Time-Dependent Variational Principle for Open Quantum Systems with Artificial Neural Networks

Moritz Reh[1,*], Markus Schmitt[2] and Martin Gärttner[1,3,4]

[1]*Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[2]*Institut für Theoretische Physik, Universität zu Köln, 50937 Köln, Germany*
[3]*Physikalisches Institut, Universität Heidelberg, Im Neuenheimer Feld 226, 69120 Heidelberg, Germany*
[4]*Institut für Theoretische Physik, Ruprecht-Karls-Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*

We develop a variational approach to simulating the dynamics of open quantum many-body systems using deep autoregressive neural networks. The parameters of a compressed representation of a mixed quantum state are adapted dynamically according to the Lindblad master equation by employing a time-dependent variational principle. We illustrate our approach by solving the dissipative quantum Heisenberg model in one dimension for up to 40 spins and in two dimensions for a 4×4 system and by applying it to the simulation of confinement dynamics in the presence of dissipation.

Introduction.—Solving the quantum many-body problem where it is analytically intractable constitutes a formidable challenge due to the inherent curse of dimensionality with growing system size. Today, two main routes are pursued to address this issue. On the one hand, the boundaries of classical computation are pushed by the development of tailored numerical techniques that build on the inherent structure of the quantum state of interest to find compressed representations using a subexponential number of variational parameters [1–4]. On the other hand, recent years have brought tremendous progress in the realization of quantum simulators as originally envisioned by Feynman [5,6], which emulate paradigmatic quantum many-body models using precisely controlled synthetic quantum systems of ultracold atoms in optical lattices [7–13], trapped ions [14–16], Rydberg atoms [17–19], and many more [20–31]. These "noisy intermediate-scale quantum" (NISQ) simulators [32] already present a valuable expansion of our scientific toolbox, enabling the discovery of new physical phenomena [12,18,19,31,33–35]. In particular, they challenge the numerical state of the art and open up largely uncharted terrain, e.g., nonequilibrium quantum matter in two spatial dimensions. As the term NISQ implies, the openness of these quantum systems will play a central role for near-term applications, and accounting for it appropriately is one of the key challenges.

In this work, we present a novel way to simulate the dynamics of open quantum systems (OQS) using a neural network encoding of the quantum state, which is relevant for two reasons: In view of the recent experimental developments, computational tools that can keep up with the system sizes of quantum simulators also in intermediate spatial dimensions are highly desired as they allow us to certify experimental observations and provide a link to

theoretical models. Simultaneously, the exploration of phenomena associated with driven dissipative systems is a major aim in itself, for which our approach opens new possibilities.

The state of an OQS is described by the density operator $\hat{\rho}$, whose dynamics, for Markovian systems, is governed by a Lindblad master equation. For a system of N spin-1/2 particles considered here the curse of dimensionality manifests in the 4^N coefficients necessary to fully represent $\hat{\rho}$, which limits exact numerical treatments to small N. Various numerical methods have been developed to reduce this complexity [36], each coming with different strengths and limitations. Stochastic Monte Carlo wave function (MCWF) methods [37–41] achieve a quadratic improvement of the N scaling at the cost of requiring statistical averaging. Semiclassical [42–45] and mean-field-like methods [46,47] provide a polynomial scaling in N but often suffer from uncontrolled approximations and numerical instabilities. Tensor network based approaches [1,2,48–56] are limited to weakly entangled states and require further approximations if applied in dimensions $d > 1$ [57–59]. A recently introduced class of methods, that can potentially resolve many of these issues are neural network quantum states (NQS) [60–75]. NQS have been applied successfully to OQS [76–79]. A natural approach is to employ a latent state purification [80]; however, this procedure has so far been restricted to shallow neural network architectures. A more recent work uses a probabilistic representation of the quantum state [64] which allows the use of deeper, more expressive networks but has the drawback of being forced to globally optimize the network parameters in each time step.

Here we introduce a numerical approach, summarized graphically in Fig. 1, that is not restricted in terms of

network architectures and operates based on explicit second-order local updates, thus overcoming structural and technical limitations of previously proposed methods. The derivation of a first-order differential equation for the time dependence of variational parameters in the context of a probabilistic formulation of quantum mechanics is a central result of our work. Thereby, our method expands the capabilities of previous approaches [64,76] in terms of system sizes and timescales reached reliably. This is demonstrated by the application to benchmark problems of spin systems in 1D and 2D geometries and by showing a first physics-motivated application.

