Neural-Network Approach to Dissipative Quantum Many-Body Dynamics

Michael J. Hartmann[1] and Giuseppe Carleo[2]

[1] *Institute of Photonics and Quantum Sciences, Heriot-Watt University Edinburgh EH14 4AS, United Kingdom**

[2] *Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

(Dated: February 15, 2019)

In experimentally realistic situations, quantum systems are never perfectly isolated and the coupling to their environment needs to be taken into account. Often, the effect of the environment can be well approximated by a Markovian master equation. However, solving this master equation for quantum many-body systems, becomes exceedingly hard due to the high dimension of the Hilbert space. Here we present an approach to the effective simulation of the dynamics of open quantum many-body systems based on machine learning techniques. We represent the mixed many-body quantum states with neural networks in the form of restricted Boltzmann machines and derive a variational Monte-Carlo algorithm for their time evolution and stationary states. We document the accuracy of the approach with numerical examples for a dissipative spin lattice system.

The description of interacting quantum many-body systems presents a formidable challenge for theoretical and numerical approaches. A pure many-body quantum state is described by the wave-function, whose complexity grows exponentially with the number of constituents. This challenge is even more pronounced for mixed quantum states, where the fundamental object describing all physical properties is the density matrix, whose degrees of freedom scale quadratically with the dimension of the Hilbert space [1]. Yet the description of experiments under realistic conditions requires modeling in terms of density matrices as the systems of interest are never perfectly isolated from their environment. The huge number of degrees of freedom of pure and mixed states however renders an exact description of large systems in general infeasible, even if one resorts to numerical approaches.

To meet this quantum complexity challenge, several approximate approaches have been developed. Tensor Networks and the Density Matrix Renormalization Group [2, 3] become efficient descriptions whenever the amount of entanglement contained in the modeled states is restricted. Despite substantial effort [4], these methods however still suffer from limitations in systems with more than one lattice dimension. Quantum Monte Carlo (QMC) methods [5, 6], in turn rely on sampling a restricted number of physically relevant configurations or perform an efficient compression of the quantum state. However, QMC approaches are effective only for a restricted number of open quantum systems and regimes [7, 8], and a severe sign problem typically emerges in the simulation of dissipative dynamics.

Recently, machine-learning inspired approaches and parameterizations of wave-functions in terms of neural networks have been introduced[9]. This variational representation, dubbed neural-network quantum states (NQS), has been used to study both system at equilibrium [9–13], and out-of-equilibrium, in the context of unitary dynamics of pure states [14–16]. The connection between NQS and tensor network representations has also been explored [11, 17, 18]. While in the past years there

has been several methodological developments to study open quantum systems using Tensor Network representations [19–26], the description of mixed states with NQS has been so-far explored for data-driven tomographic purposes [27, 28].

For modeling quantum experiments, particularly for open many-body systems [29–33], there is a strong need for efficient and accurate approaches, especially in more than one lattice dimension, where Tensor Networks face difficulties. To this end, it is instrumental to develop a flexible, and scalable numerical approach to study mixed state dynamics or stationary states of dissipative dynamics. Central to this goal is the ability to use variational density-matrix states not facing the entanglement problem, and flexible enough to describe correlations and many-body effects beyond mean-field [34, 35], and cluster approaches [36].

Here we present a machine learning approach to the simulation of dissipative quantum dynamics and its stationary states. Our approach uses a neural network parameterization for the quantum density matrix [27] and a stochastic learning method to approximate its dynamics in a time-dependent Variational Monte Carlo approach [37]. Our approach is suitable to model non-unitary dynamics of quantum systems with many degrees of freedom in a variety of settings. These include numerical characterizations of near term quantum computers where decoherence processes due to their imperfections are taken into account [38]. A second field of applications is the investigation of stationary state quantum phases and phase transitions, which have attracted increasing interest in recent years [29–33, 39, 40].

Problem and parameterization: Our aim is to solve the quantum master equation of Lindblad form,

$$\dot{\rho} = -i[H, \rho] + \sum_j \frac{\gamma_j}{2} \left(2c_j \rho c_j^\dagger - c_j^\dagger c_j \rho - \rho c_j^\dagger c_j \right) \quad (1)$$

where ρ is the density matrix of the system, H its Hamiltonian, and the γ_j and c_j the dissipation rates and jump operators of its dissipation. The index j runs over all
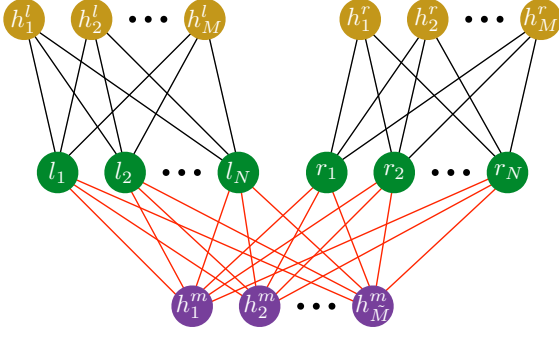
FIG. 1. Sketch of the employed neural network. Visible layer in green and hidden layers in light brown and purple. There is a hidden layer for the row indices l_i and the column indices r_i of ρ ($i = 1, 2, \ldots, N$) with hidden neurons h_j^l and h_j^r ($j = 1, \ldots, M$). A further hidden layer with neurons h_k^m ($k = 1, \ldots, \tilde{M}$) is responsible for the mixing, c.f. [27].

dissipation channels. For a large class of models, there is however only one dissipation channel per lattice site and we will restrict our treatment to this case, where j thus labels the lattice sites. As an example, we will consider a dissipative and anisotropic Heisenberg model for a lattice of N spin-1/2 degrees of freedom that has attracted significant interest recently [36].

