo samples. By increasing the number of weights up to $\alpha = 6$ in the regime of large correlation length, i.e. close to the QCPs, we cannot obtain converged results, which in fact hints to an exponential scaling of the number of required network parameters with N . Hence, in these cases, the method appears to be of no advantage as compared to exact diagonalization (see App. E).

Turning to the dTWA method we find that it qualitatively reproduces the exact results while, in general, it shows rather large deviations except at short times ($t \lesssim 0.5$) and around $h_{x,f} = 0$. This is due to the fact that quantum effects are not captured by the semi-classical approximation. Note, however, that for quenches very close to the QCPs at $h_{x,f} = \pm 1$, the dTWA provides in general a better estimate of the short-range correlation length than efficient ANN representations (see App. C).

TFIM in a longitudinal field. Having considered only the integrable TFIM so far, we now turn to the non-integrable case by adding a longitudinal field $h_{z,f} > 0$. The initial fields are the same as above, $(h_{x,i} = 100, h_{z,i} = 0)$. Since no analytical solutions are available in this case, we resort to exact numerical diagonalization for small systems and to the tDMRG.

Fig. 3(a) shows the exact time evolution of the correlation length after a sudden quench to $h_{z,f} = 2$ and different values of $h_{x,f}$ ($N = 10$). The same dynamics is

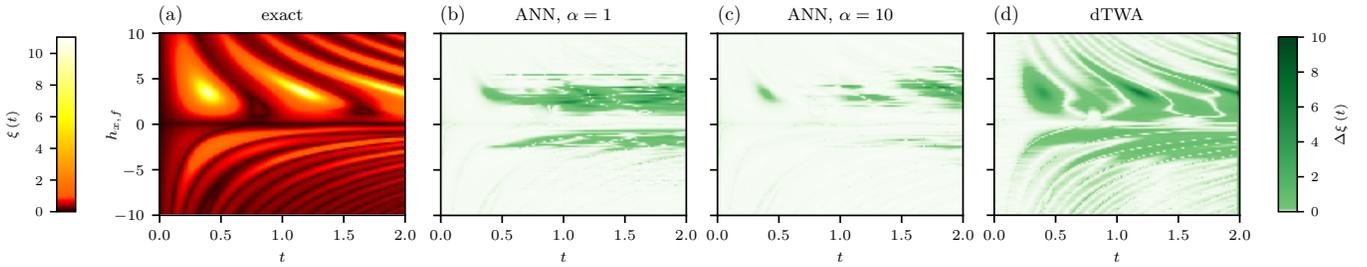


FIG. 3. (a) Exact correlation length $\xi_e(t)$ for $N = 10$ spins as a function of time after a sudden quench from $(h_{x,i} = 100, h_{z,i} = 0)$ to $h_{z,f} = 2$ and different values of $h_{x,f}$. (b),(c),(d) Deviations $\Delta\xi(t) = |\xi_s - \xi_e|$ of simulation results $\xi_s(t)$ from the exact result. The ANN calculation is not limited by Monte Carlo sampling errors. See also App. D for more final values $h_{z,f}$.

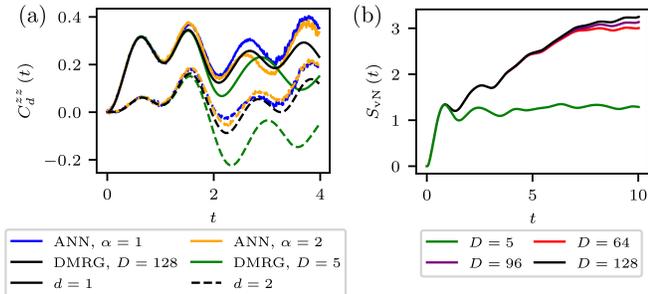


FIG. 4. (a) Time evolution of the correlation function $C_d^{zz}(t)$ after a sudden quench from $(h_{x,i} = 100, h_{z,i} = 0)$ to $(h_{x,f} = 0.5, h_{z,f} = 1)$ in a system with $N = 42$ sites. The ANN approach for $\alpha = 1$ and $\alpha = 2$ is compared to tDMRG calculations with bond dimensions $D = 128$ and $D = 5$. (b) von-Neumann entanglement entropy S_{vN} as a function of time after the same quench as in (a). The entanglement entropy is calculated using tDMRG with different bond dimensions. The deviations at late times show that in this regime tDMRG breaks down after long times due to the growing entanglement.

evaluated using the ANN approach with $\alpha = 1$, $\alpha = 10$ and the dTWA, where the deviation $\Delta\xi(t) = |\xi_s - \xi_e|$ of the correlation length $\xi_s(t)$ resulting from the different approaches as compared to the exact result $\xi_e(t)$ is shown in Fig. 3(b), (c), and (d), respectively. The ANN results are close to the exact calculations already for $\alpha = 1$ in regimes of small correlation length, namely for $|h_{x,f}| \gg 1$ and $|h_{x,f}| \ll 1$. Only for quenches to intermediate $h_{x,f}$, where the correlation length becomes large, deviations can be found for times $t > 0.5$, while the first oscillation in the correlation length is captured perfectly even here. For $\alpha = 10$ only small deviations can be found anywhere. In this case, the 111 weight parameters exceed the Hilbert space dimension $d_H = 108$, indicating that in certain parameter regimes the number of network parameters needed to achieve full convergence scales exponentially with the system size. The dTWA [Fig. 3(d)] shows similar deviations $\Delta\xi(t)$ in a wider regime, coinciding with the exact results mainly at short times.

