

## Autoregressive Neural Network for Simulating Open Quantum Systems via a Probabilistic Formulation

Di Luo,<sup>1,2,\*</sup> Zhuo Chen<sup>1,\*</sup>, Juan Carrasquilla,<sup>3,4</sup> and Bryan K. Clark<sup>1,2,5</sup>

<sup>1</sup>*Department of Physics, University of Illinois at Urbana-Champaign, Illinois 61801, USA*

<sup>2</sup>*IQUIST and Institute for Condensed Matter Theory, University of Illinois at Urbana-Champaign, Illinois 61801, USA*

<sup>3</sup>*Vector Institute for Artificial Intelligence, MaRS Centre, Toronto, Ontario, Canada*

<sup>4</sup>*Department of Physics and Astronomy, University of Waterloo, Ontario N2L 3G1, Canada*

<sup>5</sup>*NCSA Center for Artificial Intelligence Innovation, University of Illinois at Urbana-Champaign, Illinois 61801, USA*



(Received 7 October 2020; revised 19 May 2021; accepted 5 January 2022; published 28 February 2022)

The theory of open quantum systems lays the foundation for a substantial part of modern research in quantum science and engineering. Rooted in the dimensionality of their extended Hilbert spaces, the high computational complexity of simulating open quantum systems calls for the development of strategies to approximate their dynamics. In this Letter, we present an approach for tackling open quantum system dynamics. Using an exact probabilistic formulation of quantum physics based on positive operator-valued measure, we compactly represent quantum states with autoregressive neural networks; such networks bring significant algorithmic flexibility due to efficient exact sampling and tractable density. We further introduce the concept of string states to partially restore the symmetry of the autoregressive neural network and improve the description of local correlations. Efficient algorithms have been developed to simulate the dynamics of the Liouvillian superoperator using a forward-backward trapezoid method and find the steady state via a variational formulation. Our approach is benchmarked on prototypical one-dimensional and two-dimensional systems, finding results which closely track the exact solution and achieve higher accuracy than alternative approaches based on using Markov chain Monte Carlo method to sample restricted Boltzmann machines. Our Letter provides general methods for understanding quantum dynamics in various contexts, as well as techniques for solving high-dimensional probabilistic differential equations in classical setups.

DOI: [10.1103/PhysRevLett.128.090501](https://doi.org/10.1103/PhysRevLett.128.090501)

*Introduction.*—While the Universe itself is a closed quantum system, all other systems within the Universe are open quantum systems coupled to the environment around them. Open quantum systems (OQS) play a crucial role in fundamental quantum science, quantum control, and quantum engineering [1,2]. In recent years, there has been a significant interest both theoretically and experimentally in better understanding open quantum systems [3–38]. In the field of quantum engineering, coupling to the environment generates decoherence driving the destruction of entanglement within quantum devices. Quantum computers rely on the qubit-environment coupling to apply quantum gates as well as try to minimize unwanted coupling to mitigate errors on the qubits [39].

Unlike closed quantum states which can be represented by a wave function, the density matrix  $\rho$  becomes the core object of study in open quantum systems. A typical model of an OQS evolves the density matrix under both the Hamiltonian  $H$  as well as a series of dissipative operators which transfer energy and information out to a featureless bath leading to the Lindblad equation,

$$\dot{\rho} = \mathcal{L}\rho \equiv -i[H, \rho] + \sum_k \frac{\gamma_k}{2} (2\Gamma_k \rho \Gamma_k^\dagger - \{\rho, \Gamma_k^\dagger \Gamma_k\}), \quad (1)$$

where  $\gamma_k$  are the dissipation rates associated with jump operators  $\Gamma_k$ . Although there is hope that quantum algorithms [40–44] may eventually overcome the simulation bottlenecks in OQS, a direct solution to the Lindblad equation is difficult because the Hilbert space grows exponentially with the number of particles, making classical simulations largely intractable. To deal with this curse of dimensionality, OQS have historically been studied with renormalization group approaches [11–13], mean field methods [9,10,45]; or simulated with tensor networks [5–8,18,46,47] which compress the density matrix. Unfortunately, while tensor networks have proved fruitful in one dimension, their use for OQS in higher dimensions has been severely limited. Recently, inspired by the advances in the description of many-body systems in terms of neural network wave functions [48–57], ideas from machine learning have been applied to OQS studying real-time dynamics in one dimension (1D), steady states in one and two dimensions (2D) [58–61] and determining the Liouvillian gap [62] by representing the density matrix as a restricted Boltzmann machine (RBM) [63].

Here, we outline an alternative approach to using machine learning ideas to simulate the Lindblad equation. Many machine learning architectures and generative

models (such as the RBM) have fundamentally been designed to represent probability distributions (e.g., probability distributions over images on the internet) making them inadequate to store quantum states, which are complex valued in general. To overcome this, novel approaches have been devised such as using complex weights within RBM; despite these innovative ideas, effectively representing states with signs have been a key bottleneck in this field [64–66].