To find an efficient and accurate approximation to the dynamics of Eq. (1), we leverage the idea that artificial neural networks can be used to provide compact representations of quantum states [9]. Specifically, we use a parametrization of the density matrix in terms of complex-valued Restricted Boltzmann Machines (RBM), similar to the one introduced in Ref. [27]. Fig. 1 shows a sketch of the specific neural network architecture used in this work. It most prominently features three sets of hidden units, $h^{(l)}$, $h^{(r)}$ and $h^{(m)}$, whose role is to mediate correlations among, respectively, column degrees of freedom of the density matrix, row degrees of freedom, and mixed correlations between the two. Because of the bipartite structure of the RBM interactions, the hidden units can be integrated out exactly, resulting in a parametrization that guarantees a Hermitian and positive semi definite density matrix [27],

$$\rho_{\vec{l}, \vec{r}} = \exp\left[\sum_{j=1}^{N}(a_j l_j + a_j^* r_j)\right] \times \prod_{k=1}^{M} \mathcal{X}_k \times \prod_{p=1}^{\tilde{M}} \mathcal{Y}_p \quad (2)$$

$$\mathcal{X}_k = \cosh\left(b_k + \sum_{j=1}^{N} W_{k,j} l_j\right) \cosh\left(b_k^* + \sum_{j=1}^{N} W_{k,j}^* r_j\right)$$

$$\mathcal{Y}_p = \cosh\left(c_p + c_p^* + \sum_{j=1}^{N}(U_{p,j} l_j + U_{p,j}^* r_j)\right),$$

where the vector indices $\vec{l} = (l_1, l_2, \ldots)$ and $\vec{r} = (r_1, r_2, \ldots)$ contain the left (right) indices $l_j(r_j)$ for all

lattice sites j, and the variational parameters are the complex-valued weights $W_{k,j}$, $U_{p,j}$ and biases a_j, b_k and c_p. Analogously to the pure state case, increasing the number of hidden units, M and $\tilde{M}$, guarantees more expressive representations of the density matrix. Given the RBM parametrization of the density matrix, it remains to be determined how to find an approximate solution of the Lindblad master equation. The approximation of the dynamics generated by Eq. (1) can be recast as a variational optimization problem, that can be approached via a suitable extension of the stochastic reconfiguration method [41] and the time-dependent Variational Monte Carlo [37] to the dissipative case.

Stochastic Reconfiguration for Liouvillians: It is convenient to write the density matrix ρ as a vector $\vec{\rho}$ such that the right hand side of Eq. (1) can be expressed as the action of a linear operator on $\vec{\rho}$, i.e. $\partial_t \vec{\rho} = \mathcal{L}\vec{\rho}$, where $\mathcal{L}$ is the Liouvillian super-operator which, in contrast to the Hamiltonian H, is not Hermitian.

According to Eq. (2), the density matrix $\vec{\rho}$ is parameterized by a set of $(N+1)(M+\tilde{M}) + N$ complex variational parameters. In the following, we use the abbreviate notation $\vec{\alpha}$, to indicate the ensemble of these variational parameters. Most notably, the real vector $\vec{\alpha}$ contains both imaginary and real parts of the variational parameters, that are treated independently. The time derivative of the variational ρ can in turn be expressed in terms of the time derivative of the variational parameters as,

$$\partial_t \vec{\rho} = \sum_k \dot{\alpha}_k O_k \vec{\rho} \quad (3)$$

where O_k denote diagonal matrices whose non-zero matrix elements read, $[O_k]_{\vec{l}, \vec{r}; \vec{l}, \vec{r}} = \partial \ln(\rho_{\vec{l}, \vec{r}})/(\partial\alpha_k)$. To get the best approximation to the dynamics of the density matrix, our goal is to find a closed equation of motion for the variational parameters, namely the time-dependence $\alpha(t)$. To this end, at each instant in time we consider the difference between the exact Lindblad infinitesimal time evolution and the approximate variational evolution,

$$\delta = \left\|\sum_k \dot{\alpha}_k O_k \vec{\rho} - \mathcal{L}\vec{\rho}\right\|_2^2, \quad (4)$$

where the time-derivatives of the variational parameters, $\dot{\alpha}_k$, are to be determined. Minimization of δ with respect to $\dot{\alpha}_k$ leads to the system of equations

$$\sum_p S_{k,p}\,\dot{\alpha}_p = f_k \quad (5)$$

where

$$S_{k,p} = \vec{\rho}^\dagger O_k^\dagger O_p \vec{\rho} + \vec{\rho}^\dagger O_p^\dagger O_k \vec{\rho} \quad (6)$$

$$f_k = \vec{\rho}^\dagger O_k^\dagger \mathcal{L}\vec{\rho} + \vec{\rho}^\dagger \mathcal{L}^\dagger O_k \vec{\rho} \quad (7)$$

and it is easy to show that the solutions of Eqs. (5) are indeed local minima of δ, see Supplemental Material.

Alternatively to the 2-norm in Eq. (4), one can also use the Fubini-Study norm, see Supplemental Material. Eq. (5) can be written as a first order differential equation,

$$\partial_t \vec{\alpha} = \mathcal{S}^{-1} \vec{f}, \tag{8}$$

where $S_{k,p}$ are the matrix elements of the matrix $\mathcal{S}$ and f_k the elements of the vector $\vec{f}$.