Probabilistic representation.—Any quantum state $\hat{\rho}$ can be represented equivalently as a probability distribution P over measurement outcomes using positive operator valued measures (POVMs) [61,62,64,81]:

$$P^{\mathbf{a}} = \mathrm{tr}(\hat{\rho}\hat{M}^{\mathbf{a}}), \qquad (1)$$

where $\hat{M}^{\mathbf{a}} = \hat{M}^{a_1} \otimes .. \otimes \hat{M}^{a_N}$ are measurement operators associated with the outcome $\mathbf{a} = a_1..a_N$ of a tomographically complete measurement on N spins. We choose $\hat{M}^i$ to be the symmetric informationally complete-POVM (SIC-POVM), or tetrahedral POVM [61]. Its elements are obtained from the definition $M^a = (\mathbb{1} + \vec{s}^a \cdot \vec{\sigma})/4$, in which the $\vec{s}^a$ form a tetrahedron on the surface of the Bloch sphere and $\vec{\sigma}$ denotes the vector of Pauli matrices. Inverting Eq. (1) gives

$$\hat{\rho} = P^{\mathbf{a}} T^{-1\mathbf{a}\mathbf{a}'} \hat{M}^{\mathbf{a}'} \qquad (2)$$

with the overlap matrix $T^{\mathbf{a}\mathbf{a}'} = \mathrm{tr}(\hat{M}^{\mathbf{a}}\hat{M}^{\mathbf{a}'})$, where implicit summation over repeated indices is assumed from here on. Since the POVM elements $\hat{M}^{\mathbf{a}}$ form an operator basis, observables can be decomposed as $\hat{O} = \Omega^{\mathbf{a}}\hat{M}^{\mathbf{a}}$ and their expectation values become $\langle\hat{O}\rangle = P^{\mathbf{a}}\Omega^{\mathbf{a}}$. Compared to the complex-valued density matrix or its purification, the probabilistic representation has the advantage that it allows us to directly leverage the highly sophisticated toolbox for generative models developed in recent years by the machine learning (ML) community [82–84].

The dynamics of Markovian OQS is described by the Lindblad master equation [36]

$$\dot{\hat{\rho}} = -i[\hat{H}, \hat{\rho}] + \gamma\sum_i (\hat{L}^i\hat{\rho}\hat{L}^{i\dagger} - \frac{1}{2}\{\hat{L}^{i\dagger}\hat{L}^i, \hat{\rho}\}), \quad (3)$$

with $[.,.]$ ($\{\cdot,\cdot\}$) denoting the (anti-)commutator. The operators $\hat{L}^i$ are commonly referred to as jump operators and are representative of the dissipative processes that the system is subject to. Differentiating Eq. (1) and inserting Eqs. (3) and (2) allows us to state the master equation in the probabilistic formulation:
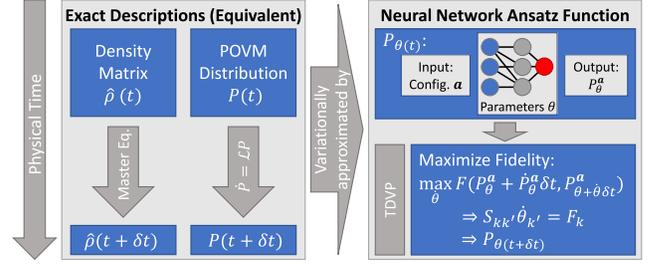
FIG. 1. Illustration of the variational approach to OQS dynamics. On the left, the standard density matrix formalism is shown along with the equivalent probabilistic formulation using POVMs. The right-hand side shows the variational approach, in which an artificial neural network is used as an ansatz function for the probability distribution over POVM outcomes. The main contribution of this work is the TDVP illustrated in the bottom box which leads to a general and accurate scheme for updating the network parameters according to the dynamics dictated by the master equation.

$$\dot{P}^{\mathbf{a}} = \mathcal{L}^{\mathbf{a}\mathbf{b}}P^{\mathbf{b}}. \qquad (4)$$

The full expression for the Lindbladian $\mathcal{L}$ is given in the Supplemental Material [85]. Crucially, $\mathcal{L}$ is sparse since the restriction to one- and two-body interactions in Eq. (3) is preserved in the probabilistic reformulation, allowing us to evaluate its action efficiently.

Nonetheless, Eq. (4) is numerically intractable for many-body systems, because of the exponentially large number of coefficients $P^{\mathbf{a}}$. In the following we employ a variational approximation by introducing a trial distribution $P^{\mathbf{a}}_{\theta}$ with variational parameters θ, see Fig. 1. The compressed representation of the state in a polynomial number of variational parameters renders the approach numerically feasible.