This motivation has inspired us to utilize the recent developments in the probabilistic formulation of quantum mechanics [67–70] to simulate the Lindblad equation. In this formulation the state is mapped to a probability distribution which we represent compactly using the Transformer [71]—a machine learning architecture from which one can efficiently sample the probability distribution exactly. Using this, we develop efficient algorithms to both update the state of the Transformer under dynamic evolution as well as find the Transformer which represents the steady state of the Lindblad equation. To perform the dynamic evolution, we combine the second-order forward-backward trapezoid method [72] with stochastic optimization on the Transformer. Since the Transformer does not naively preserve the symmetry of the true dynamic (or fixed-point) state, we further improve upon our results by developing an additional ansatz—string states—which explicitly restores some of these symmetries. We proceed to benchmark this work on a series of one- and two-dimensional systems.

*Lindblad equation as a probability equation.*—The general objective of this Letter is to develop an approach to solve for the dynamics and fixed point of the density matrix  $\rho$  in the Lindblad equation [Eq. (1)]. We test this approach on two models—the transverse-field Ising model (TFIM), where  $H = (V/4) \sum_{(i,j)} \sigma_i^{(z)} \sigma_j^{(z)} + (g/2) \sum_k \sigma_k^{(x)}$ , and the Heisenberg model, where  $H = \sum_{(i,j)} \sum_{w=x,y,z} J_w \sigma_i^{(w)} \sigma_j^{(w)} + B \sum_k \sigma_k^{(z)}$ . In both cases,  $\Gamma_k = \sigma_k^{(-)} = \frac{1}{2}(\sigma_k^{(x)} - i\sigma_k^{(y)})$ . We are interested in the expectation values of local observables given by the Pauli matrices averaged over all qubits, i.e., for a system with  $n$  qubits, we consider  $\langle \sigma_w \rangle = (1/n) \sum_i \langle \sigma_i^{(w)} \rangle$  for  $w = x, y$ , and  $z$ . Typically, the density matrix  $\rho$  is represented (explicitly or implicitly) in an orthogonal basis. In this Letter, we instead represent  $\rho$  in the positive operator-valued measure (POVM) formalism. Given an informationally complete POVM (IC POVM), a density matrix  $\rho$  of a spin-1/2 system can be uniquely mapped to a probability distribution  $p(\mathbf{a})$ , where  $\mathbf{a}$  spans over all  $4^n$  measurement outcomes in the POVM basis. An IC POVM is defined by a collection of positive semidefinite operators  $\{M_{(a)}\}$  called the frame, which specifies the probability distribution  $p(\mathbf{a}) = \text{Tr}(\rho M_{(a)})$ . The inverse transformation is given by  $\rho = \sum_b p(\mathbf{b}) N^{(b)}$ , where the dual frame  $\{N^{(b)}\}$

can be computed from the frame as  $N^{(b)} = \sum_a M_{(a)} T_{ab}^{-1}$ . The elements of the overlap matrix  $T$  are given by  $T_{ab} = \text{Tr}(M_{(a)} M_{(b)})$ , and  $T_{ab}^{-1}$  represent the elements of the inverse overlap matrix  $T^{-1}$ . Thus, we can re-express the Lindblad equation as

$$\dot{p}(\mathbf{a}) = \sum_b p(\mathbf{b}) L_a^b = \sum_b p(\mathbf{b}) (A_a^b + B_a^b), \quad (2)$$

with

$$A_a^b = -i \text{Tr}(H[N^{(b)}, M_{(a)}]);$$

$$B_a^b = \sum_k \frac{\gamma_k}{2} \text{Tr}(2\Gamma_k N^{(b)} \Gamma_k^\dagger M_{(a)} - \Gamma_k^\dagger \Gamma_k \{N^{(b)}, M_{(a)}\}). \quad (3)$$

We work with an IC POVM where the frame and dual frame are constructed from local frames acting on single spins as  $\{M_{(a)}\} = \{M_{(a_1)} \otimes M_{(a_2)} \otimes M_{(a_3)} \otimes \dots\}$  and  $\{N^{(b)}\} = \{N^{(b_1)} \otimes N^{(b_2)} \otimes N^{(b_3)} \otimes \dots\}$  with four outcomes per spin  $a_i$ . This allows us to write  $p(\mathbf{a}) = p(a_1, a_2, a_3, \dots)$ . The expectation value of observables are given by  $\langle O \rangle = \sum_b p(\mathbf{b}) \text{Tr}(O N^b) \approx (1/N_s) \sum_{b \sim p}^{N_s} \text{Tr}(O N^b)$ , where  $N_s$  is the number of samples  $\mathbf{b}$  drawn from the distribution  $p(\mathbf{b})$  used to estimate  $\langle O \rangle$ . We emphasize that a complete specification of the probability distribution  $p(\mathbf{b})$  requires  $4^n$  probability values for an  $n$ -site system.