Stochastic sampling: The expressions in Eqs. (6-7) cannot be exactly computed for systems with a large number of quantum particles. However, those quantum expectations can be conveniently interpreted as statistical expectation values over the probability distribution

$$p(\vec{l}, \vec{r}) = |\rho_{\vec{l}, \vec{r}}|^2, \tag{9}$$

in analogy to the concept in static and time-dependent variational Monte Carlo. The elements of $\mathcal{S}$ and $\vec{f}$ can thus also be written as,

$$S_{k,p} \propto \mathrm{Re}\langle O_k^\dagger O_p \rangle_p \quad \text{and} \quad f_k \propto \mathrm{Re}\langle O_k^\dagger \mathcal{L}^{\mathrm{res}} \rangle_p, \tag{10}$$

where $\langle \ldots \rangle_p$ denotes a statistical expectation value of the probability distribution p as in Eq. (9), and we have introduced the following estimator for the Liouvillian,

$$\mathcal{L}^{\mathrm{res}}_{\vec{l_1}, \vec{r_1}; \vec{l_2}, \vec{r_2}} = \sum_{\vec{l_2}, \vec{r_2}} \frac{\mathcal{L}_{\vec{l_1}, \vec{r_1}; \vec{l_2}, \vec{r_2}} \rho_{\vec{l_2}, \vec{r_2}}}{\rho_{\vec{l_1}, \vec{r_1}}} \tag{11}$$

In addition to having a stochastic strategy for solving the variational equations of motion, it is also important to provide an efficient scheme to compute expectation values of physical observables. Consider the expectation value of a generic observable X,

$$\langle X \rangle = \mathrm{Tr}\{X\rho\} = \sum_{\vec{l}, \vec{m}} X_{\vec{l}, \vec{m}} \rho_{\vec{m}, \vec{l}}. \tag{12}$$

Estimates of $\langle X \rangle$ can be obtained in this case as statistical averages over the probability distribution $q(\vec{l}) = \rho_{\vec{l}, \vec{l}}$, such that

$$\langle X \rangle \simeq \langle X^{\mathrm{loc}} \rangle_q, \quad \text{where} \quad X^{\mathrm{loc}}_{\vec{l}, \vec{l}} = \sum_{\vec{m}} \frac{X_{\vec{l}, \vec{m}} \rho_{\vec{m}, \vec{l}}}{\rho_{\vec{l}, \vec{l}}} \tag{13}$$

In all cases of physical relevance, observables X have a sparse representation, and computing the estimator $X^{\mathrm{loc}}_{\vec{l}, \vec{l}}$ can be efficiently realized. Notice that, while possible, sampling over $p(\vec{l}, \vec{m})$ to compute physical expectation values would entail a much less efficient statistical estimator for $\langle X \rangle$. This would further require to stochastically estimate the normalization factor, which is instead automatically taken into account when sampling from $q(\vec{l})$. In this work we use two independent Markov-Chain Monte Carlo schemes to obtain samples both from $p(\vec{l}, \vec{r})$

and from $q(\vec{l})$ at each instant of time, as explained in the Supplementary Material.

Results: To test the accuracy of our method, we consider an anisotropic Heisenberg model with Hamiltonian

$$H = \sum_{j=1}^{N} B\sigma_j^z + \sum_{<j,l>} \sum_{a=x,y,z} J_a \sigma_j^a \sigma_l^a \tag{14}$$

and dissipator

$$\mathcal{D}[\rho] = \frac{\gamma}{2} \sum_{j=1}^{N} \left(2\sigma_j^- \rho \sigma_j^+ - \sigma_j^+ \sigma_j^- \rho - \rho \sigma_j^+ \sigma_j^- \right) \tag{15}$$

We consider two applications. First we compare the time evolution of a density matrix as obtained from Eq. (8) to the exact time evolution of the density matrix for a small size model where the full master equation (1) can be numerically integrated. Then we show that our method correctly finds stationary states for a larger model, that can no longer be fully integrated but where a Matrix Product State (MPS) representation of ρ [42] allows to find the stationary state.

To quantify the accuracy of a time evolution for ρ as obtained from Eq. (8), we consider two quantities. **(i)** The average deviation of the matrix element of ρ from the exact density matrix ρ^e iss given by

$$\delta\rho = \frac{1}{2^{2N}} \sum_{j,l} |\rho_{j,l} - \rho_{j,l}^e|^2 = \frac{||\rho - \rho^e||_2^2}{2^{2N}} \tag{16}$$

where $||x||$ is the 2-norm of a matrix x. **(ii)** To get an accuracy test in terms of physical observables we compare the magnetization for our approximation

$$m_{\mathrm{ml}}^z = \frac{1}{N} \sum_j \langle \sigma_j^z \rangle \tag{17}$$

to the magnetization for an exact solution $m_{\mathrm{ex}}^z = \frac{1}{N} \sum_j \langle \sigma_j^z \rangle_{\mathrm{exact}}$.

Results for a linear chain with $N = 5$ spins and periodic boundary conditions are presented in Fig. 2 and clearly show that the parameterization of the density matrix ρ in terms of the neural network in Fig. 1 provides a very good approximation to the dissipative quantum dynamics of mixed states.

To show that our method correctly finds stationary states for models where the full density matrix can no longer be computed, we test whether $\mathcal{L}\rho = 0$. To this end we compute

