Editors' Suggestion | Featured in Physics

Variational Neural-Network Ansatz for Steady States in Open Quantum Systems

Filippo Vicentini,[1] Alberto Biella,[1] Nicolas Regnault,[2] and Cristiano Ciuti[1]
[1]Université de Paris, Laboratoire Matériaux et Phénomènes Quantiques, CNRS, F-75013, Paris, France
[2]Laboratoire de Physique de l'Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Sorbonne Paris Cité, F-75005, Paris, France

(Received 28 February 2019; published 28 June 2019)

We present a general variational approach to determine the steady state of open quantum lattice systems via a neural-network approach. The steady-state density matrix of the lattice system is constructed via a purified neural-network *Ansatz* in an extended Hilbert space with ancillary degrees of freedom. The variational minimization of cost functions associated to the master equation can be performed using a Markov chain Monte Carlo sampling. As a first application and proof of principle, we apply the method to the dissipative quantum transverse Ising model.

DOI: 10.1103/PhysRevLett.122.250503

In spite of the tremendous experimental progress in the isolation of quantum systems, a finite coupling to the environment [1] is unavoidable and certainly plays a crucial role in the practical implementation of quantum information and quantum simulation protocols [2]. Moreover, through an active control of the environment via the so-called reservoir engineering, an open quantum manybody system can be prepared in nontrivial phases [3–5] with also possible quantum applications [6,7]. The theoretical description of open quantum manybody systems is in general out-of-the equilibrium and much less developed than for equilibrium systems. A mixed state with a finite entropy can be described by a density matrix, whose evolution is described by a master equation. Recently, a few theoretical methods have been developed to solve the master equation of open quantum manybody systems, including analytical approaches based on the Keldysh formalism [8,9], numerical algorithms based on matrix product operator and tensor-network techniques [10–14], cluster mean-field methods [15,16], corner-space renormalization [17–19], Gutzwiller mean-field [20], full configuration-interaction Monte Carlo methods [21], permutation-invariant solvers [22], or efficient stochastic unravelings for disordered systems [23]. The research in the field is very active, since the different methods are optimal for different specific regimes. For example, the corner-space renormalization method is best suited for systems with moderate entropy, while matrix product operator techniques are best suited to systems with short-range quantum correlations.

In the last decade, the field of artificial neural networks has enjoyed a dramatic expansion and success thanks to remarkable applications in the recognition of complex patterns such as visual images or human speech (for a recent review see, e.g., Ref. [24]). The optimization (supervised learning) of the network is obtained by tuning the weights quantifying the connections between neural units via a variational minimization of a properly defined cost function. The wave function of a manybody system is in general a complex quantity, which is hard to be recognized. Recent works have proposed to exploit artificial neural networks to construct trial wave functions, where the connection weights in the network play the role of variational parameters [25,26]. Neural-network approaches have already been successfully applied to a wide number (see e.g. Refs. [27–31]) of close Hamiltonian systems. However, they have not yet been generalized to the important quantum manybody problem of open systems.

In this Letter, we present a theoretical approach based on a variational neural-network *Ansatz* in order to determine the steady state of the master equation of open quantum lattice systems. We construct the *Ansatz* for the mixed density matrix starting from a restricted Boltzman machine *Ansatz* for a pure many-body wave function in an extended Hilbert space. We determine the optimal variational parameters by minimizing a cost function which involves the Liouvillian superoperator associated to the master equation for the density matrix. As a first application, we have considered the dissipative tranverse field quantum Ising model. We present a proof-of-principle demonstration by benchmarking the neural-network calculations of the steady state against numerically exact simulations performed by quantum trajectories in the full Hilbert space [32]. Our minimization of the cost function is performed by a Markov chain Monte Carlo sampling of the gradient and is thus scalable to a large number of lattice sites. Perspectives of the present approach are discussed in the conclusions.