*Autoregressive models and string states.*—We have chosen to model the probability distribution in a compact way with an autoregressive neural network where the probability of a given configuration  $\mathbf{a}$  is expressed through its conditional probabilities  $p_\theta(\mathbf{a}) = \prod_k p_\theta(a_k | a_1, a_2, \dots, a_{k-1})$ . This representation allows for exact sampling of a configuration from the space of probability distributions without invoking Markov chain Monte Carlo techniques. Modern incarnations of autoregressive models include, among others, recurrent neural networks (RNN) [73,74], pixel convolutional neural networks (PixelCNN) [75], Transformers [71]. Recent work has effectively applied these models to quantum systems [53,54,68,69,76]. Here, we use an autoregressive Transformer, which follows the same architecture as the model in [69]. The Transformer consists of two hyperparameters: the number of transformer layers stacked on each other  $n_l$  and the hidden dimension  $n_d$ , which we adjusted for different tests.

Since our Transformer gives “ordered” measurement outcomes, when we simulate two-dimensional systems we need to choose a linear ordering of our two-dimensional sites (i.e., a string of sites). We consider two different single-string orderings [string 0 and string 1 from Fig. 1(a)]. These strings explicitly break a symmetry of our system which then would need to be restored (to the degree to which the model has the variational freedom to do so) by the Transformer itself. We can partially (or completely) restore this symmetry explicitly by choosing our ansatz to be a

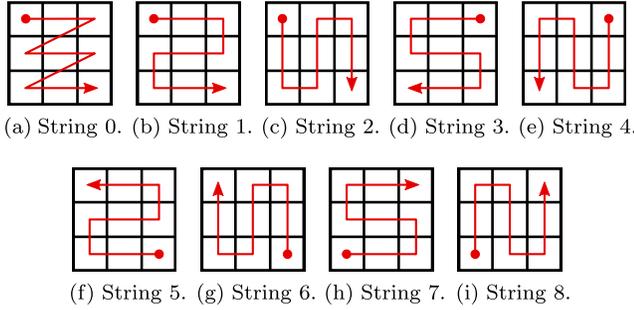


FIG. 1. Strings used for mapping 1D Transformer to 2D quantum systems. String 0 is the default mapping (which we refer to as no strings). We always refer to the first (in order)  $n$  strings (excluding string 0) when we say we used  $n$  strings.

mixture of distributions defined over multiple different symmetry-related strings—i.e.,  $p_\theta(\mathbf{a}) = \sum_S p_\theta(\mathbf{a}|\mathcal{S})p(\mathcal{S})$ , where  $p(\mathcal{S}) = 1/N_{\text{string}}$  for a total number of  $N_{\text{string}}$  strings; we call this refined ansatz a string state. This linear combination of the Transformer probabilities can be interpreted as a mixture model [77] and bears some resemblance to string bond states [78]. Restoring symmetries explicitly has proved useful in variational calculations of quantum states [53,64,79–81]. Given a set of strings and a configuration  $\mathbf{a}$  we can compute  $p(\mathbf{a})$  explicitly. Sampling an  $\mathbf{a}$  from  $p_\theta(\mathbf{a})$  is also straightforward because of linearity and the fact that each term in our average is positive. To do so, we first sample an ordered  $\{a_1, a_2, \dots, a_{k-1}\}$  from the Transformer and then randomly choose a string to map these ordered values to get the final configuration. Here, we test a subset of strings 1– $k$  for different  $k$  [see Figs. 1(b)–1(i)].

*Optimization and results.*—Equation (2) gives a prescription for applying time evolution to the density matrix by time evolving the POVM probability distribution. To solve for the time-evolved distribution, we discretize time and use a second-order forward-backward trapezoid method [72]. We designed the following objective function

$$\mathcal{C} = \frac{1}{N_s} \sum_{\mathbf{a} \sim p_{\theta(t+2\tau)}}^{N_s} \frac{1}{p_{\theta(t+2\tau)}(\mathbf{a})} \times \left| \sum_b [p_{\theta(t+2\tau)}(\mathbf{b})(\delta_a^b - \tau L_a^b) - p_{\theta(t)}(\mathbf{b})(\delta_a^b + \tau L_a^b)] \right|, \quad (4)$$

where  $N_s$  is the number of samples,  $\delta_a^b$  is the Kronecker delta function, the sum over  $\mathbf{a}$  is sampled stochastically from  $p_{\theta(t+2\tau)}$ , the sum over  $\mathbf{b}$  can be evaluated efficiently as explained in Supplemental Material [82] Sec. IX and the gradient of the objective function  $\mathcal{C}$  with respect to the parameters in  $p_{\theta(t+2\tau)}(\mathbf{b})$  is computed using PyTorch’s [83] automatic differentiation. To optimize the objective function we use Adam [84]. In the limit where  $\mathcal{C}$  is zero, we get exact time evolution up to the discretization error induced