$$\delta\mathcal{L} = \langle |\mathcal{L}^{\mathrm{res}}| \rangle_p. \tag{18}$$

Since $\delta\mathcal{L} = \sum_{\vec{l_1}, \vec{r_1}} |\rho_{\vec{l_1}, \vec{r_1}}| |(\mathcal{L}\rho)_{\vec{l_1}, \vec{r_1}}|$ this tests whether all matrix elements $(\mathcal{L}\rho)_{\vec{l_1}, \vec{r_1}}$ vanish. Moreover, the measure $\delta\mathcal{L}$ weights the matrix elements of $\mathcal{L}\rho$ according to the relevance for the state ρ and is very economic to compute.
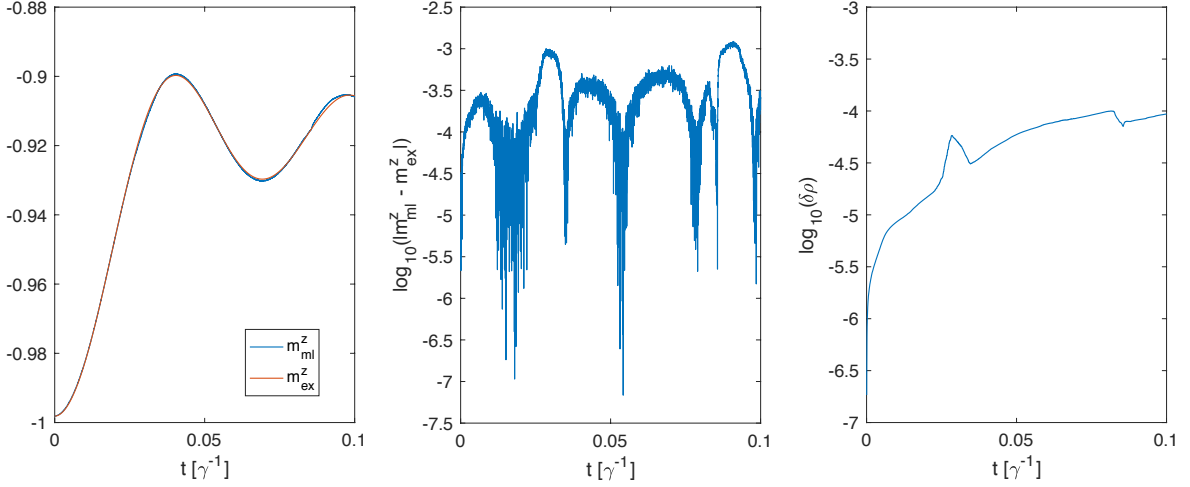
FIG. 2. Results for a chain of 5 spins with periodic boundary conditions and $B = 10\gamma$, $J_x = 20\gamma$, $J_y = 0$ and $J_z = 10\gamma$. $M = \tilde{M} = 20$, the sample size was $N_S = 10^6$ and the time step of 4-th order Runge-Kutta integration was $\delta t = 2 \times 10^{-5}\gamma^{-1}$. **(a)** Average deviation of the matrix element of ρ from the exact density matrix ρ_e as given by Eq. (16). **(b)** Magnetization for the neural network approximation, $\langle\sigma^z\rangle$ (blue) and the exact solution, $\langle\sigma^z\rangle_e$ (orange). **(c)** Log plot of difference between $\langle\sigma^z\rangle$ and $\langle\sigma^z\rangle_e$.

$\delta\mathcal{L}$ can thus be computed as a test for the convergence to the stationary state, even if the properties of the latter are completely unknown. Notice that since $\delta\mathcal{L}$ can be efficiently estimated (as well as other related quantities such as $|\mathcal{L}\rho|^2 \propto \langle|\mathcal{L}^{\mathrm{res}}|^2\rangle_p$, at the same cost of appliying $\mathcal{L}$ once, it is in principle possible to devise an alternative variational optimization scheme that directly minimizes $\delta\mathcal{L}$, if only the stationary state is of interest.

In addition to computing $\delta\mathcal{L}$, we also test whether the magnetization, see Eq. (17) approaches the steady state magnetization $m_{\mathrm{ss}}^z = \lim_{t\to\infty} \frac{1}{N}\sum_j\langle\sigma_j^z\rangle$, which we obtain from an integration with a MPS representation of the density matrix ρ. Results for the approach to the stationary state of a chain with $N = 16$ spins and open boundary conditions are presented in Fig. 3 and show that the stationary state is found with high accuracy. For finding stationary states, we make use of the fact that the parameterization (2) always guarantees a physically valid state. If we do not require to correctly model the dynamics for all times, we can thus choose the integration time step larger and still find convergence to the correct stationary state.

The example presented here, features a moderately correlated state where an MPS representation of the steady state with bond dimension $D = 20$ suffices to compute m_{ss}^z with an accuracy of 10^{-3}. Scenarios with stronger spin-spin interactions would require more variational parameters and larger sample sizes, increasing the numerical effort of the method.

Conclusions: We have introduced a neural-network based approach to numerically modeling the quantum dynamics and stationary quantum states of open or dis-
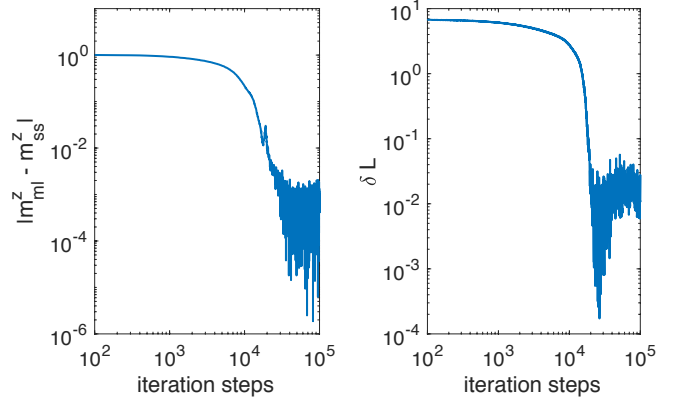

FIG. 3. Results for a chain of 16 spins with open boundary conditions $B = 10\gamma$, $J_x = \gamma$, $J_y = 0$ and $J_z = 0$. $M = \tilde{M} = 12$, the sample size was $N_S = 2 \times 10^5$ and a 2nd order Runge-Kutta integration with adaptive step size was used. **(a)** Difference between m_{ml}^z, see Eq. (17), and $m_{\mathrm{ss}}^z = -15.9286$ as found from an integration with MPS. and **(b)** magnitude of $\mathcal{L}\rho$ as quantified by δL given in Eq. (18)

sipative quantum many-body systems. Our results show that both, the dynamics and stationary states of such systems can be obtained with high accuracy. In this work we have shown one-dimensional systems, in order to provide benchmarks with existing approaches. Several extensions of our approach can be envisaged for future research. From the point of view of applications, the study of two-dimensional lattices does not present conceptual difficulties, and will represent an interesting opportunity for our method. From the methodological point of view,

schemes targeting only the stationary state can also be efficiently implemented, using the same ideas introduced to compute $\delta\mathcal{L}$ in this work.