An important question concerns the predictive power of the ANN approach in regimes where MPS-based ap-

proaches such as tDMRG are limited to short times due to an extensively growing entanglement entropy. As a representative case, we show, in Fig. 4(a), the correlation function after a quench to $(h_{x,f} = 0.5, h_{z,f} = 1)$ for a spin chain with $N = 42$ sites, where the ANN approach for $\alpha = 1$ and $\alpha = 2$ is compared to converged tDMRG results for bond dimension $D = 128$ and approximate data for $D = 5$. Clearly, deviations appear for times $t \geq 1$, while the qualitative behavior is captured for longer times. Increasing α from 1 (44 network parameters) to 2 (87 parameters) does not improve the convergence significantly. We note that, for $D = 128$, the number of variational parameters ($2D^2$) in the MPS vastly exceeds the number of parameters in the ANN approach ($1 + \alpha + M$). For bond dimension $D = 5$, the tDMRG variational ansatz has 50 parameters, resulting in a similar quality as the ANN. Fig. 4(b) shows the von Neumann entanglement entropy [$S_{vN} = -\text{Tr}(\rho_A \log \rho_A)$, with half-chain reduced density matrix ρ_A] obtained by means of tDMRG with different bond dimensions D , as a function of time after the same quench as before. Deviations appear at late times, showing that also tDMRG struggles due to the linearly growing entanglement entropy, which in turn requires an exponentially growing bond dimension.

Conclusions. Considering sudden quenches in the integrable transverse-field Ising model as well as in the non-integrable Ising model in a transverse and a longitudinal field, we compare the dynamics obtained within the discrete truncated Wigner approximation (dTWA) [42, 44] and a variational approach based on artificial neural networks (ANN) [20] with analytical and exact numerical results. Both the dTWA and the ANN methods reproduce the short-time dynamics correctly and perform well for small transverse fields. The ANN approach agrees with exact calculations in a much wider parameter regime than dTWA. Near criticality and at the presently attainable level of accuracy, it shows similar deviations as the semi-classical approach.

Acknowledgements. The authors thank A. Baumbach, H. Cakir, G. Carleo, J. C. Halimeh, M. Karl, M. Kastner, K. Meier, M. K. Oberthaler, J. M. Pawłowski, and A. Piñeiro Orioli for discussions and collaboration on the topics described here. The authors thank J. Schachen-

mayer for sharing his tDMRG codes. This work was supported by the Horizon-2020 framework programme of the European Union (FET-Proactive, AQuS, No. 640800),

by Deutsche Forschungsgemeinschaft (SFB 1225 ISO-QUANT), by Heidelberg University, and by the state of Baden-Württemberg through bwHPC.