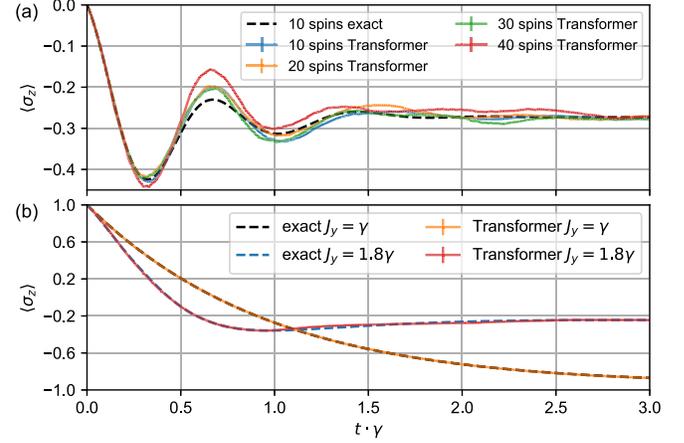


FIG. 2. The expectation value  $\langle \sigma_z \rangle$  as a function of time (a) for the 1D Heisenberg model with  $B = \gamma, J_x = 2\gamma, J_y = 0$ , and  $J_z = \gamma$  using a time step  $\tau = 0.005\gamma^{-1}$ . The initial state is the product state  $\prod_{i=1}^N |\leftarrow\rangle$  ( $\langle \sigma_y \rangle = -1$ ). (b) for the  $3 \times 3$  Heisenberg model with  $B = 0, J_x = 0.9\gamma, J_y = 1.0\gamma, 1.8\gamma$ , and  $J_z = \gamma$  using a time step  $\tau = 0.008\gamma^{-1}$ . The initial state is the product state  $\prod_{i=1}^N |\uparrow\rangle$  ( $\langle \sigma_z \rangle = 1$ ). Both models use periodic boundary conditions. Exact curves are produced using QuTip [87,88]. The Transformer has one encoder layer and 32 hidden dimensions, and is trained using a forward-backward trapezoid method with a sample size  $N_s = 12000$ .

by the trapezoid rule. Typically, it will be impossible for the Transformer to exactly represent the time-evolved state; instead by minimizing  $\mathcal{C}$  the optimization continuously projects onto a nearby state in the manifold of distributions represented by our Transformer. This can be viewed as a higher order generalization of imaginary-time supervised wave function optimization [85] and the method in Ref. [86] but here applied instead to a probability distribution. The dominant source of error in performing our dynamics comes from the limited set of states that the Transformer can represent. Additionally, it is possible that even within this manifold of states, one may not reach the optimal value if there are optimization issues such as local minima. Over multiple time steps, errors will naturally accumulate due to the unitary dynamics of the system and be suppressed by the dissipative operators which should drive all dynamics to a fixed point. We test this dynamic evolution on the 1D and a 2D Heisenberg model (see Fig. 2) using the tetrahedral POVM basis (see Supplemental Material [82] Sec. II) where we find that the dynamics matches closely to the exact result. We capture both the qualitative behavior (i.e., the peaks and oscillation of the observables) as well as their quantitative values. The values are especially accurate in both the limit of small and large time. In our results, we have simulated one-dimensional chains up to  $N = 40$  and two-dimensional chains for  $3 \times 3$  lattices.

One approach to finding the fixed point of the Liouvillian superoperator  $\mathcal{L}$  is through a sufficiently long

time evolution (for an example, see the large time limit of Fig. 2). Interestingly, our approximate time evolution fluctuates around a fixed value of the observable, though it may not reach a true fixed point [i.e.,  $p_\theta(t + 2\tau) = p_\theta(t)$ ] even in the limit of small  $\tau$  (see Supplemental Material [82] Sec. VII).

Alternatively, we can search for the fixed point by direct minimization of the  $L_1$  norm of  $\dot{p}_\theta$  giving

$$\|\dot{p}_\theta\|_1 = \sum_a \left| \sum_b p_\theta(\mathbf{b}) L_a^b \right| \approx \frac{1}{N_s} \sum_{\mathbf{a} \sim p_\theta} \frac{|\sum_b p_\theta(\mathbf{b}) L_a^b|}{p_\theta(\mathbf{a})}, \quad (5)$$

where the second line offers a stochastic approach to evaluate the  $\|\dot{p}_\theta\|_1$  by sampling  $\mathbf{a}$  from  $p_\theta(\mathbf{a})$ . The gradient in Eq. (5) is taken with respect to the parameters in  $p_\theta(\mathbf{b})$  using PyTorch’s [83] automatic differentiation. Notice that because the gradients of Eq. (4) and (5) (see Supplemental Material [82] Sec. VII) are different (except in the limit where the manifold of states representable by the Transformer span the full space), they will converge to different answers.