———————

* present address: Google Inc., Erika-Mann-Str. 33, 80636 München, Germany

[1] Heinz-Peter Breuer, *The Theory of Open Quantum Systems* (Oxford University Press, USA, Oxford, 2007).

[2] Ulrich Schollwöck, "The density-matrix renormalization group in the age of matrix product states," Annals of Physics January 2011 Special Issue, **326**, 96–192 (2011).

[3] F. Verstraete, V. Murg, and J. I. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," Advances in Physics **57**, 143–224 (2008).

[4] Augustine Kshetrimayum, Hendrik Weimer, and Román Orús, "A simple tensor network algorithm for two-dimensional steady states," Nature Communications **8**, 1291 (2017).

[5] David Ceperley and Berni Alder, "Quantum Monte Carlo," Science **231**, 555–560 (1986).

[6] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Reviews of Modern Physics **73**, 33–83 (2001).

[7] Zheng Yan, Lode Pollet, Jie Lou, Xiaoqun Wang, Yan Chen, and Zi Cai, "Interacting lattice systems with quantum dissipation: A quantum Monte Carlo study," Physical Review B **97**, 035148 (2018).

[8] Alexandra Nagy and Vincenzo Savona, "Driven-dissipative quantum Monte Carlo method for open quantum systems," Physical Review A **97**, 052129 (2018).

[9] Giuseppe Carleo and Matthias Troyer, "Solving the quantum many-body problem with artificial neural networks," Science **355**, 602–606 (2017).

[10] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma, "Machine learning topological states," Physical Review B **96**, 195145 (2017).

[11] Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez, and J. Ignacio Cirac, "Neural-Network Quantum States, String-Bond States, and Chiral Topological States," Physical Review X **8**, 011006 (2018).

[12] Raphael Kaubruegger, Lorenzo Pastori, and Jan Carl Budich, "Chiral topological phases from artificial neural networks," Physical Review B **97**, 195136 (2018).

[13] Kenny Choo, Giuseppe Carleo, Nicolas Regnault, and Titus Neupert, "Symmetries and Many-Body Excitations with Neural-Network Quantum States," Physical Review Letters **121**, 167204 (2018).

[14] Markus Schmitt and Markus Heyl, "Quantum dynamics in transverse-field Ising models from classical networks," SciPost Physics **4**, 013 (2018).

[15] Stefanie Czischek, Martin Gärttner, and Thomas Gasenzer, "Quenches near Ising quantum criticality as a challenge for artificial neural networks," Physical Review B

[16] Bjarni Jónsson, Bela Bauer, and Giuseppe Carleo, "Neural-network states for the classical simulation of quantum computing," arXiv:1808.05232 [cond-mat, physics:physics, physics:quant-ph] (2018), arXiv: 1808.05232.

[17] Jing Chen, Song Cheng, Haidong Xie, Lei Wang, and Tao Xiang, "Equivalence of restricted Boltzmann machines and tensor network states," Physical Review B **97**, 085104 (2018).

[18] Lorenzo Pastori, Raphael Kaubruegger, and Jan Carl Budich, "Generalized Transfer Matrix States from Artificial Neural Networks," arXiv:1808.02069 [cond-mat, physics:physics, physics:quant-ph] (2018), arXiv: 1808.02069.

[19] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, "Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems," Physical Review Letters **93**, 207204 (2004).

[20] Michael Zwolak and Guifré Vidal, "Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm," Physical Review Letters **93**, 207205 (2004).

[21] R. Orús and G. Vidal, "Infinite time-evolving block decimation algorithm beyond unitary evolution," Physical Review B **78**, 155117 (2008).

[22] Jian Cui, J. Ignacio Cirac, and Mari Carmen Bañuls, "Variational Matrix Product Operators for the Steady State of Dissipative Quantum Systems," Physical Review Letters **114**, 220601 (2015).

[23] Eduardo Mascarenhas, Hugo Flayac, and Vincenzo Savona, "Matrix-product-operator approach to the nonequilibrium steady state of driven-dissipative quantum arrays," Physical Review A **92**, 022116 (2015).

[24] A.H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, "Positive Tensor Network Approach for Simulating Open Quantum Many-Body Systems," Physical Review Letters **116**, 237201 (2016).

[25] Adil A. Gangat, Te I, and Ying-Jer Kao, "Steady States of Infinite-Size Dissipative Quantum Chains via Imaginary Time Evolution," Physical Review Letters **119**, 010501 (2017).

[26] Daniel Jaschke, Simone Montangero, and Lincoln D. Carr, "One-dimensional many-body entangled open quantum systems with tensor network methods," Quantum Science and Technology **4**, 013001 (2018).

[27] Giacomo Torlai and Roger G. Melko, "Latent Space Purification via Neural Density Operators," Physical Review Letters **120**, 240503 (2018).

[28] Juan Carrasquilla, Giacomo Torlai, Roger G. Melko, and Leandro Aolita, "Reconstructing quantum states with generative models," arXiv:1810.10584 [cond-mat, physics:quant-ph] (2018), arXiv: 1810.10584.

[29] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller, "Quantum states and phases in driven open quantum systems with cold atoms," Nature Physics **4**, 878 EP – (2008).

[30] Julio T. Barreiro, Markus Müller, Philipp Schindler, Daniel Nigg, Thomas Monz, Michael Chwalla, Markus Hennrich, Christian F. Roos, Peter Zoller, and Rainer Blatt, "An open-system quantum simulator with trapped ions," Nature **470**, 486 EP – (2011).