The general task that we wish to solve is the determination of the steady state of an open quantum system described by the Lindblad master equation [1] for the system reduced density matrix $\hat{\rho}$, which reads (setting $\hbar = 1$):

$$\dot{\hat{\rho}} = \mathcal{L}\hat{\rho} = -i[\hat{H},\hat{\rho}] + \sum_j \frac{\gamma_j}{2}[2\hat{L}_j\hat{\rho}\hat{L}_j^\dagger - \{\hat{L}_j^\dagger\hat{L}_j,\hat{\rho}\}], \quad (1)$$

where $\mathcal{L}$ is the so-called Liouvillian superoperator depending on the system Hamiltonian operator $\hat{H}$. The coupling to the environment is represented by interaction channels with the reservoir characterized by dissipation rates γ_j and jump operators $\hat{L}_j$ acting on the system. Here we will focus on situations where the steady state ($\partial_t\hat{\rho}_{\rm SS} = 0$) is unique. In this case, the steady-state density matrix can be obtained as $\hat{\rho}_{\rm SS} = \lim_{t\to\infty}\hat{\rho}(t)$ regardless of the initial condition. Although it is possible to engineer peculiar Liouvillians with more than one steady state [33], typical physical systems with a finite Hilbert space dimension have a unique steady state [34–36].

For the many-body problem an analytical expression for $\hat{\rho}_{\rm SS}$ can be found in very few cases [37,38]. In general, because of the exponential growth of the Hilbert space with the number of lattice sites, describing the full density matrix requires exponentially many complex numbers, which in practice can be done exactly only for a small number of sites. If one wants to attack the problem within a variational framework, the density matrix can be represented by an *Ansatz* $\hat{\rho}_{\boldsymbol{\nu}}$ depending on a set of variational parameters $\boldsymbol{\nu}$. If $\{|\boldsymbol{\sigma}\rangle = |\sigma_1,\sigma_2,\ldots,\sigma_N\rangle\}$ denotes a basis of states for the system Hilbert space, the density matrix can be expressed in the form

$$\hat{\rho}(\boldsymbol{\nu}) = \sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'}\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}')|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}'|. \quad (2)$$

In order to construct our neural-network *Ansatz* for the density matrix, we consider an extended Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_A$ where $\mathcal{H}_{S,A}$ represents respectively the system and ancillary Hilbert spaces. Such extended space is spanned by the basis set $\{|\boldsymbol{\sigma},\boldsymbol{a}\rangle\}$ where $\boldsymbol{a} = (a_1,a_2,\ldots,a_{N_a})$ labels the ancillary degrees of freedom. We start by considering a pure state in the extended Hilbert space, represented by the wave function $\psi_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{a})$. In this framework the reduced density matrix of the system S is obtained by tracing out the ancillary degrees freedom [39], namely

$$\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}') = \sum_{\boldsymbol{a}}\psi_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{a})\psi^\star_{\boldsymbol{\nu}}(\boldsymbol{\sigma}',\boldsymbol{a}). \quad (3)$$

The next step is to represent $\psi_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{a})$ via a neural-network *Ansatz*. This purified procedure automatically ensures that $\hat{\rho}_{\boldsymbol{\nu}}$ is Hermitian and positive semidefinite, as required for a density matrix. In a recent paper, Torlai and Melko [39] proposed to describe purified wave functions as

$$\psi_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{a}) = \sqrt{\mathcal{P}_{\boldsymbol{\nu}_A}(\boldsymbol{\sigma},\boldsymbol{a})}\exp\left\{-\frac{1}{2}\log[\mathcal{P}_{\boldsymbol{\nu}_\theta}(\boldsymbol{\sigma},\boldsymbol{a})]\right\}. \quad (4)$$