Time dependent variational principle for POVMs.—The main theoretical contribution of our work is a time dependent variational principle (TDVP) for POVM-probability distributions which dictates the time dependence of the network parameters $\theta(t)$ by determining the closest approximation of the Lindbladian dynamics within the variational manifold. The starting point is a distance measure $\mathcal{D}(P, Q)$ for probability distributions P, Q. Assuming a small time step τ at time t with associated network parameters $\theta(t)$, the aim is to minimize the distance between the updated POVM-probability distribution $P^{\mathbf{a}}_{\theta(t)+\dot{\theta}\tau}$ and the time-propagated distribution $P^{\mathbf{a}}_{\theta(t)} + \tau\mathcal{L}^{\mathbf{a}\mathbf{b}}P^{\mathbf{b}}_{\theta(t)}$ that follows from Eq. (4). We found that two natural choices for the distance measure $\mathcal{D}$ are equivalent for this purpose, because they describe locally identical geometries. The first is the Hellinger distance $\mathcal{D}_H(P, Q) = 1 - F(P, Q)$, which is defined via the Bhattacharyya coefficient (or classical fidelity) $F(P, Q) = \sum_{\mathbf{a}}\sqrt{P^{\mathbf{a}}Q^{\mathbf{a}}}$. The second is the Kullback-Leibler divergence $\mathcal{D}_{\mathrm{KL}}(P, Q) = \sum_{\mathbf{a}} P^{\mathbf{a}}\log(P^{\mathbf{a}}/Q^{\mathbf{a}})$. In both cases, a second order

consistent small-τ expansion of $\mathcal{D}_{H/KL}(P^{\mathbf{a}}_{\theta(t)+\dot\theta\tau}, P^{\mathbf{a}}_{\theta(t)} + \tau \mathcal{L}^{\mathbf{ab}} P^{\mathbf{b}}_{\theta(t)})$ and subsequently demanding stationarity to find the optimal parameter update $\dot\theta$ yields the TDVP equation

$$S_{kk'}\dot\theta_{k'} = F_k. \qquad (5)$$

Here, S denotes the Fisher-metric $S_{kk'} = \langle O^{\mathbf{a}}_k O^{\mathbf{a}}_{k'}\rangle^c_{\mathbf{a}\sim P}$, $F_k = \langle O^{\mathbf{a}}_k \mathcal{L}^{\mathbf{ab}}(P^{\mathbf{b}}/P^{\mathbf{a}})\rangle^c_{\mathbf{a}\sim P}$ and the repeated indices $\mathbf{a}$ inside the brackets are not summed over. The brackets denote connected correlation functions $\langle AB \rangle^c = \langle AB \rangle - \langle A \rangle \langle B \rangle$ of expectation values with respect to the POVM-distribution P and $O^{\mathbf{a}}_k = \partial_{\theta_k} \log P^{\mathbf{a}}$. We provide a detailed derivation of Eq. (5) in the Supplemental Material [85]. It is worth noting that for models that are normalized by default [such as recurrent neural networks (RNNs)] $\langle O^{\mathbf{a}}_k \rangle = 0$. Usually, Eq. (5) is ill conditioned and needs to be regularized. Here, advanced regularization schemes such as described in Ref. [63] are applicable but they did not turn out to be crucial for the test cases we consider.

The TDVP equation (5) exhibits a number of features beneficial for the numerical time evolution of $\theta(t)$. As a result of the employed short-time expansion, the variational optimization problem becomes convex and information about the local geometry of the variational manifold is taken into account in the form of the Fisher-metric S. Upon inverting S, the differential equation can be solved straightforwardly with explicit integration schemes. Importantly, Monte Carlo sampling is only required once per time step. These features are in contrast to the implicit integration scheme presented in Ref. [64], where Monte Carlo sampling is required at each optimization step performed for the iterative global minimization of a nonconvex cost function.

Network architecture.—Neural networks are highly nonlinear universal function approximators in the limit of large networks [86–88]. For the purpose of generative modeling autoregressive networks are advantageous, because they enable direct generation of uncorrelated samples; therefore, various autoregressive architectures have recently been explored for NQS [64,65,89–91]. In the following, we employ RNNs, which belong to this family of network architectures (see Supplemental Material [85] for details).

Numerical results.—To illustrate the accuracy and scalability of our method we apply it to the anisotropic Heisenberg model,

$$\hat{H} = \sum_{\langle ij \rangle}(J_x \hat{X}_i \hat{X}_j + J_y \hat{Y}_i \hat{Y}_j + J_z \hat{Z}_i \hat{Z}_j) + \sum_i h_z \hat{Z}_i, \qquad (6)$$