[31] Mattias Fitzpatrick, Neereja M. Sundaresan, Andy C. Y.

Li, Jens Koch, and Andrew A. Houck, "Observation of a dissipative phase transition in a one-dimensional circuit qed lattice," Phys. Rev. X **7**, 011016 (2017).

[32] Michele C. Collodo, Anton Potočnik, Simone Gasparinetti, Jean-Claude Besse, Marek Pechal, Mahdi Sameti, Michael J. Hartmann, Andreas Wallraff, and Christopher Eichler, "Observation of the Crossover from Photon Ordering to Delocalization in Tunably Coupled Resonators," arXiv e-prints , arXiv:1808.00889 (2018), arXiv:1808.00889 [quant-ph].

[33] Ruichao Ma, Brendan Saxberg, Clai Owens, Nelson Leung, Yao Lu, Jonathan Simon, and David I. Schuster, "A Dissipatively Stabilized Mott Insulator of Photons," arXiv e-prints , arXiv:1807.11342 (2018), arXiv:1807.11342 [cond-mat.quant-gas].

[34] Jamir Marino and Sebastian Diehl, "Quantum dynamical field theory for nonequilibrium phase transitions in driven open systems," Physical Review B **94**, 085150 (2016).

[35] Wim Casteels, Ryan M. Wilson, and Michiel Wouters, "Gutzwiller Monte Carlo approach for a critical dissipative spin model," Physical Review A **97**, 062107 (2018).

[36] Jiasen Jin, Alberto Biella, Oscar Viyuela, Leonardo Mazza, Jonathan Keeling, Rosario Fazio, and Davide Rossini, "Cluster Mean-Field Approach to the Steady-State Phase Diagram of Dissipative Spin Systems," Physical Review X **6**, 031011 (2016).

[37] Giuseppe Carleo, Federico Becca, Marco Schiro, and Michele Fabrizio, "Localization and Glassy Dynamics Of Many-Body Quantum Systems," Scientific Reports **2**, 243 (2012).

[38] John Preskill, "Quantum Computing in the NISQ era and beyond," Quantum **2**, 79 (2018).

[39] E. M. Kessler, G. Giedke, A. Imamoglu, S. F. Yelin, M. D. Lukin, and J. I. Cirac, "Dissipative phase transition in a central spin system," Physical Review A **86**, 012116 (2012).

[40] Michael J. Hartmann, "Quantum simulation with interacting photons," Journal of Optics **18**, 104005 (2016).

[41] Sandro Sorella, Michele Casula, and Dario Rocca, "Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods," The Journal of Chemical Physics **127**, 014105 (2007).

[42] Michael J. Hartmann, "Polariton Crystallization in Driven Arrays of Lossy Nonlinear Resonators," Physical Review Letters **104**, 113601 (2010).

SUPPLEMENTAL MATERIAL

Proof of Minimum

$\mathcal{S}$ and $\vec{f}$ can be written as,

$$S_{k,p} = \vec{\rho}^\dagger O_k^\dagger O_p \vec{\rho} + \vec{\rho}^\dagger O_p^\dagger O_k \vec{\rho} = \frac{\partial \vec{\rho}^\dagger}{\partial \alpha_k} \frac{\partial \vec{\rho}}{\partial \alpha_p} + \frac{\partial \vec{\rho}^\dagger}{\partial \alpha_p} \frac{\partial \vec{\rho}}{\partial \alpha_k}$$

$$f_k = \vec{\rho}^\dagger O_k^\dagger \mathcal{L} \vec{\rho} + \vec{\rho}^\dagger \mathcal{L}^\dagger O_k \vec{\rho} = \frac{\partial \vec{\rho}^\dagger}{\partial \alpha_k} \mathcal{L} \vec{\rho} + \vec{\rho}^\dagger \mathcal{L}^\dagger \frac{\partial \vec{\rho}}{\partial \alpha_k}$$

The $\dot{\alpha}_k$ are real and we have for any vector with real elements v_k,

$$\sum_{k,p} v_k S_{k,p} v_p = \sum_{k,p} \left(v_k \frac{\partial \vec{\rho}^\dagger}{\partial \alpha_k} \frac{\partial \vec{\rho}}{\partial \alpha_p} v_p + v_p \frac{\partial \vec{\rho}^\dagger}{\partial \alpha_p} \frac{\partial \vec{\rho}}{\partial \alpha_k} v_k \right) = 2 \vec{\nu}^\dagger \vec{\nu} \geq 0 \tag{19}$$

where $\vec{\nu} = \sum_k \frac{\partial \vec{\rho}}{\partial \alpha_k} v_k$. Therefore, the matrix S is positive semidefinite and the solution to Eq. (5) is indeed the minimum of δ.

Fubini-Study norm

Alternatively to the 2-norm, the Fubini-Study norm can be employed to derive the approximation. For this approach one minimizes

$$\gamma(\vec{\sigma}, \vec{\mu}) = \arccos \sqrt{\frac{|\vec{\sigma}^\dagger \cdot \vec{\mu}|^2}{(\vec{\sigma}^\dagger \cdot \vec{\sigma})(\vec{\mu}^\dagger \cdot \vec{\mu})}} \tag{20}$$

for $\sigma = \sum_k \dot{\alpha}_k O_k \vec{\rho}$ and $\mu = \mathcal{L} \vec{\rho}$. This leads to the same system of ODEs as given in Eq. (7) of the main text, but where S and f read,

$$S_{k,k'} = \vec{\rho}^\dagger O_k^\dagger O_{k'} \vec{\rho} + \vec{\rho}^\dagger O_{k'}^\dagger O_k \vec{\rho} - \vec{\rho}^\dagger O_k^\dagger \vec{\rho} \vec{\rho}^\dagger O_{k'} \vec{\rho} - \vec{\rho}^\dagger O_{k'}^\dagger \vec{\rho} \vec{\rho}^\dagger O_k \vec{\rho} \tag{21}$$

$$f_k = \vec{\rho}^\dagger O_k^\dagger \mathcal{L} \vec{\rho} + \vec{\rho}^\dagger O_k \mathcal{L}^\dagger \vec{\rho} - \vec{\rho}^\dagger O_k^\dagger \vec{\rho} \vec{\rho}^\dagger \mathcal{L} \vec{\rho} - \vec{\rho}^\dagger O_k \vec{\rho} \vec{\rho}^\dagger \mathcal{L}^\dagger \vec{\rho} \tag{22}$$