In Fig. 3, we consider the one-dimensional TFIM, with the 4-Pauli POVM basis (see Supplemental Material [82] Sec. II), and compute the expectation value of all three

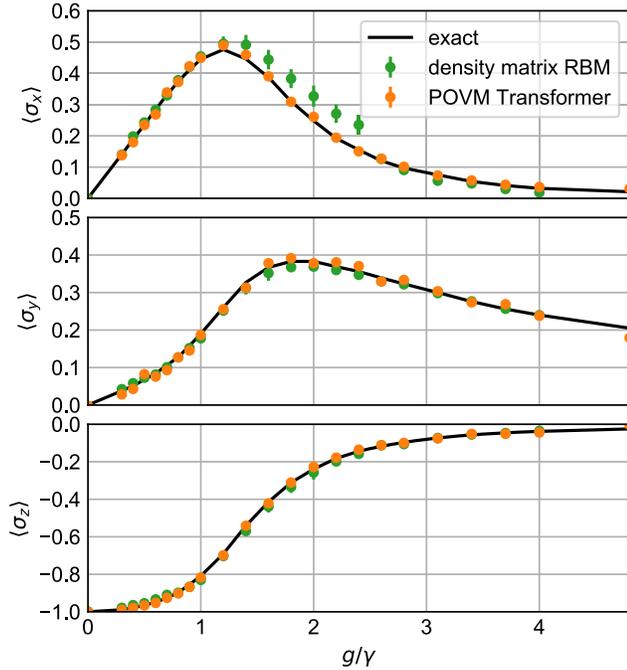


FIG. 3. Variational steady-state solution for a 16-site TFIM chain with periodic boundary condition and  $V = 2\gamma$  (orange dots). The initial state is the product state  $\prod_{i=1}^N |\uparrow\rangle$  ( $\langle \sigma_z \rangle = 1$ ). The Transformer has one encoder layer and 32 hidden dimensions, and is trained using Adam [84] in 500 iterations with  $N_s = 12000$ . Green points are the fixed point solution representing the density matrix as an RBM; both the exact curve (black line) and density matrix results are digitized from Ref. [58].

Pauli matrices at various values of  $\gamma$ . We find strong agreement with the exact method. In addition, we find that this approach performs particularly well in the regime of  $1 < g/\gamma < 2.5$  which has proven particularly challenging for the RBM method [58]. We can further improve the performance by averaging over multiple simulations (see Supplemental Material [82] Sec. V). In Fig. 4, we consider optimizing a  $3 \times 3$  Heisenberg model using Eq. (5) with various different variational ansatz (here we use the tetrahedral POVM basis (see Supplemental Material [82] Sec. II). In looking at the quality of  $\langle \sigma_z \rangle$  we find that increasing the size of the Transformer both in depth and hidden dimension improves the result although this improvement is marginal until we reach two transformer layers and a hidden dimension of 64. Interestingly, we find that the use of strings has a significant effect on our results [see Fig. 1(i)]. To begin with, the use of string 1 is marginally superior to string 0. We expect this is because string 1 better addresses local correlations. More importantly, we find that there is a significant improvement (for any Transformer) by including more symmetry related strings out to the maximum of eight strings we considered. In fact, eight strings with one hidden layer and a hidden

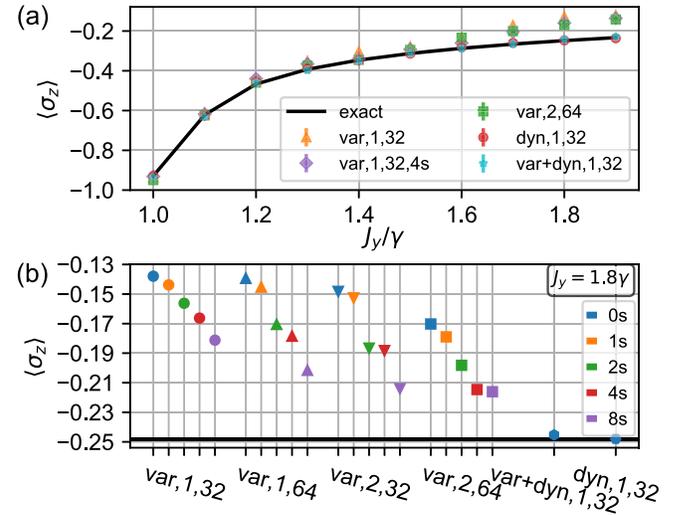


FIG. 4. Steady-state solutions for  $3 \times 3$  Heisenberg model with periodic boundary condition with  $B = 0$ ,  $J_x = 0.9\gamma$ , and  $J_z = \gamma$ . The exact curves (black lines) are produced using QuTiP [87,88]. (a)  $\langle \sigma_z \rangle$  for different values of  $J_y$  for POVM variational results (var), POVM dynamics (dyn) and POVM dynamics starting from the variational results (var + dyn). The two integers in the legend label are the number of transformer layers and hidden dimensions. (b) Steady-state solution  $J_y = 1.8\gamma$  comparing different variational ansatz. “0s” and “1s” use one string (string 0 and string 1); “2s”, “4s”, and “8s” use strings 1-2, 1-4, and 1-8, respectively [see Fig. 1(i)]. All initial states are  $\prod_{i=1}^N |\uparrow\rangle$  ( $\langle \sigma_z \rangle = 1$ ). The dynamics and variational plus dynamics approaches use the time step  $\tau = 0.008\gamma^{-1}$ . The results of two transformer layers are computed exactly under all POVM frame elements.