-
- [1] P. Pfeuty, *Ann. Phys. (NY)* **57**, 79 (1970).
- [2] P. Calabrese, F. H. L. Essler, and M. Fagotti, *J. Stat. Mech. Theor. Exp.*, P07016 (2012).
- [3] P. Calabrese, F. H. L. Essler, and M. Fagotti, *J. Stat. Mech. Theor. Exp.*, P07022 (2012).
- [4] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [5] U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).
- [6] G. Vidal, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [7] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, *J. Stat. Mech. Theor. Exp.* **2004**, P04005 (2004).
- [8] S. Sharma, S. Suzuki, and A. Dutta, *Phys. Rev. B* **92**, 104306 (2015).
- [9] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, *Phys. Rev. B* **94**, 165116 (2016).
- [10] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, *Nature* **549**, 195 (2017).
- [11] V. Dunjko and H. J. Briegel, [arXiv:1709.02779 \[quant-ph\]](https://arxiv.org/abs/1709.02779) (2017).
- [12] E. Zahedinejad, J. Ghosh, and B. C. Sanders, *Phys. Rev. Applied* **6**, 054005 (2016).
- [13] M. August and X. Ni, *Phys. Rev. A* **95**, 012335 (2017).
- [14] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, *Nat. Phys.* **13**, 435 (2017).
- [15] J. Carrasquilla and R. G. Melko, *Nat. Phys.* **13**, 431 (2017).
- [16] R. Biswas, L. Blackburn, J. Cao, R. Essick, K. A. Hodge, E. Katsavounidis, K. Kim, Y.-M. Kim, E.-O. Le Bigot, C.-H. Lee, J. J. Oh, S. H. Oh, E. J. Son, Y. Tao, R. Vaulin, and X. Wang, *Phys. Rev. D* **88**, 062003 (2013).
- [17] M. Pasquato, [arXiv:1606.08548 \[astro-ph.GA\]](https://arxiv.org/abs/1606.08548) (2016).
- [18] G. Torlai and R. G. Melko, *Phys. Rev. Lett.* **119**, 030501 (2017).
- [19] P. Baireuther, T. E. O'Brien, B. Tarasinski, and C. W. J. Beenakker, *Quantum* **2**, 48 (2018).
- [20] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [21] H. Saito, *J. Phys. Soc. Jpn.* **86**, 093001 (2017).
- [22] Z. Cai and J. Liu, *Phys. Rev. B* **97**, 035116 (2018).
- [23] D.-L. Deng, X. Li, and S. Das Sarma, *Phys. Rev. X* **7**, 021021 (2017).
- [24] D.-L. Deng, X. Li, and S. Das Sarma, *Phys. Rev. B* **96**, 195145 (2017).
- [25] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, *Phys. Rev. B* **96**, 205152 (2017).
- [26] R. Kaubruegger, L. Pastori, and J. C. Budich, [arXiv:1710.04713 \[cond-mat.str-el\]](https://arxiv.org/abs/1710.04713) (2017).
- [27] P. Teng, [arXiv:1710.03213 \[quant-ph\]](https://arxiv.org/abs/1710.03213) (2017).
- [28] X. Gao and L.-M. Duan, *Nat. Commun.* **8**, (2017).
- [29] N. Freitas, G. Morigi, and V. Dunjko, [arXiv:1803.02118 \[quant-ph\]](https://arxiv.org/abs/1803.02118) (2018).
- [30] G. Carleo, Y. Nomura, and M. Imada, [arXiv:1802.09558 \[cond-mat.dis-nn\]](https://arxiv.org/abs/1802.09558) (2018).
- [31] Y. Huang and J. E. Moore, [arXiv:1701.06246 \[cond-mat.dis-nn\]](https://arxiv.org/abs/1701.06246) (2017).
- [32] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, *Phys. Rev. X* **8**, 011006 (2018).
- [33] S. R. Clark, *J. Phys. A: Math. Theor.* **51**, 135301 (2018).
- [34] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, *Phys. Rev. B* **97**, 085104 (2018).
- [35] M. Karl, H. Cakir, J. C. Halimeh, M. K. Oberthaler, M. Kastner, and T. Gasenzer, *Phys. Rev. E* **96**, 022110 (2017).
- [36] A. A. Ovchinnikov, D. V. Dmitriev, V. Y. Krivnov, and V. O. Chervanovskii, *Phys. Rev. B* **68**, 214406 (2003).
- [37] G. E. Hinton, "A practical guide to training restricted Boltzmann machines (Version 1)," (2010).
- [38] K. Sohn and H. Lee, in *Proceedings of the 29th International Conference on Machine Learning*, ICML'12 (Omnipress, USA, 2012) pp. 1339–1346.
- [39] B. Rubenstein, "Introduction to the variational monte carlo method in quantum chemistry and physics," in *Variational Methods in Molecular Modeling*, edited by J. Wu (Springer Singapore, Singapore, 2017) pp. 285–313.
- [40] R. Orús, *Ann. Phys.* **349**, 117 (2014).
- [41] Y. Levine, D. Yakira, N. Cohen, and A. Shashua, [arXiv:1704.01552 \[cs.LG\]](https://arxiv.org/abs/1704.01552) (2017).
- [42] J. Schachenmayer, A. Pikovski, and A. M. Rey, *New J. Phys.* **17**, 065009 (2015).
- [43] J. Schachenmayer, A. Pikovski, and A. M. Rey, *Phys. Rev. X* **5**, 011022 (2015).
- [44] L. Pucci, A. Roy, and M. Kastner, *Phys. Rev. B* **93**, 174302 (2016).
- [45] A. Piñeiro Orioli, A. Safavi-Naini, M. L. Wall, and A. M. Rey, *Phys. Rev. A* **96**, 033607 (2017).
- [46] M. Schmitt and M. Heyl, [arXiv:1707.06656 \[cond-mat.str-el\]](https://arxiv.org/abs/1707.06656) (2017).
- [47] W. K. Wootters, *Ann. Phys.* **176**, 1 (1987).
- [48] E. Wigner, *Phys. Rev.* **40**, 749 (1932).
- [49] A. Polkovnikov, *Ann. Phys.* **325**, 1790 (2010).
- [50] S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).
- [51] S. Sorella, *Phys. Rev. B* **64**, 024512 (2001).
- [52] G. Carleo, F. Becca, L. Sanchez-Palencia, S. Sorella, and M. Fabrizio, *Phys. Rev. A* **89**, 031602 (2014).
- [53] G. Carleo, F. Becca, M. Schiό, and M. Fabrizio, *Sci. Rep.* **2**, (2012).

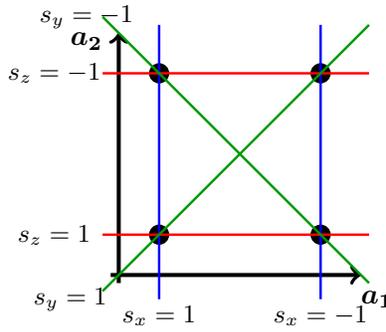


FIG. 5. Discrete quantum phase space for a spin- $\frac{1}{2}$ system spanned by two variables a_1 , a_2 . The colored lines denote the three sets of parallel lines in the 2×2 dimensional finite mathematical field. With each set a spin operator σ^α , $\alpha \in \{x, y, z\}$ is associated and each line is identified with one eigenvalue of the corresponding operator.

Appendix

Appendix A: The discrete truncated Wigner approximation

The discrete truncated Wigner approximation (dTWA) is based on sampling an initial spin state from a Wigner function on a discrete phase space and classically evolving it in time. By repeating this, one can create lots of trajectories and by averaging the outcoming observables, their time evolution can be approximated semi-classically [42, 44].