with nearest neighbor interactions and periodic boundary conditions, which was also used in Ref. [64] as a benchmark system. The considered decoherence channel is spontaneous decay given by the jump operator $\hat{L} = \hat{\sigma}^- = (\hat{X} - i\hat{Y})/2$ acting on each spin. We obtain benchmark data using exact simulations for $N = 10$ spins and test our approach in the case of $N = 10$ [85] and $N = 40$ [Figs. 2(a) and 2(b)] spins, where we compare magnetizations and next-nearest neighbor correlators. Since finite-size effects are negligible to good approximation for systems with more than $N = 10$ spins, we can use the exact data for $N = 10$ spins as comparison for the case of $N = 40$ spins studied in the main text. The noise in the correlation signal is due to the finite number of samples that are used to evaluate the observables. One observes slight deviations in the correlation functions, which may be attributed to both an imperfect choice of hyperparameters and the stochastic
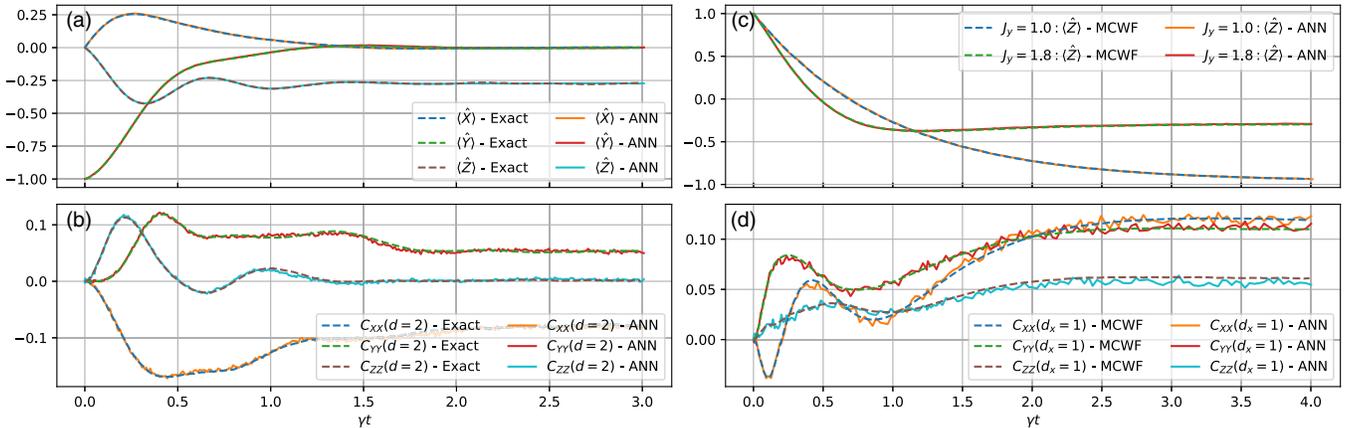
FIG. 2. (a) and (b) Mean magnetizations and next-nearest neighbor connected correlation functions [e.g., $C_{XX}(d=2) = \sum_i \langle \hat{X}_i \hat{X}_{i+2}\rangle^c/N$] as a function of time in the anisotropic 1D Heisenberg model for $N = 40$ spins starting in the product state $\langle \hat{Y} \rangle = -1$. Nearest neighbor couplings are given by $\vec{J}/\gamma = (2, 0, 1)$, $h_z/\gamma = 1$ and the dissipation channel is $\hat{L} = \hat{\sigma}^- = \frac{1}{2}(\hat{X} - i\hat{Y})$. The exact data are obtained for $N = 10$ spins. (c) and (d) Mean z magnetizations and nearest neighbor connected correlation functions (for $J_y/\gamma = 1.8$) in a 4×4 anisotropic 2D Heisenberg lattice with nearest neighbor couplings $\vec{J}/\gamma = [0.9, 1.0(1.8), 1.0]$ and the same decay as in (a) and (b), starting in the product state $\langle \hat{Z} \rangle = 1$.

nature of the proposed method. We found that the sample size needed to reach a given precision does not need to be increased when transitioning to larger systems, as the overall noise decreases thanks to a self-averaging effect in the translationally invariant system.

Figures 2(c) and 2(d) show results for a 4×4 lattice initialized in a product state with $\langle \hat{Z} \rangle = 1$. In panel (c) we compare the magnetization for two different parameter choices to results from MCWF stochastic integration with 500 trajectories, showing perfect agreement. Exact integration of Eq. (3) would be exceedingly costly in this case. We provide a comparison of a 3×3 lattice to exact dynamics in the Supplemental Material [85], which shall serve as a numerically exact benchmark. Nearest neighbor correlations shown in panel (d) for the case of $J_y/\gamma = 1.8$ show small deviations at late times, which we attribute to the finite number of samples used for estimating the updates $\dot{\theta}$ [85].