Logarithmic derivatives

The logarithmic derivatives read,

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Re}[a_n]} = l_n + r_n \tag{23}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Im}[a_n]} = i(l_n - r_n) \tag{24}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Re}[b_k]} = \xi_{b,W}(k,\vec{l}) + [\xi_{b,W}(k,\vec{r})]^* \tag{25}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Im}[b_k]} = i\xi_{b,W}(k,\vec{l}) - i[\xi_{b,W}(k,\vec{r})]^* \tag{26}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Re}[W_{k,n}]} = l_n\xi_{b,W}(k,\vec{l}) + r_n[\xi_{b,W}(k,\vec{r})]^* \tag{27}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Im}[W_{k,n}]} = il_n\xi_{b,W}(k,\vec{l}) - ir_n[\xi_{b,W}(k,\vec{r})]^* \tag{28}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Re}[c_p]} = 2\,\xi_{c,U}(p,\vec{l},\vec{r}) \tag{29}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Re}[U_{p,n}]} = (l_n + r_n)\,\xi_{c,U}(p,\vec{l},\vec{r}) \tag{30}$$

$$\frac{\partial \ln(\rho_{\vec{l},\vec{r}})}{\partial \text{Im}[U_{p,n}]} = i(l_n - r_n)\,\xi_{c,U}(p,\vec{l},\vec{r}) \tag{31}$$

where

$$\xi_{b,W}(k,\vec{n}) = \tanh\left[b_k + \sum_{j=1}^{N} W_{k,j}n_j\right] \tag{32}$$

$$\xi_{c,U}(p,\vec{l},\vec{r}) = \tanh\left[c_p + c_p^* + \sum_{j=1}^{N} U_{p,j}l_j + \sum_{j=1}^{N} U_{p,j}^* r_j\right] \tag{33}$$

Since they all involve sums of exponentially many terms, we will compute all expectation values stochastically via Monte Carlo sampling.

Sampling for expectation values

We here provide the arguments why one needs to sample from the distribution formed by the diagonal elements of the density matrix for computing expectation values of observables. Sampling over $p(\vec{l},\vec{m})$ would not work for computing such expectation values since this would require finding the normalization via

$$\mathcal{N} = \sum_{\vec{l}} \rho_{\vec{l},\vec{l}} = \sum_{\vec{l},\vec{m}} p(\vec{l},\vec{m}) \frac{\rho_{\vec{m},\vec{m}}}{|\rho_{\vec{l},\vec{m}}|^2} \approx \frac{1}{N_S} \sum_{\vec{l},\vec{r}\in\mathcal{S}} \frac{\rho_{\vec{m},\vec{m}}}{|\rho_{\vec{l},\vec{m}}|^2} \tag{34}$$

which may be flawed by the statistical behavior. Indeed if we look at the variance,

$$\Delta\mathcal{N}^2 = \sum_{\vec{l},\vec{m}} p(\vec{l},\vec{m}) \left(\frac{\rho_{\vec{m},\vec{m}}}{|\rho_{\vec{l},\vec{m}}|^2}\right)^2 - \mathcal{N}^2 = \sum_{\vec{l},\vec{m}} \frac{\rho_{\vec{m},\vec{m}}^2}{|\rho_{\vec{l},\vec{m}}|^2} - \mathcal{N}^2 \tag{35}$$

we find that it could diverge for $|\rho_{\vec{l},\vec{m}}| \to 0$. This can become problematic due to the Schwarz inequality $|\rho_{\vec{l},\vec{m}}|^2 \leq \rho_{\vec{l},\vec{l}}\rho_{\vec{m},\vec{m}}$.

Local estimator of the Liouvillian

For computing the matrices S and f for the stochastic reconfiguration approach, we need to calculate the so called local Liouvillian, see Eq. (10) of the main text. We first consider the term $-i[H, \rho]$, for which the local Liouvillian reads,

$$\frac{-i\langle\vec{l}|[H,\rho]|\vec{r}\rangle}{\rho_{\vec{l},\vec{r}}} = -i\sum_{\vec{m}} H_{\vec{l},\vec{m}} \frac{\rho_{\vec{m},\vec{r}}}{\rho_{\vec{l},\vec{r}}} + i\sum_{\vec{m}} \frac{\rho_{\vec{l},\vec{m}}}{\rho_{\vec{l},\vec{r}}} H_{\vec{m},\vec{r}}$$

which can be computed efficiently for k-local Hamiltonians, with the same complexity of the local energy in standard variational calculations for the ground-state. For the Hamiltonian in Eq. (14), we have in one dimension

$$H_{\vec{m},\vec{n}} = \sum_{j=1}^{N} \langle\vec{m}|H_j|\vec{n}\rangle$$

$$\langle\vec{m}|H_j|\vec{n}\rangle = B\langle\vec{m}|\sigma_j^z|\vec{n}\rangle + J_x\langle\vec{m}|\sigma_j^x\sigma_{j+1}^x|\vec{n}\rangle + J_y\langle\vec{m}|\sigma_j^y\sigma_{j+1}^y|\vec{n}\rangle + J_z\langle\vec{m}|\sigma_j^z\sigma_{j+1}^z|\vec{n}\rangle$$