A discrete phase space of a quantum spin- $\frac{1}{2}$ system is based on a 2×2 dimensional finite mathematical field, which consists of three sets of parallel lines. As in a continuous phase space, with each set of parallel lines one operator is associated and each line is identified with an eigenvalue of the corresponding operator [47]. For a spin- $\frac{1}{2}$ system we associate the spin operators, which are the Pauli operators σ^x , σ^y and σ^z , with the sets of parallel lines and each line is identified with either the $+1$ or -1 eigenvalue, as illustrated in Fig. 5. For each point $\alpha = (a_1, a_2)$ in the phase space, a phase point operator A_α can be defined, which maps each point in the Hilbert space onto a point in phase space. Here, a convenient choice for the phase point operators is [42, 47]

$$A_\alpha = \frac{1}{2} \left[(-1)^{a_1} \sigma^x + (-1)^{a_1+a_2} \sigma^y + (-1)^{a_2} \sigma^z + \mathbf{1} \right], \quad (\text{A1})$$

which is consistent with the association of the spin operator eigenvalues in Fig. 5.

By mapping the density operator ρ onto the phase space, the Wigner function is defined as

$$W_\alpha = \frac{1}{2} \text{Tr} (\rho A_\alpha), \quad (\text{A2})$$

which is a quasi-probability distribution over the phase space. This means it gives a probability for each point in phase space and shows properties of a probability distribution, but it might have negative values [47–49]. By combining two phase spaces with different phase point operators, all eight possible orientations of the discrete spin can be captured and a quasi-probability for each orientation is given [42, 44].

From the Wigner function on the discrete phase space, initial spin states can be sampled by creating one phase space for each site and sampling the orientation of the corresponding spin. This is only possible if the Wigner function is non-negative, which limits the flexibility of the simulation method. The time evolution of the sampled spin state can then be approximated by treating the quantum spins as classical spins \mathbf{s}_i and evolving them in time individually using classical equations of motion [42],

$$\dot{s}_i^\alpha = \{s_i^\alpha, \mathcal{H}\}, \quad (\text{A3})$$

with $\alpha \in \{x, y, z\}$ and $i \in \{1, \dots, N\}$ for an N -site spin system. The brackets denote the Poisson bracket, which for two classical spins is defined as

$$\{s_i^\alpha, s_j^\beta\} = 2\delta_{ij} \sum_\gamma \epsilon^{\alpha\beta\gamma} s_i^\gamma. \quad (\text{A4})$$

In the dTWA, R classical trajectories are calculated, where the initial state is sampled from the Wigner function for each trajectory individually. The observables resulting from the trajectories are then averaged to approximate quantum dynamics in a semi-classical way, since a classical time evolution is used, but quantum fluctuations come into the system by sampling the initial states.

In the main text, sudden quenches in the transverse field Ising model are considered. The quenches are chosen to start from a large initial transverse field $h_{x,i} = 100$, which is sufficiently large to create a fully x -polarized ground state, for which the Wigner function is non-negative, making it possible to calculate the dTWA for the considered quenches. To reach more accurate results, we improved the method by trying higher orders in the approximation of the equations of motion [44]. For this we observed that taking only the next order into account leads to numerical instabilities, so that even higher orders would be necessary.

Appendix B: The artificial neural network approach

The recently introduced artificial neural network (ANN) approach is based on representing a quantum state in terms of the weights in a restricted Boltzmann machine [20]. Using the setup introduced in the main text with $M \times N$ complex connecting weights $W_{i,j}$ between the N visible and M hidden neurons, as well as N visible biases a_i and M hidden biases b_j , the wave function $|\Psi\rangle$ of a quantum spin state can be written as

$$|\Psi\rangle = \sum_{\mathbf{v}} c_{\mathbf{v}}(\mathcal{W}) |\mathbf{v}\rangle, \quad (\text{B1})$$

$$\text{with } c_{\mathbf{v}}(\mathcal{W}) = \sum_{h_j \in \{\pm 1\}} \exp\left(\sum_{i=1}^N \sum_{j=1}^M v_i W_{i,j} h_j + \sum_{i=1}^N a_i v_i + \sum_{j=1}^M b_j h_j\right) \quad (\text{B2})$$

$$= \exp\left(\sum_{i=1}^N a_i v_i\right) \prod_{j=1}^M 2 \cosh\left(b_j + \sum_{i=1}^N v_i W_{i,j}\right), \quad (\text{B3})$$

with the vector $\mathcal{W} = (a_1, \dots, a_N, b_1, \dots, b_M, W_{1,1}, \dots, W_{N,M})$ of all weight variables.

In a general restricted Boltzmann machine with real weights, the expression for $c_{\mathbf{v}}(\mathcal{W})$ would correspond to the probability assigned to a configuration of visible variables [37], but since we have complex weights here, also $c_{\mathbf{v}}(\mathcal{W})$ is complex and does not describe a probability. Instead its square corresponds to the probability of the configuration \mathbf{v} , as $c_{\mathbf{v}}(\mathcal{W})$ is the prefactor of the product state $|\mathbf{v}\rangle$ in the wave function.