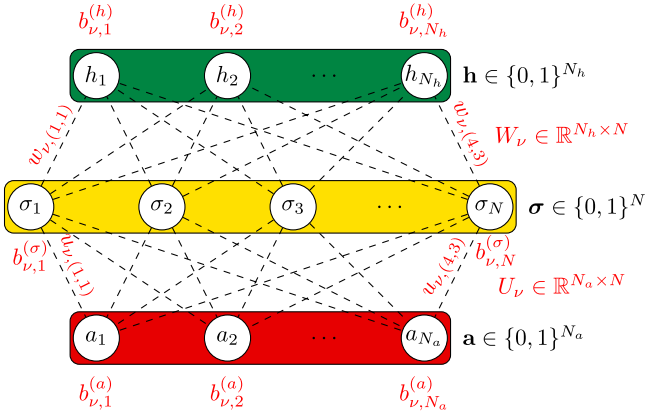

FIG. 1. Graph representation of the artificial neural network used for the density matrix *Ansatz*. The vector $\boldsymbol{\sigma} = (\sigma_1,\sigma_2,\ldots,\sigma_N)$ contains the variables of the physical system (visible layer). The vector $\boldsymbol{a} = (a_1,a_2,\ldots,\sigma_{N_a})$ describes the ancillary degrees of freedom of the extended Hilbert space (ancilla layer), where a purified density matrix is considered [see Eq. (3)]. The vector $\boldsymbol{h} = (h_1,h_2,\ldots,h_{N_h})$ contains variables of auxiliary nodes (hidden layer). The network parameters are $\boldsymbol{\nu}_\nu = (\mathbf{b}^{(\sigma)}_\nu,\boldsymbol{b}^{(a)}_\nu,\boldsymbol{b}^{(h)}_\nu,\boldsymbol{W}_\nu,\boldsymbol{U}_\nu)$. One network is used for the representation of the amplitude ($\nu = \boldsymbol{\nu}_A$) of the purified wave function in Eq. (3), while another independent network with the same topology is used to represent the phase ($\nu = \boldsymbol{\nu}_\theta$).

Both the amplitude $\mathcal{P}_{\boldsymbol{\nu}_A}(\boldsymbol{\sigma},\boldsymbol{a})$ and phase-related function $\mathcal{P}_{\boldsymbol{\nu}_\theta}(\boldsymbol{\sigma},\boldsymbol{a})$ of the purified wave function are given by the Boltzmann-like expression $\mathcal{P}_\nu(\boldsymbol{\sigma},\boldsymbol{a}) = \sum_{\boldsymbol{h}}\exp[-E_\nu(\boldsymbol{\sigma},\boldsymbol{a},\boldsymbol{h})]$ (with $\nu \in \{\boldsymbol{\nu}_A,\boldsymbol{\nu}_\theta\}$), where the associated dimensionless energy reads

$$E_\nu(\boldsymbol{\sigma},\boldsymbol{a},\boldsymbol{h}) = \boldsymbol{\sigma}\cdot\mathbf{b}^{(\sigma)}_\nu + \boldsymbol{a}\cdot\boldsymbol{b}^{(a)}_\nu + \boldsymbol{h}\cdot\boldsymbol{b}^{(h)}_\nu + \boldsymbol{\sigma}^T\boldsymbol{W}_\nu\boldsymbol{h} + \boldsymbol{\sigma}^T\boldsymbol{U}_\nu\boldsymbol{a}. \quad (5)$$

Note that the *Ansatz* parameters are $\boldsymbol{\nu} = (\boldsymbol{\nu}_A,\boldsymbol{\nu}_\theta)$ where $\boldsymbol{\nu}_\nu = (\mathbf{b}^{(\sigma)}_\nu,\boldsymbol{b}^{(a)}_\nu,\boldsymbol{b}^{(h)}_\nu,\boldsymbol{W}_\nu,\boldsymbol{U}_\nu)$. The rectangular matrix $\boldsymbol{W}_\nu$ weighs the connections between the system variables (visible layer) to the auxiliary variables (hidden layer), while the weight matrix $\boldsymbol{U}_\nu$ quantifies the connection between the system variables and the ancillary ones (ancillary layer). Such a neural-network *Ansatz* is represented by a tripartite restricted Boltzmann machine depicted in Fig. 1. In other words, there are two independent artificial neural networks, one for the amplitude ($\nu = A$) and one for the phase ($\nu = \theta$). By substituting those formulas into Eq. (3) and carrying out the sum over the ancillary degrees of freedom, one obtains a closed formula for the entries of the density matrix:

$$\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}') = \exp[\Gamma^-_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \Gamma^+_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \Pi_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma}')], \quad (6)$$

where the expression of $\Gamma^{+,-}$ and Π can be found in the Supplemental Material [40]. The representation power [47–49] of this *Ansatz* can be systematically improved by

increasing the density of the hidden ($\alpha = N_h/N$) and ancillary layer ($\beta = N_a/N$). It is worth pointing out that this scheme is not specific to this network topology, but relies only on the general fact that if two visible layers are connected by a shallow ancillary layer, the ancilla can be traced out analytically and an efficient neural-network description of the density matrix can be obtained.