For the dissipator, $\mathcal{D}[\rho]$, we get,

$$\mathcal{D}[\rho] = \sum_{j=1}^{N} \mathcal{D}_j[\rho]$$

$$\frac{\langle\vec{l}|\mathcal{D}_j[\rho]|\vec{r}\rangle}{\rho_{\vec{l},\vec{r}}} = \gamma\sum_{\vec{l}',\vec{r}'} \langle\vec{l}|\sigma_j^-|\vec{l}'\rangle\langle\vec{r}'|\sigma_j^+|\vec{r}\rangle \frac{\rho_{\vec{l}',\vec{r}'}}{\rho_{\vec{l},\vec{r}}} - \frac{\gamma}{2}\sum_{\vec{m}} \langle\vec{l}|\sigma_j^+\sigma_j^-|\vec{m}\rangle \frac{\rho_{\vec{m},\vec{r}}}{\rho_{\vec{l},\vec{r}}} - \frac{\gamma}{2}\sum_{\vec{m}} \langle\vec{m}|\sigma_j^+\sigma_j^-|\vec{r}\rangle \frac{\rho_{\vec{l},\vec{m}}}{\rho_{\vec{l},\vec{r}}}.$$

We use here a notation, where $n_j = \pm 1$. Thus we get,

$$\langle\vec{m}|\sigma_j^z|\vec{n}\rangle = n_j\,\delta_{\vec{m},\vec{n}}$$

$$\langle\vec{m}|\sigma_j^x\sigma_{j+1}^x|\vec{n}\rangle = \delta_{n_j,-m_j}\delta_{n_{j+1},-m_{j+1}} \prod_{i\neq j,j+1} \delta_{m_i,n_i}$$

$$\langle\vec{m}|\sigma_j^y\sigma_{j+1}^y|\vec{n}\rangle = \delta_{n_j,-m_j}\delta_{n_{j+1},-m_{j+1}}(\delta_{n_j,-n_{j+1}} - \delta_{n_j,n_{j+1}}) \prod_{i\neq j,j+1} \delta_{m_i,n_i}$$

$$\langle\vec{m}|\sigma_j^z\sigma_{j+1}^z|\vec{n}\rangle = n_jn_{j+1}\,\delta_{\vec{m},\vec{n}}$$

$$\langle\vec{m}|\sigma_j^+\sigma_j^-|\vec{n}\rangle = \delta_{n_j,1}\,\delta_{\vec{m},\vec{n}}$$

$$\langle\vec{m}|\sigma_j^-|\vec{n}\rangle = \delta_{n_j,1}\delta_{m_j,-1} \prod_{i\neq j} \delta_{m_i,n_i}$$

$$\langle\vec{m}|\sigma_j^+|\vec{n}\rangle = \delta_{n_j,-1}\delta_{m_j,1} \prod_{i\neq j} \delta_{m_i,n_i}$$

where the $\delta_{j,l}$ are Kronecker deltas. Using the notation $\rho(l_1,\ldots,l_n;r_1,\ldots,r_n) = \rho_{l_1,\ldots,l_n;r_1,\ldots,r_n}$ with $n = 2^N$, we get,

$$\frac{\langle\vec{l}|\mathcal{D}_j[\rho]|\vec{r}\rangle}{\rho_{\vec{l},\vec{r}}} = \gamma\delta_{l_j,-1}\delta_{r_j,-1} \frac{\rho(\ldots,l_{j-1},1,l_{j+1},\ldots;\ldots,r_{j-1},1,r_{j+1},\ldots)}{\rho(\ldots,l_{j-1},-1,l_{j+1},\ldots;\ldots,r_{j-1},-1,r_{j+1},\ldots)} - \frac{\gamma}{2}(\delta_{l_j,1} + \delta_{r_j,1})$$

and

$$\frac{-i\langle\vec{l}|[H_j,\rho]|\vec{r}\rangle}{\rho_{\vec{l},\vec{r}}} = -iB(l_j - r_j) - iJ_z(l_jl_{j+1} - r_jr_{j+1})$$

$$- iJ_x\left[\frac{\rho(\ldots,-l_j,-l_{j+1},\ldots;\ldots)}{\rho(\ldots,l_j,l_{j+1},\ldots;\ldots)} - \frac{\rho(\ldots;\ldots,-r_j,-r_{j+1},\ldots)}{\rho(\ldots;\ldots,r_j,r_{j+1},\ldots)}\right]$$

$$- iJ_y\left[(\delta_{l_j,-l_{j+1}} - \delta_{l_j,l_{j+1}})\frac{\rho(\ldots,-l_j,-l_{j+1},\ldots;\ldots)}{\rho(\ldots,l_j,l_{j+1},\ldots;\ldots)} - (\delta_{r_j,-r_{j+1}} - \delta_{r_j,r_{j+1}})\frac{\rho(\ldots;\ldots,-r_j,-r_{j+1},\ldots)}{\rho(\ldots;\ldots,r_j,r_{j+1},\ldots)}\right]$$

Moves of the Metropolis Sampling

For sampling from the distribution $p(\vec{l}, \vec{r})$, we considered four types of moves,

1. one index, either left or right (l_j or r_j) is flipped.

2. the left and right indices l_j and r_j, corresponding to one site j are both flipped.

3. neighboring left indices l_j and l_{j+1} or right indices r_j and r_{j+1} are flipped.

4. a new configuration is drawn from a uniform distrbution.

Whereas the moves 1-3 occur with the same probability, the likelihood for move 4 was chosen to be 100 times smaller. In turn for the sampling from the distribution $q(\vec{l}, \vec{l})$, we considered three types of moves,

1. one index l_j is flipped.

2. neighboring indices l_j and l_{j+1} are flipped.

3. a new configuration is drawn from a uniform distrbution.

Here again, the last move was chosen to occur 100 times less often than the other two.