With this representation, the ground state wave function of a quantum spin- $\frac{1}{2}$ system can be found using an iterative scheme based on a Stochastic Reconfiguration method [50, 51]. The weight parameters \mathcal{W} at iteration step $p+1$ can be calculated using the vector of forces F and the covariance matrix S [20],

$$\mathcal{W}(p+1) = \mathcal{W}(p) - \gamma S^{-1}(p) F(p), \quad (\text{B4})$$

$$\text{with } S_{kk'} = \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle - \langle \mathcal{O}_k^* \rangle \langle \mathcal{O}_{k'} \rangle, \quad (\text{B5})$$

$$F_k(p) = \langle E_{\text{loc}} \mathcal{O}_k^* \rangle - \langle E_{\text{loc}} \rangle \langle \mathcal{O}_k^* \rangle, \quad (\text{B6})$$

with an iteration step size γ and the star denoting complex conjugation. To calculate the inverse of S in a stable way, a regularization method is used [20]. The local energy E_{loc} and the variational derivative \mathcal{O}_k are defined as

$$E_{\text{loc}}(\mathbf{v}) = \frac{\langle \mathbf{v} | \mathcal{H} | \Psi \rangle}{c_{\mathbf{v}}}, \quad (\text{B7})$$

$$\mathcal{O}_k(\mathbf{v}) = \frac{1}{c_{\mathbf{v}}} \partial_{\mathcal{W}_k} c_{\mathbf{v}}. \quad (\text{B8})$$

Here, the Hamiltonian \mathcal{H} of the quantum spin system under consideration enters into the algorithm. The expectation values are generally calculated as

$$\langle \mathcal{O} \rangle = \langle \Psi | \mathcal{O} | \Psi \rangle \quad (\text{B9})$$

$$= \sum_{\mathbf{v}, \tilde{\mathbf{v}}} \langle \mathbf{v} | \mathcal{O} | \tilde{\mathbf{v}} \rangle c_{\tilde{\mathbf{v}}} c_{\mathbf{v}}, \quad (\text{B10})$$

where the summations run over all states in the Hilbert space, which are 2^N states for an N -site spin system.

Since such a large number of states can not be taken into account for large spin systems, it is more convenient to use only a subset of states, which is generated by a Markov chain set up by sampling $|c_{\mathbf{v}}|^2$ using a Metropolis-Hastings algorithm [20]. The Markov chain is created starting from a random initial state \mathbf{v}^k by flipping a random spin to get a new configuration $\tilde{\mathbf{v}}$. The new configuration is accepted with probability $A(\mathbf{v}^k \rightarrow \tilde{\mathbf{v}}) = \min\left(1, |c_{\tilde{\mathbf{v}}}/c_{\mathbf{v}^k}|^2\right)$. If accepted, the configuration is updated, $\mathbf{v}^{k+1} = \tilde{\mathbf{v}}$, if rejected, the state stays the same, $\mathbf{v}^{k+1} = \mathbf{v}^k$. This way, the Markov chain is set up in an iterative scheme and in the end creates a subset of states with large coefficients $|c_{\mathbf{v}}|^2$. Since these states have high contributions to the expectation values, this gives a reasonable approximation, since states with smaller contributions can be neglected. This technique is used routinely in variational Monte Carlo and is known to be stable and efficient.

For small system sizes, as considered in the main text, this sampling procedure is not necessary, since the expectation values can be calculated exactly due to the small dimension of the Hilbert space.

The iteration scheme to find the ground state can be interpreted as an effective imaginary time evolution and in an analogous way an iteration scheme to approximate the real time evolution of the quantum spin system can be derived using a time-dependent variational Monte Carlo approach [20, 52, 53]. The equations of motion for the weight parameters \mathcal{W} again depend on the vector of forces F and the covariance matrix S , which are defined in the same way as above and give

$$\dot{\mathcal{W}}(t) = -iS^{-1}(t)F(t). \quad (\text{B11})$$

These can be easily integrated numerically. Here, S^{-1} denotes the Moore-Penrose pseudo-inverse of S , which is not necessarily invertible since it is not guaranteed to have full rank. To get stable results even for large spin systems and in the vicinity of a quantum critical point, we found that the pseudo-inverse needs to be combined with the regularization method used in the ground state calculation.

Given this, the simulations of the dynamics after sudden quenches as considered in the main text can be calculated by starting with random weights, converging to the initial ground state and calculating the time evolution starting from the ground state weights using the parameters after the quench.

By introducing symmetries into the ANN setup, the number of weight parameters can be reduced. In the case of spin chains with periodic boundary conditions, as discussed in the main text, we find translation invariance. To include this into the ANN approach, we shift all visible variables by a factor d around the ring and force the new configuration to have the same weight parameters as the old configuration. Doing so reduces the number of weight parameters from $M + N + MN$ to $M + \frac{M}{N} + 1$ [20, 38].

Appendix C: Transverse Field Ising Model

The one-dimensional transverse field Ising model (TFIM) with N spin- $\frac{1}{2}$ sites is described by the Hamiltonian

$$H = -J \sum_{i=1}^N \sigma_i^z \sigma_{(i+1) \bmod N}^z - h_x \sum_{i=1}^N \sigma_i^x, \quad (\text{C1})$$

with the Pauli matrices σ_i^α , and we choose $J = 1$ without loss of generality. This model is integrable, since it can be mapped onto non-interacting fermions, from which spectrum and energy eigenstates can be calculated analytically [1–3]. The model shows a quantum phase transition at $h_c = \pm 1$ between a paramagnetic ($|h_x| > 1$) and a ferromagnetic ($|h_x| < 1$) phase. This quantum phase transition is quantified by a gapless dispersion relation [2, 3, 35]. It has been shown that for sudden quenches from a large transverse field into the vicinity of the quantum critical point, as we are considering them in the main text, the correlation length shows the behavior of a generalized Gibbs ensemble (GGE), so it increases for smaller distances from the quantum critical point [2, 3, 35].