Having benchmarked our approach on generic spin models, we now apply it to a physical scenario to gauge the effect of decoherene for large problem instances. It was recently shown that confinement dynamics, as found for quarks in quantum chromodynamics, can be realized in the Ising spin model with transverse and longitudinal fields [92]

$$\hat{H} = \sum_{\langle ij \rangle} J_z \hat{Z}_i \hat{Z}_j + \sum_i (h_z \hat{Z}_i + h_x \hat{X}_i). \qquad (7)$$

In this system pairs of domain walls form after a quench. For $h_z = 0$ these domain walls can propagate freely while for finite h_z the separation between them comes with an energy cost leading to confinement. This phenomenon manifests in a buildup of dominant spin-spin correlations that is limited to short distances and a much weaker light-cone spreading due to the propagation of bound domain-wall pairs. Signatures of this effect have been observed recently for moderate system sizes on IBMQ [93].

Here we study in what way dissipation influences the signature spreading of spatio-temporal correlations. We consider single particle dephasing $\hat{L} = \hat{Z}$ as the dissipation channel, significantly altering the nature of the spreading on timescales $\gamma t \gtrsim 1$.

Figure 3 shows results for a typical scenario with a dephasing rate of $\gamma = 0.25J$ and $N = 32$. The magnetization [panel (a)] initially shows coherent oscillations (inset) which are quickly damped out followed by a slow relaxation of all magnetizations towards zero. The dashed lines show MCWF simulations for $N = 16$.

The top half of panel (b) shows the correlation dynamics in the considered dissipative scenario using the hitherto described numerical approach. For comparison, we show the corresponding unitary dynamics simulated using matrix product states (MPS) [1,2] on the bottom half of panel (b). In the unitary case correlations initially show a light-cone
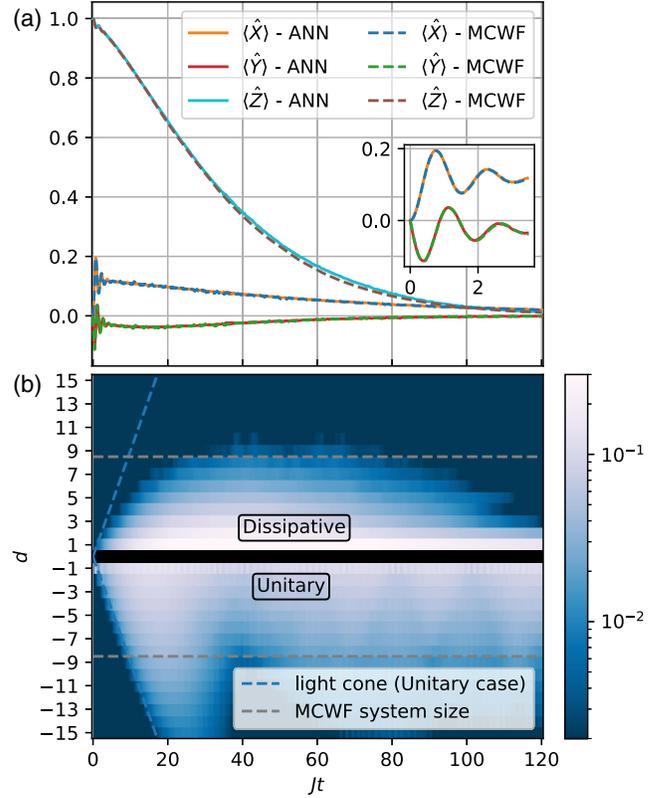
FIG. 3. (a) Mean magnetizations in a spin chain of length $N = 32$ with the quench parameters $h_x/J_z = 0.25$, $h_z/J_z = 0.05$ and the dissipation channel $\hat{L} = \hat{Z}$ with relative strength $\gamma/J_z = 0.25$ compared to MCWF-data for $N = 16$ spins starting in the product state $\langle \hat{Z} \rangle = 1$. (b) Spreading of correlations in the spin chain. Top: Dissipative system with $\gamma = 0.25$, Bottom: MPS simulation of the unitary system where $\gamma = 0.0$. After an initial linear light-cone spreading, the nature of the dissipative propagation grows more diffusive, before all correlations eventually vanish. Notice that the slight deviations in panel (a) coincide with the time at which the dissipative correlations cross the MCWF system size boundary.

spreading. In contrast, the dissipative dynamics deviates from this light cone even for short times as the dissipation results in correlation growth that we find to be consistent with diffusive spreading on intermediate timescales [85]. At long times all correlations decay and the system approaches the featureless steady state $\hat{\rho}(t \to \infty) \propto \mathbb{1}$.

The ability to simulate these dynamics is a direct consequence of the polynomial scaling of the described ansatz. The system size of the MCWF-approach ($N = 16$) is plotted in panel (b) as a dashed gray line. As is obvious from the chosen color-scale cutoff, the MCWF-approach suffers from finite-size effects at around $Jt = 40$ when correlations beyond $d = 8$ build up. This is also the time at which the z magnetization in panel (a) deviates, suggesting that this deviation is due to finite-size effects present in the MCWF simulation.