Having defined a variational *Ansatz* $\hat{\rho}(\boldsymbol{\nu})$, we now wish to define a variational principle to determine the optimal parameters. In particular, we have to recast the search for the steady state into a minimization problem for a real, positive cost function $\mathcal{C}(\boldsymbol{\nu})$, which has a global minimum when the master equation $\mathcal{L}\hat{\rho}_{\boldsymbol{\nu}} = 0$ is satisfied [50]. Moreover, in order to be able to deal with large Hilbert spaces, we need a quantity that can be sampled and computed efficiently. These requirements are met by the following cost function expressed in terms of the 2-norm of the time derivative of the density matrix:

$$\mathcal{C}(\boldsymbol{\nu}) = \frac{||d\hat{\rho}_{\boldsymbol{\nu}}/dt||_2^2}{||\hat{\rho}_{\boldsymbol{\nu}}||_2^2} = \frac{\mathrm{Tr}[\hat{\rho}_{\boldsymbol{\nu}}^\dagger \mathcal{L}^\dagger \mathcal{L}\hat{\rho}_{\boldsymbol{\nu}}]}{\mathrm{Tr}[\hat{\rho}_{\boldsymbol{\nu}}^\dagger \hat{\rho}_{\boldsymbol{\nu}}]}, \tag{7}$$

as (i) $\mathcal{C}(\boldsymbol{\nu}_{\mathrm{SS}}) = 0 \Leftrightarrow \hat{\rho}(\boldsymbol{\nu}_{\mathrm{SS}}) = \hat{\rho}_{\mathrm{SS}}$ and (ii) $\mathcal{C}(\boldsymbol{\nu}) \geq 0$.

It is useful to rewrite Eq. (7) as a sum over the whole space of bounded operators on the Hilbert space:

$$\mathcal{C}(\boldsymbol{\nu}) = \sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}} p_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) |\mathcal{C}^{\mathrm{loc}}(\boldsymbol{\nu},\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})|^2, \tag{8}$$

这部分类似求loc能量的内容

where $p_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) = |\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})|^2/Z$ corresponds to a probability distribution as $Z = \sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}} |\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})|^2$ [51]. The local contribution reads [52]:

$$\mathcal{C}^{\mathrm{loc}}(\boldsymbol{\nu},\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) = \sum_{\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}'} \mathcal{L}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}};\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}') \frac{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}')}{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})}. \tag{9}$$

In order to find the global minimum of the cost function [Eq. (7)] by means of gradient-based iterative schemes, we need to compute its gradient

$$\begin{aligned}\nabla_{\boldsymbol{\nu}}\mathcal{C}(\boldsymbol{\nu}) = &\sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}} p_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) \mathcal{C}^{\mathrm{loc}}(\boldsymbol{\nu},\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})^\star \\ &\times \left[\sum_{\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}'} \mathcal{L}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}};\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}') \frac{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}')}{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})} \mathcal{O}_{\boldsymbol{\nu}}(\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}')\right] \\ &- \mathcal{C}(\boldsymbol{\nu})\mathcal{O}_{\boldsymbol{\nu}},\end{aligned} \tag{10}$$

where we have defined the log-derivatives of the density matrix $\mathcal{O}_{\boldsymbol{\nu}} = \sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}} \mathcal{O}_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})$ and $\mathcal{O}_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) = \nabla_{\boldsymbol{\nu}} \log \rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})$, which can be efficiently computed for the considered neural network.

The computational complexity of evaluating $\nabla\mathcal{C}(\boldsymbol{\nu})$ exactly grows exponentially with the size of the system. This cost can be considerably reduced if one only uses an estimate of $\nabla\mathcal{C}(\boldsymbol{\nu})$ obtained by sampling the values $(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})$ according to the probability $p_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})$. Because the normalization factor Z is not fixed, we cannot sample the distribution directly and have to resort to a Markov chain Monte Carlo method [53] with Metropolis update rules [54]. At every sampling step, we propose to update the configuration $(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}) \to (\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}')$ by switching a random number of spins and accept the new configuration with probability $\min\left(\exp[p_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})/p_{\boldsymbol{\nu}}(\boldsymbol{\sigma}',\tilde{\boldsymbol{\sigma}}')], 1\right)$.