Fig. 6(a) shows this behavior of the correlation length at a fixed time $t = 1$ after quenches within the paramagnetic phase for a spin chain with $N = 10$ sites, where the distance $\epsilon = (h_{x,f} - h_c)/h_c$ from the quantum critical point is plotted on a logarithmic scale. The different simulation methods are compared with the exact solution of the model. The correlation length is here extracted from the equal-time correlation function $C_d^{zz}(t)$ by fitting an exponential function to the short distance decay of $C_d^{zz} \propto \exp(-d\xi^{-1})$ for small d . The solid line shows the GGE behavior, which is given by [2, 3]

$$\xi_{GGE}(\epsilon) = \frac{1}{\ln(2(\epsilon h_c + h_c))}. \quad (\text{C2})$$

The exact solution describes the GGE behavior well except for the region around $\epsilon = 10^{-1}$. Here, the exact solution is not yet saturated and a longer evolution time is needed until it is converged onto the GGE curve. Since we are

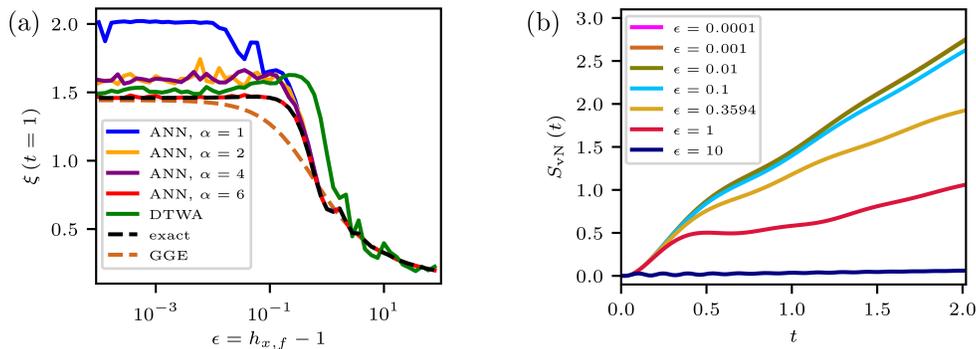


FIG. 6. Correlation length at a fixed time (a), time evolution of the von-Neumann entanglement entropy (b) after a sudden quench from $(h_{x,i} = 100, h_{z,i} = 0)$ to different distances ϵ from the quantum critical point in the TFIM with $h_{z,f} = 0$. In the region of large correlation lengths, large α are needed in the ANN approach to capture the exact solution. In the same region also volume-law entanglement is found.

only considering systems with $N = 10$ sites here, the correlation length does not saturate before finite size effects appear, so the GGE curve is never reached completely. As already discussed in the main text, the ANN approach needs larger α to capture the exact solution in the regime of large correlation lengths, while it works very well even for $\alpha = 1$ at small correlation lengths. This behavior can also be observed in Fig. 6(a). It can also be seen that the dTWA is closer to the exact solution in the vicinity of the QCP, but gets worse when the transverse field increases, as discussed in the main text.

Fig. 6(b) shows the time evolution of the von-Neumann entanglement entropy, which we calculate using tDMRG. There one can see that for $h_{x,f} < 2$ the entropy grows linearly with time, while it stays constant for $h_{x,f} \geq 2$. To better see this linear growth, the tDMRG calculations are done for a chain with $N = 40$ sites, so that finite size effects do not appear on this time scale. Here one can directly see that in the regime where large α are necessary in the ANN approach, not only the correlation length is large, but also volume law entanglement is found. This volume law entanglement also limits tDMRG calculations in this regime for large spin chains and longer times, so that the ANN approach is limited in the same regime as the tDMRG calculations in the TFIM.

If we quench the transverse field to $h_{x,f} < 0$, we again find a gapless dispersion relation at $h_{c,2} = -1$, where the gap now closes at the edge of the Brillouin zone, while it closes in the middle for $h_c = 1$. Hence, we find another quantum phase transition between a paramagnetic ($h_{x,f} < -1$) and a ferromagnetic ($h_{x,f} > -1$) phase. In this regime, the nearest-neighbor correlation function gets negative. The correlation length can then be extracted from the absolute value of the correlation function in the same way as before, which results in the symmetric curve around $h_{x,f} = 0$ plotted in the main text in Fig. 2(a).

Appendix D: Ising Model in Transverse and Longitudinal Field

If a longitudinal field h_z is added to the TFIM, the Hamiltonian becomes

$$H = -J \sum_{i=1}^N \sigma_i^z \sigma_{(i+1) \bmod N}^z - h_x \sum_{i=1}^N \sigma_i^x - h_z \sum_{i=1}^N \sigma_i^z, \quad (\text{D1})$$

and the model is not integrable any more. Again we can choose $J = 1$ without loss of generality and we again find a paramagnetic phase for large h_x and small h_z and a ferromagnetic phase for small h_x . The difference to the TFIM is now that there is no quantum phase transition between the phases, since the ground state in the ferromagnetic regime is not degenerate anymore due to the longitudinal field and hence there is no spontaneous symmetry breaking between the phases [36].