Discussion and outlook.—We have introduced a novel method that allows the variational simulation of open quantum dynamics based on the efficient encoding of the quantum state in an artificial neural network and demonstrated its potential. Our method scales approximately cubically with the number of spins N, since the computationally intense part is obtaining F in Eq. (5). For that purpose, the nonvanishing entries of $\mathcal{L}^{\mathbf{ab}}$, the cost of evaluating a plain RNN, and the incorporation of translational symmetry each contribute a factor of N to the computational cost. Importantly, however, the algorithm admits massive parallelization, e.g., on GPU clusters [63,94], which allows for extensive control of the total compute wall time. Here, different levels of parallelization are exploited: The algorithm permits treating the samples independently from each other, allowing us to employ multiple accelerators which communicate via MPI. At the same time, the batched network evaluation of the configurations allows for convenient vectorization of the computations on the single accelerator level.

For future research it will be crucial to better understand the limitations of finite network architectures to represent physically relevant quantum states. An obvious shortcoming of the probabilistic state representation is that the positivity of the density operator is not guaranteed; understanding the consequences will be key for further progress.

The presented combination of the probabilistic formulation of mixed quantum states with a TDVP opens new possibilities for the investigation of driven-dissipative many-body systems in regimes that are challenging for other approaches, for example, to study two-dimensional systems or the propagation of information across large distances [95,96]. Future work employing the TDVP for OQS could address the emergence of glassy dynamics [97–100] or self-organization in OQS [101,102]. Furthermore, the developed technique is not restricted to solving the Lindblad equation; it could be generalized for other use cases of master equations with large discrete configuration space, such as disease dynamics models [103,104] or the chemical master equation [105].

*moritz.reh@kip.uni-heidelberg.de

[1] U. Schollwöck, Ann. Phys. (Amsterdam) **326**, 96 (2011).
[2] R. Orús, Ann. Phys. (Amsterdam) **349**, 117 (2014).
[3] I. Cirac, D. Perez-Garcia, N. Schuch, and F. Verstraete, arXiv:2011.12127.
[4] J. Carrasquilla, Adv. Phys. X **5**, 1797528 (2020).
[5] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[6] S. Lloyd, Science **273**, 1073 (1996).
[7] N. R. Cooper, J. Dalibard, and I. B. Spielman, Rev. Mod. Phys. **91**, 015005 (2019).
[8] D. Mitra, P. T. Brown, E. Guardado-Sanchez, S. S. Kondov, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, Nat. Phys. **14**, 173 (2018).
[9] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, Phys. Rev. X **7**, 041047 (2017).
[10] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Luschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Science **349**, 842 (2015).
[11] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
[12] J. y. Choi, S. Hild, J. Zeiher, P. Schauss, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, Science **352**, 1547 (2016).
[13] W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, Nature (London) **462**, 74 (2009).
[14] 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, Nature (London) **534**, 516 (2016).
[15] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
[16] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Nat. Phys. **13**, 781 (2017).
[17] A. Browaeys and T. Lahaye, Nat. Phys. **16**, 132 (2020).
[18] 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. Vuletic, and M. D. Lukin, Nature (London) **595**, 227 (2021).
[19] 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, Nature (London) **595**, 233 (2021).
[20] D. I. Tsomokos, S. Ashhab, and F. Nori, Phys. Rev. A **82**, 052311 (2010).
[21] M. J. Hartmann, F. G. S. L. Brandão, and M. B. Plenio, Nat. Phys. **2**, 849 (2006).
[22] M. Ortner, A. Micheli, G. Pupillo, and P. Zoller, New J. Phys. **11**, 055045 (2009).

[23] G. Roumpos, C. P. Master, and Y. Yamamoto, Phys. Rev. B 75, 094415 (2007).

[24] T. Byrnes, N. Y. Kim, K. Kusudo, and Y. Yamamoto, Phys. Rev. B 78, 075320 (2008).

[25] A. D. Greentree, C. Tahan, J. H. Cole, and L. C. L. Hollenberg, Nat. Phys. 2, 856 (2006).

[26] X. Peng, J. Du, and D. Suter, Phys. Rev. A 71, 012307 (2005).

[27] S. Mostame and R. Schützhold, Phys. Rev. Lett. 101, 220501 (2008).

[28] D. G. Angelakis, M. F. Santos, and S. Bose, Phys. Rev. A 76, 031805(R) (2007).

[29] J. Zhang, F. M. Cucchietti, C. M. Chandrashekar, M. Laforest, C. A. Ryan, M. Ditty, A. Hubbard, J. K. Gamble, and R. Laflamme, Phys. Rev. A 79, 012305 (2009).