dimension of 32 provides a similar accuracy to 1 string with 2 hidden layers and a hidden dimension of 64. Additionally, we compared the results obtained through time evolution at long time to the fixed point method and found that the steady state approached by the time-evolved state provides significantly more accurate results. While the evaluation of the dynamics is computationally slower, we find that supplementing the fixed-point method with further dynamical evolution achieves the same steady-state solution as the dynamical approach at an overall reduced computational time.

*Conclusion.*—We have demonstrated an approach, whose run-time complexity per iteration step is polynomial on the system size and the hidden dimensions, to simulate the real-time dynamics of open quantum systems via an exact probabilistic formulation. By parametrizing the quantum state using an autoregressive Transformer, we accurately track the dynamics and steady state in 1D and 2D transverse field Ising and Heisenberg models. For 2D systems, we introduce string states which partially restore the symmetry of the Transformer.

Our methods constitute an important step in the machine learning approach for quantum many-body dynamics simulation. It provides the first exact sampling method for neural networks in OQS, which is a crucial improvement over the standard Markov chain Monte Carlo techniques with RBM, as well as an efficient stochastic optimization method for high dimensional differential equations. Our approach is versatile and applicable to general quantum dynamics in various contexts, including closed systems quantum dynamics, finite temperature dynamics of the density matrix, as well as challenging fermionic transport problems [89,90] with interactions to the environment [91]. Because of the probabilistic formulation as a quantum-classical mapping, our work has applications beyond quantum mechanics and demonstrates how to efficiently solve high-dimensional probabilistic differential equations with autoregressive neural networks. Such probabilistic equations appear in a wide variety of classical contexts and our Letter represents an important step forward in the direction.

D. L. is grateful for the insightful discussion with Filippo Vicentini, and appreciates very much the help from Filippo Vicentini, Alberto Biella, and Cristiano Ciuti on providing the original data from their paper [58]. D. L. would also like to thank Mohamed Hibat-Allah for sharing his insights on the RNN wave function. Z. C. is indebted to Qiwei Zhang for her contribution in digitizing Fig. 3 for the exact result and drawing string figures in Fig. 1. J. C. acknowledges support from Natural Sciences and Engineering Research Council of Canada (NSERC), the Shared Hierarchical Academic Research Computing Network (SHARCNET), Compute Canada, Google Quantum Research Award, and the Canadian Institute for Advanced Research (CIFAR) AI chair program. B. K. C. acknowledges support from the

Department of Energy Grant No. DOE desc0020165. This work utilized resources supported by the National Science Foundations Major Research Instrumentation program, Grant No. 1725729, as well as the University of Illinois at Urbana-Champaign. Z. C. acknowledges support from the A. C. Anderson Summer Research Award.

\*These authors contributed equally to this work.

- [1] F. Verstraete, M. M. Wolf, and J. Ignacio Cirac, *Nat. Phys.* **5**, 633 (2009).
- [2] J. T. Barreiro, M. Müller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, *Nature (London)* **470**, 486 (2011).
- [3] L. M. Sieberer, M. Buchhold, and S. Diehl, *Rep. Prog. Phys.* **79**, 096001 (2016).
- [4] M. F. Maghrebi and A. V. Gorshkov, *Phys. Rev. B* **93**, 014307 (2016).
- [5] E. Mascarenhas, H. Flayac, and V. Savona, *Phys. Rev. A* **92**, 022116 (2015).
- [6] J. Cui, J. I. Cirac, and M. C. Bañuls, *Phys. Rev. Lett.* **114**, 220601 (2015).
- [7] D. Jaschke, S. Montangero, and L. D. Carr, *Quantum Sci. Technol.* **4**, 013001 (2018).
- [8] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, *Phys. Rev. Lett.* **116**, 237201 (2016).
- [9] A. Biella, J. Jin, O. Viyuela, C. Ciuti, R. Fazio, and D. Rossini, *Phys. Rev. B* **97**, 035103 (2018).
- [10] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio, and D. Rossini, *Phys. Rev. X* **6**, 031011 (2016).
- [11] S. Finazzi, A. Le Boité, F. Storme, A. Baksic, and C. Ciuti, *Phys. Rev. Lett.* **115**, 080604 (2015).
- [12] R. Rota, F. Storme, N. Bartolo, R. Fazio, and C. Ciuti, *Phys. Rev. B* **95**, 134431 (2017).
- [13] R. Rota, F. Minganti, C. Ciuti, and V. Savona, *Phys. Rev. Lett.* **122**, 110405 (2019).
- [14] W. Casteels, R. M. Wilson, and M. Wouters, *Phys. Rev. A* **97**, 062107 (2018).
- [15] A. Nagy and V. Savona, *Phys. Rev. A* **97**, 052129 (2018).
- [16] N. Shammah, S. Ahmed, N. Lambert, S. De Liberato, and F. Nori, *Phys. Rev. A* **98**, 063815 (2018).
- [17] F. Vicentini, F. Minganti, A. Biella, G. Orso, and C. Ciuti, *Phys. Rev. A* **99**, 032115 (2019).
- [18] A. Kshetrimayum, H. Weimer, and R. Orús, *Nat. Commun.* **8**, 1291 (2017).
- [19] I. Carusotto and C. Ciuti, *Rev. Mod. Phys.* **85**, 299 (2013).
- [20] M. J. Hartmann, *J. Optics* **18**, 104005 (2016).
- [21] C. Noh and D. G. Angelakis, *Rep. Prog. Phys.* **80**, 016401 (2017).
- [22] N. Bartolo, F. Minganti, W. Casteels, and C. Ciuti, *Phys. Rev. A* **94**, 033841 (2016).
- [23] A. Biella, F. Storme, J. Lebreuilly, D. Rossini, R. Fazio, I. Carusotto, and C. Ciuti, *Phys. Rev. A* **96**, 023839 (2017).
- [24] M. Biondi, G. Blatter, H. E. Türeci, and S. Schmidt, *Phys. Rev. A* **96**, 043809 (2017).
- [25] H. J. Carmichael, *Phys. Rev. X* **5**, 031028 (2015).
- [26] W. Casteels, R. Fazio, and C. Ciuti, *Phys. Rev. A* **95**, 012128 (2017).