Considering now quenches from $(h_{x,i} = 100, h_{z,i} = 0)$ to different $h_{x,f}$ and $h_{z,f}$, the spin dynamics show Rabi oscillations due to the interaction of h_x and h_z . These oscillations can also be found in the correlation length extracted from the equal-time correlation function $C_d^{zz}(t)$ in the same way as discussed earlier. The time evolution of the correlation length is plotted in Fig. 7 for quenches to different $h_{x,f}$ with $h_{z,f} = 0$ (a), $h_{z,f} = 1$ (b), $h_{z,f} = 2$ (c) and $h_{z,f} = 3$ (d). We compare the simulation results with exact diagonalization calculations, since the model is not analytically solvable anymore. We can observe in all plots that the ANN approach shows fluctuations for $\alpha = 1$ at times $t > 0.5$ in the regime where the correlation lengths get large at shorter times. For $\alpha = 10$, the method can

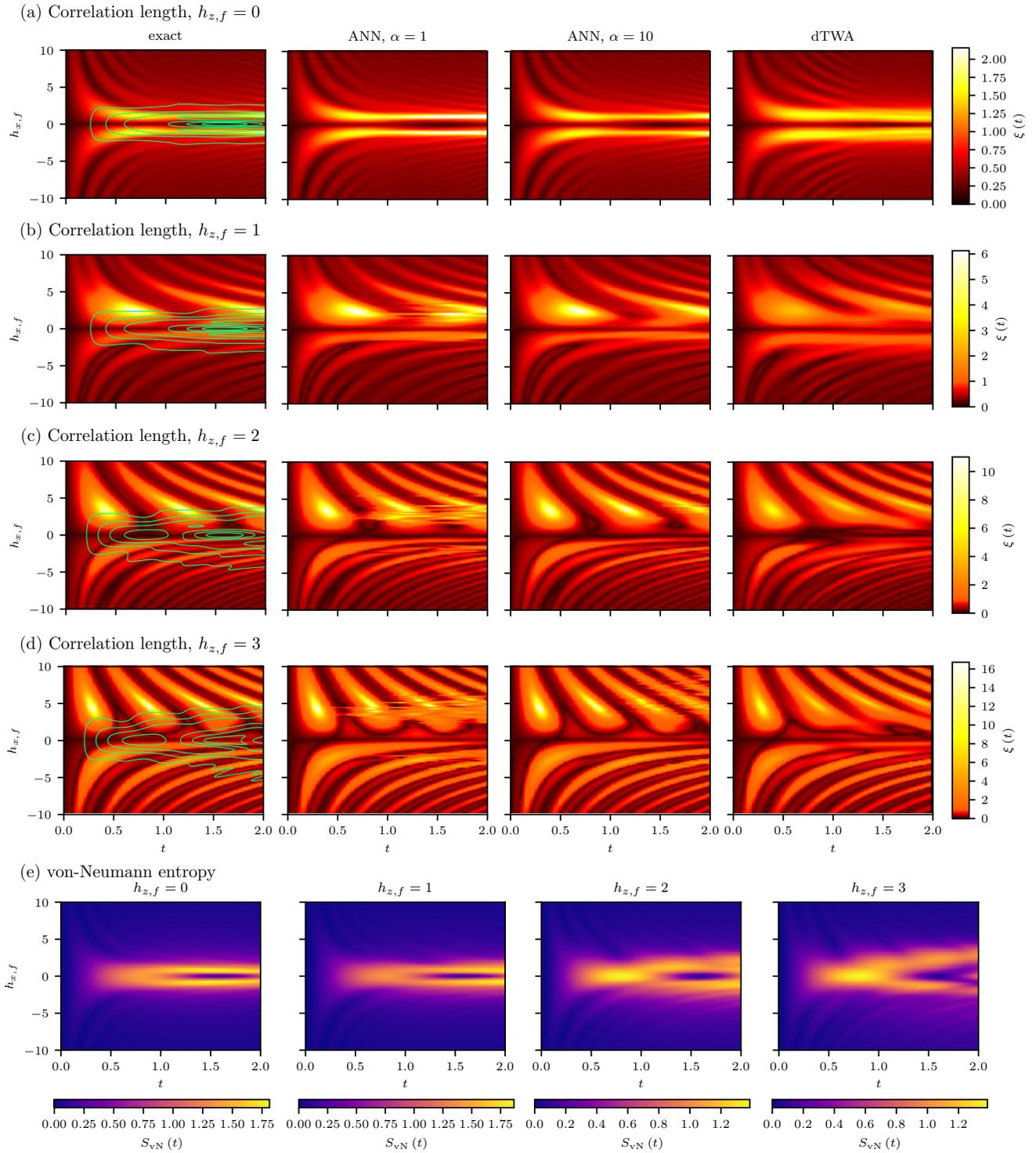


FIG. 7. Time evolution of correlation length and von-Neumann entropy after sudden quenches from $(h_{x,i} = 100, h_{z,i} = 0)$ to different values of $h_{x,f}$ with $h_{z,f} = 0$ (a), $h_{z,f} = 1$ (b), $h_{z,f} = 2$ (c) and $h_{z,f} = 3$ (d). The ANN approach with $\alpha = 1$ and $\alpha = 10$ as well as the dTWA are compared with exact results. (e) shows the time evolution of the von-Neumann entanglement entropy extracted from the tDMRG calculations for $h_{z,f} = 1$, $h_{z,f} = 2$ and $h_{z,f} = 3$. The entanglement entropy is also plotted as contours in the color plot of the exact correlation length in (a), (b), (c) and (d). There one can directly see that large entanglement entropy and volume-law entanglement can be found in the same regimes as large correlation lengths.