[30] J. Cai, A. Retzker, F. Jelezko, and M. B. Plenio, Nat. Phys. 9, 168 (2013).

[31] 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, Nature (London) 543, 221 (2017).

[32] J. Preskill, Quantum 2, 79 (2018).

[33] 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, Nature (London) 551, 579 (2017).

[34] M. Prüfer, P. Kunkel, H. Strobel, S. Lannig, D. Linnemann, C.-M. Schmied, J. Berges, T. Gasenzer, and M. K. Oberthaler, Nature (London) 563, 217 (2018).

[35] S. Erne, R. Bücker, T. Gasenzer, J. Berges, and J. Schmiedmayer, Nature (London) 563, 225 (2018).

[36] H. Weimer, A. Kshetrimayum, and R. Orús, Rev. Mod. Phys. 93, 015008 (2021).

[37] K. Mølmer, Y. Castin, and J. Dalibard, J. Opt. Soc. Am. B 10, 524 (1993).

[38] M. B. Plenio and P. L. Knight, Rev. Mod. Phys. 70, 101 (1998).

[39] J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. 68, 580 (1992).

[40] R. Dum, P. Zoller, and H. Ritsch, Phys. Rev. A 45, 4879 (1992).

[41] M. Kornyik and A. Vukics, Comput. Phys. Commun. 238, 88 (2019).

[42] I. Carusotto and C. Ciuti, Rev. Mod. Phys. 85, 299 (2013).

[43] F. Vicentini, F. Minganti, A. Biella, G. Orso, and C. Ciuti, Phys. Rev. A 99, 032115 (2019).

[44] F. Vicentini, F. Minganti, R. Rota, G. Orso, and C. Ciuti, Phys. Rev. A 97, 013853 (2018).

[45] G. Dagvadorj, J. M. Fellows, S. Matyjaśkiewicz, F. M. Marchetti, I. Carusotto, and M. H. Szymańska, Phys. Rev. X 5, 041028 (2015).

[46] R. Liboff, Kinetic Theory: Classical, Quantum, and Relativistic Descriptions (Springer-Verlag, Berlin, 2003).

[47] P. Navez and R. Schützhold, Phys. Rev. A 82, 063603 (2010).

[48] C. J. Wood, J. D. Biamonte, and D. G. Cory, Quantum Inf. Comput. 15, 759 (2015).

[49] I. A. Luchnikov, S. V. Vintskevich, H. Ouerdane, and S. N. Filippov, Phys. Rev. Lett. 122, 160401 (2019).

[50] C. M. Keever and M. H. Szymańska, Phys. Rev. X 11, 021035 (2021).

[51] D. Kilda, A. Biella, M. Schiro, R. Fazio, and J. Keeling, SciPost Phys. Core 4, 5 (2021).

[52] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Phys. Rev. Lett. 116, 237201 (2016).

[53] D. Jaschke, S. Montangero, and L. D. Carr, Quantum Sci. Technol. 4, 013001 (2018).

[54] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, New J. Phys. 12, 025012 (2010).

[55] J. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Ann. Phys. (Amsterdam) 378, 100 (2017).

[56] E. Mascarenhas, H. Flayac, and V. Savona, Phys. Rev. A 92, 022116 (2015).

[57] J. Haferkamp, D. Hangleiter, J. Eisert, and M. Gluza, Phys. Rev. Research 2, 013010 (2020).

[58] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. 98, 140506 (2007).

[59] A. Kshetrimayum, H. Weimer, and R. Orús, Nat. Commun. 8, 1291 (2017).

[60] G. Carleo and M. Troyer, Science 355, 602 (2017).

[61] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Nat. Mach. Intell. 1, 155 (2019).

[62] J. Carrasquilla, D. Luo, F. Pérez, A. Milsted, B. K. Clark, M. Volkovs, and L. Aolita, Phys. Rev. A 104, 032610 (2021).

[63] M. Schmitt and M. Heyl, Phys. Rev. Lett. 125, 100503 (2020).

[64] D. Luo, Z. Chen, J. Carrasquilla, and B. K. Clark, arXiv:2009.05580.

[65] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, Phys. Rev. Research 2, 023358 (2020).

[66] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X 7, 021021 (2017).

[67] S. Czischek, M. Garttner, and T. Gasenzer, Phys. Rev. B 98, 024311 (2018).

[68] M. Neugebauer, L. Fischer, A. Jäger, S. Czischek, S. Jochim, M. Weidemüller, and M. Gärttner, Phys. Rev. A 102, 042604 (2020).

[69] X. Gao and L.-M. Duan, Nat. Commun. 8, 662 (2017).

[70] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B 96, 195145 (2017).

[71] Z.-A. Jia, Y.-H. Zhang, Y.-C. Wu, L. Kong, G.-C. Guo, and G.-P. Guo, Phys. Rev. A 99, 012307 (2019).