Finally, in order to find the global minimum of the cost function, we employ a standard stochastic gradient descent algorithm [55]. In order to improve the performance of the stochastic gradient descent (i.e., to reduce the number of iterations needed to converge to the global minima of the cost function) we update the variational parameters according to the metric of the space of density matrices exploiting the stochastic reconfiguration approach [41]. During the optimization procedure we sample the physical observables of interest through another Markov chain as

$$\langle\hat{\Theta}\rangle = \frac{\mathrm{Tr}[\hat{\rho}\,\hat{\Theta}]}{\mathrm{Tr}[\hat{\rho}]} = \sum_{\boldsymbol{\sigma}} p_{\boldsymbol{\nu}}^{\mathrm{obs}}(\boldsymbol{\sigma}) \sum_{\tilde{\boldsymbol{\sigma}}} \frac{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}})\Theta(\tilde{\boldsymbol{\sigma}},\boldsymbol{\sigma})}{\rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma})}, \tag{11}$$

where $p_{\boldsymbol{\nu}}^{\mathrm{obs}}(\boldsymbol{\sigma}) = \rho_{\boldsymbol{\nu}}(\boldsymbol{\sigma},\boldsymbol{\sigma})/\mathrm{Tr}[\hat{\rho}]$.

In order to benchmark our neural-network approach for open quantum systems, we consider here the dissipative quantum transverse Ising model, whose Hamiltonian is

$$H = \frac{V}{4} \sum_{\langle j,l\rangle} \hat{\sigma}_j^z \hat{\sigma}_l^z + \frac{g}{2} \sum_j \hat{\sigma}_j^x, \tag{12}$$

being $\hat{\sigma}_j^\alpha$ the Pauli matrices ($\alpha \in \{x,y,z\}$) acting on the jth site. The first term represents the nearest-neighbor spin-spin interaction depending only on the z components, V being the coupling strength. The second term accounts for a local and uniform magnetic field along the transverse direction x. We consider local dissipative spin-flip processes described by the site-dependent jump operator $\hat{L}_j^{(z)} = \hat{\sigma}_j^- = \frac{1}{2}(\hat{\sigma}_j^x - i\hat{\sigma}_j^y)$, which fully determine the master equation in Eq. (1).

Numerical results for steady-state observables of the dissipative quantum transverse Ising model on a 1D periodic chain are reported in Fig. 2. In particular, we report the spatial components of the averaged magnetization as a function of the magnetic field g (in units of the dissipation rate γ) for $V/\gamma = 2$. For $N = 16$ lattice sites the predictions of the neural-network variational method (circles) are compared to the results obtained with a brute-force exact integration of the master equation in the whole Hilbert space, showing a good agreement over all the parameter range. For $g \lesssim \gamma$ and $g \gtrsim 2.5\gamma$ a remarkable precision is reached for all the local observables with a low density of the hidden and ancillary layer $\alpha = \beta = 1$ and $O(10^2)$ minimization steps. For $1 \lesssim g/\gamma \lesssim 2.5$ a higher