- [27] W. Casteels, F. Storme, A. Le Boité, and C. Ciuti, *Phys. Rev. A* **93**, 033824 (2016).
- [28] J. M. Fink, A. Dombi, A. Vukics, A. Wallraff, and P. Domokos, *Phys. Rev. X* **7**, 011012 (2017).
- [29] T. Fink, A. Schade, S. Höfling, C. Schneider, and A. Imamoglu, *Nat. Phys.* **14**, 365 (2018).
- [30] M. Fitzpatrick, N. M. Sundaresan, A. C. Y. Li, J. Koch, and A. A. Houck, *Phys. Rev. X* **7**, 011016 (2017).
- [31] M. Foss-Feig, P. Niroula, J. T. Young, M. Hafezi, A. V. Gorshkov, R. M. Wilson, and M. F. Maghrebi, *Phys. Rev. A* **95**, 043826 (2017).
- [32] E. M. Kessler, G. Giedke, A. Imamoglu, S. F. Yelin, M. D. Lukin, and J. I. Cirac, *Phys. Rev. A* **86**, 012116 (2012).
- [33] J. Marino and S. Diehl, *Phys. Rev. Lett.* **116**, 070407 (2016).
- [34] V. Savona, *Phys. Rev. A* **96**, 033826 (2017).
- [35] L. M. Sieberer, S. D. Huber, E. Altman, and S. Diehl, *Phys. Rev. Lett.* **110**, 195301 (2013).
- [36] F. Vicentini, F. Minganti, R. Rota, G. Orso, and C. Ciuti, *Phys. Rev. A* **97**, 013853 (2018).
- [37] T. E. Lee, S. Gopalakrishnan, and M. D. Lukin, *Phys. Rev. Lett.* **110**, 257204 (2013).
- [38] R. Rota, F. Minganti, A. Biella, and C. Ciuti, *New J. Phys.* **20**, 045003 (2018).
- [39] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, *Rev. Mod. Phys.* **93**, 025005 (2021).
- [40] N. Yoshioka, Y. O. Nakagawa, K. Mitarai, and K. Fujii, *Phys. Rev. Research* **2**, 043289 (2020).
- [41] Z. Liu, L. M. Duan, and D.-L. Deng, [arXiv:2008.05488](https://arxiv.org/abs/2008.05488).
- [42] C.-K. Lee, P. Patil, S. Zhang, and C.-Y. Hsieh, *Phys. Rev. Research* **3**, 023095 (2021).
- [43] N. Ramusat and V. Savona, *Quantum* **5**, 399 (2021).
- [44] H.-Y. Liu, T.-P. Sun, Y.-C. Wu, and G.-P. Guo, *Chin. Phys. Lett.* **38**, 080301 (2021).
- [45] O. Scarlatella, A. A. Clerk, R. Fazio, and M. Schir, *Phys. Rev. X* **11**, 031018 (2021).
- [46] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [47] M. Zwolak and G. Vidal, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [48] I. E. Lagaris, A. Likas, and D. I. Fotiadis, *Comput. Phys. Commun.* **104**, 1 (1997).
- [49] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [50] D. Luo and B. K. Clark, *Phys. Rev. Lett.* **122**, 226401 (2019).
- [51] D. Pfau, J. S. Spencer, A. G. de G. Matthews, and W. M. C. Foulkes, *Phys. Rev. Research* **2**, 033429 (2020).
- [52] J. Hermann, Z. Schtzle, and F. No, [arXiv:1909.08423](https://arxiv.org/abs/1909.08423).
- [53] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, *Phys. Rev. Research* **2**, 023358 (2020).
- [54] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, *Phys. Rev. Lett.* **124**, 020503 (2020).
- [55] S. Lu, X. Gao, and L.-M. Duan, *Phys. Rev. B* **99**, 155136 (2019).
- [56] X. Gao and L.-M. Duan, *Nat. Commun.* **8**, 662 (2017).
- [57] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, *Phys. Rev. X* **8**, 011006 (2018).
- [58] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, *Phys. Rev. Lett.* **122**, 250503 (2019).
- [59] N. Yoshioka and R. Hamazaki, *Phys. Rev. B* **99**, 214306 (2019).
- [60] M. J. Hartmann and G. Carleo, *Phys. Rev. Lett.* **122**, 250502 (2019).