[72] Y.-H. Zhang, Z.-A. Jia, Y.-C. Wu, and G.-C. Guo, arXiv:1809.08631.

[73] Y. Huang and J. E. Moore, Phys. Rev. Lett. 127, 170601 (2021).

[74] S. Lu, X. Gao, and L.-M. Duan, Phys. Rev. B 99, 155136 (2019).

[75] W.-C. Gan and F.-W. Shu, Int. J. Mod. Phys. D 26, 1743020 (2017).

[76] M. J. Hartmann and G. Carleo, Phys. Rev. Lett. 122, 250502 (2019).

[77] N. Yoshioka and R. Hamazaki, Phys. Rev. B 99, 214306 (2019).

[78] A. Nagy and V. Savona, Phys. Rev. Lett. 122, 250501 (2019).

[79] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Phys. Rev. Lett. 122, 250503 (2019).

[80] G. Torlai and R. G. Melko, Phys. Rev. Lett. 120, 240503 (2018).

[81] A. Peres, *Quantum Theory: Concepts and Methods* (Springer, Dordrecht, 2002).

[82] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning internal representations by error propagation, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations* (MIT Press, Cambridge, MA, USA, 1986), p. 318–362.

[83] S. Hochreiter and J. Schmidhuber, Neural Comput. **9**, 1735 (1997).

[84] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, arXiv:1706 .03762.

[85] See Supplemental Material at http://link.aps.org/ supplemental/10.1103/PhysRevLett.127.230501 for details on the POVM formalism and the derivation of the TDVP equation.

[86] G. Cybenko, Math. Control Signals Syst. **2**, 303 (1989).

[87] K. Hornik, Neural Netw. **4**, 251 (1991).

[88] A. Pinkus, Acta Numer. **8**, 143 (1999).

[89] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, Phys. Rev. Lett. **124**, 020503 (2020).

[90] D. Luo, Z. Chen, K. Hu, Z. Zhao, V. M. Hur, and B. K. Clark, arXiv:2101.07243.

[91] S.-H. Lin and F. Pollmann, arXiv:2104.10696.

[92] M. Kormos, M. Collura, G. Takács, and P. Calabrese, Nat. Phys. **13**, 246 (2017).

[93] J. Vovrosh and J. Knolle, Sci. Rep. **11**, 11577 (2021).

[94] L. Yang, W. Hu, and L. Li, arXiv:2011.12453.

[95] M. J. Kastoryano and J. Eisert, J. Math. Phys. (N.Y.) **54**, 102201 (2013).

[96] R. Sweke, J. Eisert, and M. Kastner, J. Phys. A **52**, 424003 (2019).

[97] D. Poletti, P. Barmettler, A. Georges, and C. Kollath, Phys. Rev. Lett. **111**, 195301 (2013).

[98] G. Kucsko, S. Choi, J. Choi, P. C. Maurer, H. Zhou, R. Landig, H. Sumiya, S. Onoda, J. Isoya, F. Jelezko, E. Demler, N. Y. Yao, and M. D. Lukin, Phys. Rev. Lett. **121**, 023601 (2018).

[99] J. Choi, S. Choi, G. Kucsko, P. C. Maurer, B. J. Shields, H. Sumiya, S. Onoda, J. Isoya, E. Demler, F. Jelezko, N. Y. Yao, and M. D. Lukin, Phys. Rev. Lett. **118**, 093601 (2017).

[100] B. Everest, I. Lesanovsky, J. P. Garrahan, and E. Levi, Phys. Rev. B **95**, 024310 (2017).

[101] B. Zhu, J. Schachenmayer, M. Xu, F. Herrera, J. G. Restrepo, M. J. Holland, and A. M. Rey, New J. Phys. **17**, 083063 (2015).

[102] S. Helmrich, A. Arias, G. Lochead, T. M. Wintermantel, M. Buchhold, S. Diehl, and S. Whitlock, Nature (London) **577**, 481 (2020).

[103] G. Jenkinson and J. Goutsias, PLoS One **7**, e36160 (2012).

[104] M. J. Keeling and J. V. Ross, J. R. Soc. Interface **5**, 171 (2008).

[105] V. Wolf, R. Goel, M. Mateescu, and T. A. Henzinger, BMC Syst. Biol. **4**, 42 (2010).

[106] J. Johansson, P. Nation, and F. Nori, Comput. Phys. Commun. **183**, 1760 (2012).

[107] M. Schmitt and M. Reh, arXiv:2108.03409.

[108] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: Composable transformations of Python + NumPy programs (2018), http://github.com/google/jax.

[109] Jülich Supercomputing Centre, J. Large-Scale Res. Facil. **5** (2019).