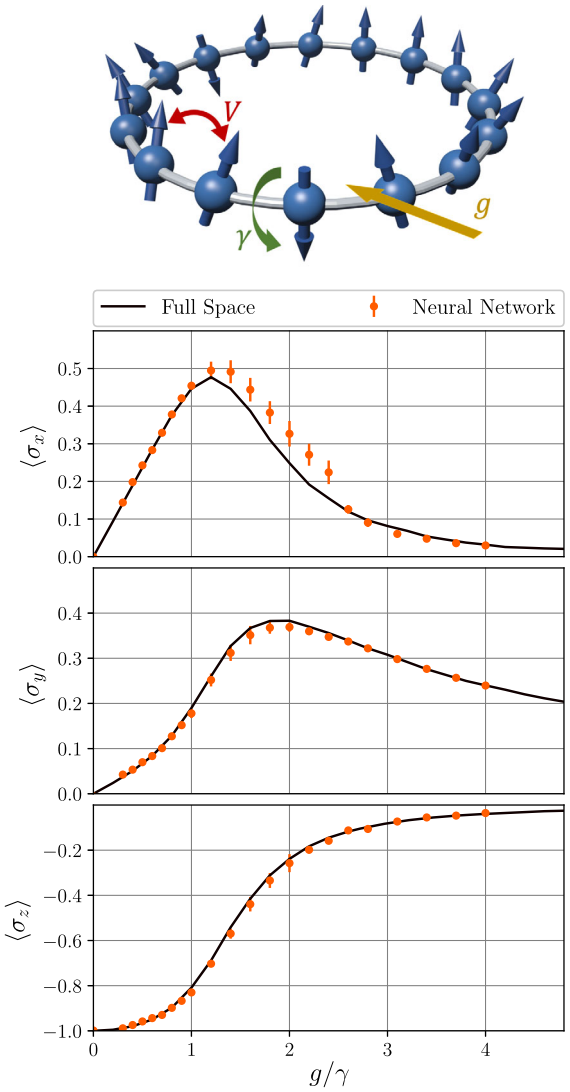

FIG. 2. Top panel: sketch of the considered physical system described by the dissipative quantum transverse Ising 1D model with periodic boundary conditions. The quantity g denotes an applied magnetic field, V the spin-spin coupling, and γ the spin flip rate. Bottom panel: the prediction of the neural-network variational calculations (circles) are compared to the results obtained by quantum trajectory simulations of the master equation by considering the whole Hilbert space (solid lines). The top, middle, and bottom panels depict the expectation values of the three components of the averaged magnetization as a function of the applied magnetic field g (in units of γ). Model parameters: $V/\gamma = 2$ (spin-spin coupling), $N = 16$ (number of lattice sites). Neural-network parameters: $\alpha = \beta = 1$ for $g \leq \gamma$ and $g \geq 2.5\gamma$ while $\alpha = 1$ and $\beta = 4$ for the remaining points. The parameters required for the convergence of the Monte Carlo calculations depend on the value of g/γ, with the intermediate region being the most demanding. The maximum number of accepted Monte Carlo samples is 8640 and the maximum number of steps for the stochastic gradient descent is 10^4. For points outside the intermediate region, 3000 accepted Monte Carlo samples and 10^3 iteration steps have been performed.

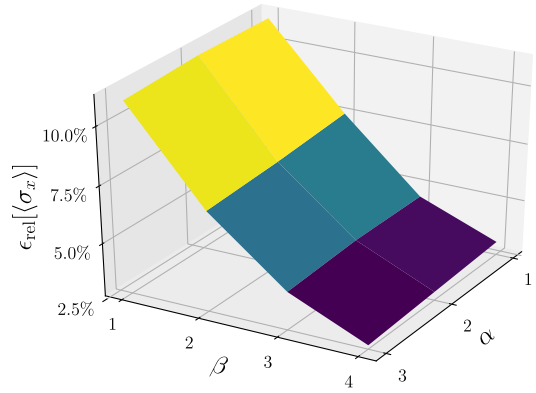

FIG. 3. Relative error with respect to the exact result for the observable $\langle\sigma_x\rangle$ as a function of α and β. Parameters are set as in Fig. 2 but for a fixed value $g/\gamma = 1.2$.

number of variational parameters is required. In particular, as shown in Fig. 3, a systematic improvement of the relative error $\epsilon_{\rm rel}[\langle\sigma_x\rangle]$ with respect to the exact solution can be obtained by increasing β. Interestingly, for $1 \lesssim g/\gamma \lesssim 2.5$, we note that the gradient-descent procedure requires more iterations. This region corresponds to the range of g/γ where the smallest nonzero eigenvalue of $\mathcal{L}^\dagger\mathcal{L}$ decreases significantly [56]. In this range the steady-state density matrix also displays nontrivial correlations and nonthermal mixness properties [56]. Remarkably, the fidelity of the reconstructed local density matrix with respect to the exact one is always larger than 0.998 for all the values of g/γ considered. Finally, as an example of convergence, the top panel of Fig. 4 depicts a typical evolution of the cost function in the iterative minimization procedure for a fixed set of parameters ($g/\gamma = 1$), showing a good convergence

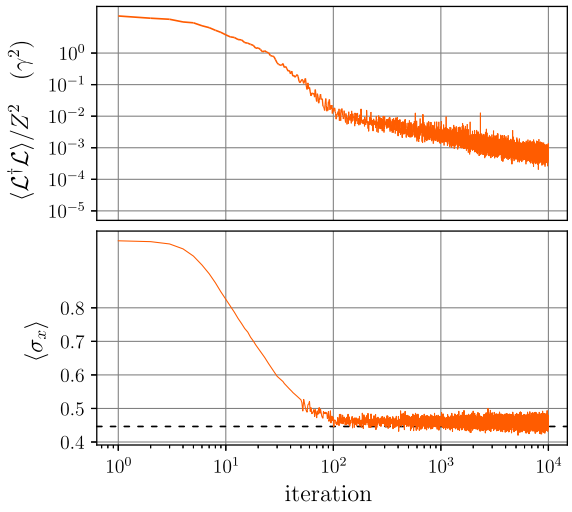

FIG. 4. Same parameters as in Fig. 2 but for a fixed value $g/\gamma = 1$. Top panel: the cost function is shown as a function of the iteration steps. Bottom panel: the corresponding evolution of the x component of the average magnetization during the stochastic minimization is shown.

towards the global minimum. In the bottom panel of Fig. 4, the convergence of the x component of the averaged magnetization is also reported.