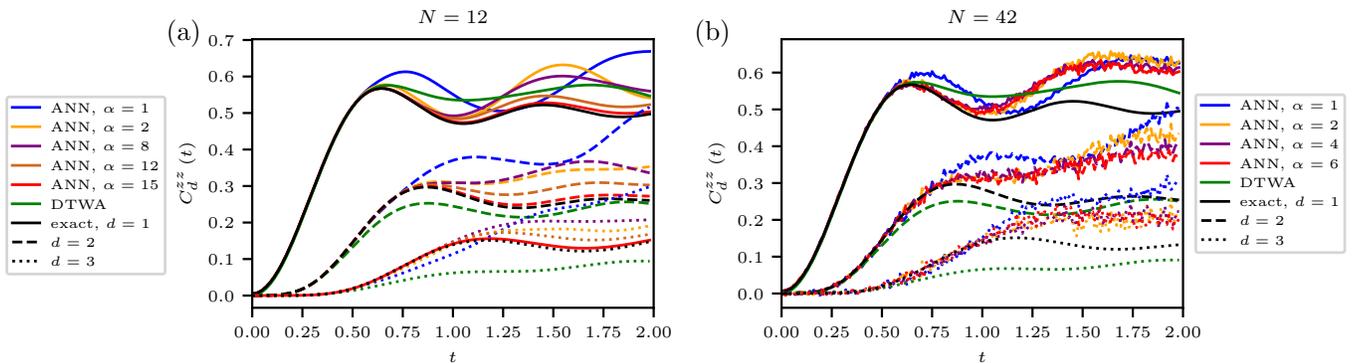


FIG. 8. Time evolution of the correlation function $C_d^{zz}(t)$ after a sudden quench from $(h_{x,i} = 100, h_{z,i} = 0)$ to $(h_{x,f} = 1, h_{z,f} = 0)$ in a spin chain with $N = 12$ (a) and $N = 42$ (b) sites. The correlation function is shown for distances $d = 1$ (solid lines), $d = 2$ (dashed lines) and $d = 3$ (dotted lines) between the considered spins. The number of weight parameters in the ANN approach is increased to see that $\alpha = 15$ is necessary to capture the exact dynamics for $N = 12$, while full convergence can not be reached for $N = 42$ within suitable computation time.

capture the exact solution quite well, only small deviations appear at very late times. It is interesting to see that even for $\alpha = 1$ the first oscillation is always captured perfectly, it only breaks down at later oscillations. For $h_{z,f} = 0$ deviations in the $\alpha = 1$ calculations only appear in the regime of large correlation length, where the simulations reach too large values. This is directly at the quantum phase transition at $h_{x,f} = 1$, where also volume-law entanglement can be found.

In Fig. 7(e) the time evolution of the von-Neumann entanglement entropy is shown for $h_{z,f} = 1$, $h_{z,f} = 2$ and $h_{z,f} = 3$, which is calculated using tDMRG. Here we can see that the entanglement entropy gets large approximately in the same regime as the correlation length, which is also where deviations in the ANN approach for small α can be found.

The dTWA also captures the oscillations in the correlation length, but at later times deviations can be found for all $h_{x,f}$. The maximum values of the correlation length are never reached in these calculations and deviations are found even at shorter times than in the ANN approach.

Appendix E: Artificial Neural Network Approach for Large Spin Systems

In the main text we mostly considered small spin chains with $N = 10$ sites. For these systems we found that in the ANN approach the number of weight parameters needs to be as large as the dimension of the Hilbert space to represent the exact dynamics in regimes of large correlations. To see how the necessary number of weight parameters depends on the system size, we consider the quench from $(h_{x,i} = 100, h_{z,i} = 0)$ to $(h_{x,f} = 1, h_{z,f} = 0)$ and compare the ANN and dTWA simulations to exact solutions for spin chains with $N = 12$ and $N = 42$ sites. For $N = 12$, the number of weight parameters can still be increased until the Hilbert space dimension is reached, while this is not possible anymore for $N = 42$. Also the Monte Carlo sampling needs to be used for the $N = 42$ calculations, since a summation over all configurations is not possible.

We have already shown in the main text that even for $N = 42$, $\alpha = 1$ is sufficient for quenches into regimes of small correlations, but larger α is needed in regimes of larger correlations even if a longitudinal field is added. For the quench we are considering here, we found in the main text that $\alpha = 6$ is necessary for $N = 10$ sites. Fig. 8 shows the time evolution of the correlation function for $N = 12$ (a) and $N = 42$ (b). There we find that for $N = 12$, $\alpha = 15$ is necessary to capture the exact dynamics, which corresponds to $1 + \alpha + M = 196$ weight parameters. This is again of the order of the Hilbert space dimension after symmetrization ($d_H = 352$), as we also found it for $N = 10$ sites. This suggests that the necessary number of weight parameters scales exponentially with the system size.

In Fig. 8(b), we increase α as far as possible within suitable computation time for $N = 42$ sites. One can see the convergence to the exact dynamics with increasing α , but for $\alpha = 6$ the result is still far away from the exact solution. The small fluctuations in the ANN calculations are caused by the finite Monte Carlo sampling. A much larger α would be necessary here. This directly shows the limitations of the ANN approach for large spin systems in regimes of large correlations due to the exponential scaling of α with system size, while we have shown in the main text that the method works perfectly fine for large spin systems and small α in regimes of small correlations.