- [61] A. Nagy and V. Savona, *Phys. Rev. Lett.* **122**, 250501 (2019).
- [62] D. Yuan, H. Wang, Z. Wang, and D.-L. Deng, *Phys. Rev. Lett.* **126**, 160401 (2021).
- [63] G. Torlai and R. G. Melko, *Phys. Rev. Lett.* **120**, 240503 (2018).
- [64] F. Ferrari, F. Becca, and J. Carrasquilla, *Phys. Rev. B* **100**, 125131 (2019).
- [65] T. Westerhout, N. Astrakhantsev, K. S. Tikhonov, M. I. Katsnelson, and A. A. Bagrov, *Nat. Commun.* **11**, 1593 (2020).
- [66] A. Szabó and C. Castelnovo, *Phys. Rev. Research* **2**, 033075 (2020).
- [67] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, *Nat. Phys.* **5**, 27 (2009).
- [68] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, *Nat. Mach. Intell.* **1**, 155 (2019).
- [69] J. Carrasquilla, D. Luo, F. Prez, A. Milsted, B. K. Clark, M. Volkovs, and L. Aolita, *Phys. Rev. A* **104**, 032610 (2021).
- [70] E. O. Kiktenko, A. O. Malyshev, A. S. Mastiukova, V. I. Manko, A. K. Fedorov, and D. Chrućiski, *Phys. Rev. A* **101**, 052320 (2020).
- [71] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, [arXiv:1706.03762](https://arxiv.org/abs/1706.03762).
- [72] A. Iserles, Euler's method and beyond, in *A First Course in the Numerical Analysis of Differential Equations*, 2nd ed., Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, England, 2008), p. 813.
- [73] S. Hochreiter and J. Schmidhuber, *Neural Comput.* **9**, 1735 (1997).
- [74] K. Cho, B. van Merriënboer, Ç. Gülçehre, F. Bougares, H. Schwenk, and Y. Bengio, [arXiv:1406.1078](https://arxiv.org/abs/1406.1078).
- [75] A. van den Oord, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves, and K. Kavukcuoglu, [arXiv:1606.05328](https://arxiv.org/abs/1606.05328).
- [76] P. Cha, P. Ginsparg, F. Wu, J. Carrasquilla, P. L. McMahon, and E.-A. Kim, *Mach. Learn.: Sci. Technol.* **3**, 01LT01 (2021).
- [77] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics (Springer New York Inc., New York, NY, USA, 2001).
- [78] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, *Phys. Rev. Lett.* **100**, 040501 (2008).
- [79] A. Mahajan and S. Sharma, *J. Phys. Chem. A* **123**, 3911 (2019).
- [80] D. Tahara and M. Imada, *J. Phys. Soc. Jpn.* **77**, 114701 (2008).
- [81] Y. Qiu, T. M. Henderson, J. Zhao, and G. E. Scuseria, *J. Chem. Phys.* **147**, 064111 (2017).
- [82] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.128.090501> for detailed explanation of the algorithms and additional data.
- [83] A. Paszke *et al.*, [arXiv:1912.01703](https://arxiv.org/abs/1912.01703).
- [84] D. P. Kingma and J. Ba, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [85] D. Kochkov and B. K. Clark, [arXiv:1811.12423](https://arxiv.org/abs/1811.12423).
- [86] I. L. Gutierrez and C. B. Mendl, [arXiv:1912.08831](https://arxiv.org/abs/1912.08831).

- [87] J. Johansson, P. Nation, and F. Nori, *Comput. Phys. Commun.* **184**, 1234 (2013).  
[88] J. Johansson, P. Nation, and F. Nori, *Comput. Phys. Commun.* **183**, 1760 (2012).  
[89] F. M. Souza and L. Sanz, *Phys. Rev. A* **96**, 052110 (2017).  
[90] Y. Yan, *J. Chem. Phys.* **140**, 054105 (2014).  
[91] T. C. Berkelbach and M. Thoss, *J. Chem. Phys.* **152**, 020401 (2020).