In conclusion, we have presented a general variational approach for the steady-state density matrix of open quantum manybody systems based on an artificial neural-network scheme. Our method is scalable since the cost function associated to the Liouvillian of the master equation can be calculated via Monte Carlo sampling. We have demonstrated a proof of principle of the theoretical scheme by a successful benchmarking to brute-force finite-size simulations in the full Hilbert space for arrays of spins described by the dissipative quantum transverse Ising model. We would like to point out that the present approach does not depend on the specific network topology. Indeed, the variational procedure presented in this Letter is general and can be applied to many other neural networks or physically inspired variational *Ansätze*. There are many future developments at the horizon, including the study of dynamical properties, the use of deep neural networks and/or alternative cost functions, comparison with other existing techniques as well as the study of disordered systems without translational invariance. The neural-network approach has the potential to pave the way to the theoretical study of a wide spectrum of open quantum manybody systems.

Numerical code for this Letter has been written in *Julia* [57] and is accessible online [58].

We thank G. Carleo, V. Savona, and G. Orso for fruitful discussions. Full space simulations have been made with QuantumOptics.jl [59] and with QuTiP [60,61]. We acknowledge support from ERC (via Consolidator Grant CORPHO No. 616233). This work was granted access to the HPC resources of TGCC under the allocation 2018-A0050510601 attributed by GENCI (Grand Equipement National de Calcul Intensif).

Note added.—Recently, we became aware of related independent theoretical works that have been carried on in parallel [62–64].

[1] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, England, 2007), https://doi.org/10.1093/acprof:oso/9780199213900.001.0001.
[2] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169 (2013).
[3] I. Carusotto and C. Ciuti, Rev. Mod. Phys. **85**, 299 (2013).
[4] C. Noh and D. G. Angelakis, Rep. Prog. Phys. **80**, 016401 (2017).
[5] M. J. Hartmann, J. Opt. **18**, 104005 (2016).
[6] F. Verstraete, M. M. Wolf, and J. Ignacio Cirac, Nat. Phys. **5**, 633 (2009).
[7] 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).
[8] L. M. Sieberer, M. Buchhold, and S. Diehl, Rep. Prog. Phys. **79**, 096001 (2016).
[9] M. F. Maghrebi and A. V. Gorshkov, Phys. Rev. B **93**, 014307 (2016).
[10] E. Mascarenhas, H. Flayac, and V. Savona, Phys. Rev. A **92**, 022116 (2015).
[11] J. Cui, J. I. Cirac, and M. C. Bañuls, Phys. Rev. Lett. **114**, 220601 (2015).
[12] D. Jaschke, S. Montangero, and L. D. Carr, Quantum Sci. Technol. **4**, 013001 (2019).
[13] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Phys. Rev. Lett. **116**, 237201 (2016).
[14] A. Kshetrimayum, H. Weimer, and R. Orús, Nat. Commun. **8**, 1291 (2017).
[15] A. Biella, J. Jin, O. Viyuela, C. Ciuti, R. Fazio, and D. Rossini, Phys. Rev. B **97**, 035103 (2018).
[16] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio, and D. Rossini, Phys. Rev. X **6**, 031011 (2016).
[17] S. Finazzi, A. Le Boité, F. Storme, A. Baksic, and C. Ciuti, Phys. Rev. Lett. **115**, 080604 (2015).
[18] R. Rota, F. Storme, N. Bartolo, R. Fazio, and C. Ciuti, Phys. Rev. B **95**, 134431 (2017).
[19] R. Rota, F. Minganti, C. Ciuti, and V. Savona, Phys. Rev. Lett. **122**, 110405 (2019).
[20] W. Casteels, R. M. Wilson, and M. Wouters, Phys. Rev. A **97**, 062107 (2018).
[21] A. Nagy and V. Savona, Phys. Rev. A **97**, 052129 (2018).
[22] N. Shammah, S. Ahmed, N. Lambert, S. De Liberato, and F. Nori, Phys. Rev. A **98**, 063815 (2018).
[23] F. Vicentini, F. Minganti, A. Biella, G. Orso, and C. Ciuti, Phys. Rev. A **99**, 032115 (2019).
[24] Y. LeCun, Y. Bengio, and G. Hinton, Nature (London) **521**, 436 (2015).
[25] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[26] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X **8**, 011006 (2018).
[27] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[28] F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
[29] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Phys. Rev. Lett. **121**, 167204 (2018).
[30] S. Czischek, M. Gärttner, and T. Gasenzer, Phys. Rev. B **98**, 024311 (2018).
[31] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, arXiv:1902.04057.
[32] A. J. Daley, Adv. Phys. **63**, 77 (2014).
[33] V. V. Albert and L. Jiang, Phys. Rev. A **89**, 022118 (2014).
[34] H. Spohn, Lett. Math. Phys. **2**, 33 (1977).
[35] F. Minganti, A. Biella, N. Bartolo, and C. Ciuti, Phys. Rev. A **98**, 042118 (2018).
[36] D. Nigro, J. Stat. Mech. (2019) 043202.
[37] T. Prosen, Phys. Rev. Lett. **112**, 030603 (2014).
[38] T. Prosen, Phys. Rev. Lett. **107**, 137201 (2011).
[39] G. Torlai and R. G. Melko, Phys. Rev. Lett. **120**, 240503 (2018).

[40] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.122.250503 for more details, which includes Refs. [11,41–46].
[41] F. Becca and S. Sorella, in *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, England, 2017), pp. 131–155, https://doi.org/10.1017/9781316417041.007.
[42] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).
[43] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).
[44] J. Martens, arXiv:1412.1193.
[45] R. Pascanu and Y. Bengio, arXiv:1301.3584.
[46] S.-C. T. Choi and M. A. Saunders, ACM Trans. Math. Softw. **40**, 1 (2014).
[47] N. L. Roux and Y. Bengio, Neural Comput. **20**, 1631 (2008).
[48] L. Younes, Appl. Math. Lett. **9**, 109 (1996).
[49] G. Montufar and N. Ay, Neural Comput. **23**, 1306 (2011).
[50] H. Weimer, Phys. Rev. Lett. **114**, 040402 (2015).
[51] Being able to rewrite the cost function [Eq. (7)] into this form allows us to sample it efficiently. This is the main advantage with respect to using the trace norm of $\mathcal{L}\hat{\rho}$ [50] as a cost function, which requires computing the singular value decomposition of $\mathcal{L}\rho$ at each iteration.
[52] This is not the only possible expression that allows to sample the cost function $\mathcal{C}(\boldsymbol{v})$. However, this choice of the local cost function $\mathcal{C}^{\mathrm{loc}}$ is particularly convenient since it respects the zero-variance property, it is more numerically stable and cheaper to compute. For further details see Sec. III of the Supplemental Material [40].
[53] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, England, 2017), https://doi.org/10.1017/9781316417041.
[54] We point out that a promising direction of research would be to devise particular trial wave functions where Z is fixed or cheaper to compute, so that a direct sampling of the distribution $p_{\sigma,\tilde{\sigma}}$ without a Markov Chain would lead to easier convergence properties. Indeed, it has been recently shown [31] that a direct sampling is possible in some types of networks.
[55] L. Bottou, in *Proceedings of COMPSTAT'2010*, edited by Y. Lechevallier and G. Saporta (Physica-Verlag HD, Heidelberg, 2010), pp. 177–186.
[56] J. Jin, A. Biella, O. Viyuela, C. Ciuti, R. Fazio, and D. Rossini, Phys. Rev. B **98**, 241108(R) (2018).
[57] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, SIAM Rev. **59**, 65 (2017).
[58] F. Vicentini, PhilipVinc/NeuralQuantum.jl (GitHub Repository), this paper refers to v0.1 on Zenodo, https://doi.org/10.5281/zenodo.3238294 (2019).
[59] S. Krämer, D. Plankensteiner, L. Ostermann, and H. Ritsch, Comput. Phys. Commun. **227**, 109 (2018).
[60] J. Johansson, P. Nation, and F. Nori, Comput. Phys. Commun. **183**, 1760 (2012).
[61] J. Johansson, P. Nation, and F. Nori, Comput. Phys. Commun. **184**, 1234 (2013).
[62] M. J. Hartmann and G. Carleo, preceeding Letter, Phys. Rev. Lett. **122**, 250502 (2019).
[63] N. Yoshioka and R. Hamazaki, Phys. Rev. B **99**, 214306 (2019).
[64] A. Nagy and V. Savona, this issue, Phys. Rev. Lett. **122**, 250501 (2019).