Variational Quantum Monte Carlo Method with a Neural-Network Ansatz for Open Quantum Systems

Alexandra Nagy[1] and Vincenzo Savona[1]

[1]*Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland*

The possibility to simulate the properties of many-body open quantum systems with a large number of degrees of freedom is the premise to the solution of several outstanding problems in quantum science and quantum information. The challenge posed by this task lies in the complexity of the density matrix increasing exponentially with the system size. Here, we develop a variational method to efficiently simulate the non-equilibrium steady state of Markovian open quantum systems based on variational Monte Carlo and on a neural network representation of the density matrix. Thanks to the stochastic reconfiguration scheme, the application of the variational principle is translated into the actual integration of the quantum master equation. We test the effectiveness of the method by modeling the two-dimensional dissipative XYZ spin model on a lattice.

Open quantum systems have evolved into a major field of studies in recent years. Focus of these studies are the characterization of emergent phenomena and dissipative phase transitions [1–23], as well as the ongoing debate about whether quantum computing schemes are still hard to simulate classically – and thus achieve quantum supremacy – when in presence of some degree of noise-induced decoherence [24–27].

Assuming a Markovian interaction with the environment, the dynamics of open quantum systems is governed by the quantum master equation in Lindblad form [28]. Only few models within this description admit an analytical solution [29, 30]. The quest for efficient numerical methods to simulate the dynamics and the asymptotic steady state resulting from the Lindblad master equation, is a research field that is still in its infancy. Many recent tools have been developed following in the footsteps of well established numerical methods for the simulations of closed, Hamiltonian quantum systems. In particular, matrix-product state and tensor network schemes [31–34], a real-space renormalization approach [35], cluster mean-field [19], and other ad-hoc approximation schemes [23, 36] have recently emerged.

A groundbreaking progress in the numerical simulation of both the ground state and the dynamics of closed quantum systems has recently been made with the introduction of the neural-network variational ansatz [37–43], which efficiently represents highly correlated quantum states and whose parameters are easily optimized by means of the variational Monte Carlo (VMC) method. Recently, a self-adjoint and positive semidefinite parametrization of the density matrix, in terms of a neural network has been introduced [44].

The steady state of an open quantum system can be characterized by a variational principle [31, 33, 45, 46], whereby the dissipative part of the real-time dynamics, under quite general conditions, drives the system towards a unique steady state, in analogy to the imaginary-time Schrödinger equation that leads to the ground state of Hamiltonian systems.

In this Letter, we present a VMC approach to simulate the non-equilibrium steady state (NESS) of open quantum systems governed by the quantum master equation in Lindblad form. The density matrix is parametrized using a neural network ansatz [44] and parameters are varied using an extension of the stochastic reconfiguration method [47], which is shown to approximate the real-time dynamics of the system. We apply the present VMC to study the steady-state properties of the dissipative XYZ spin model [19–23], that displays a prototypical second-order dissipative phase transition. Thanks to the Monte-Carlo sampling of expectation values, this method holds promise for the efficient simulation of open quantum systems with a large number of degrees of freedom.

Dynamics of open quantum systems – The dynamics of the density matrix $\hat{\rho}$ of an open quantum system is governed by the quantum master equation which – in case of Markovian coupling to the environment – takes the Lindblad form

$$\frac{d\hat{\rho}}{dt} = -i[\hat{H}, \hat{\rho}] - \sum_i \frac{\gamma_i}{2} \left[\left\{ \hat{F}_i^\dagger \hat{F}_i, \hat{\rho} \right\} - 2\hat{F}_i \hat{\rho} \hat{F}_i^\dagger \right], \quad (1)$$

where the curly brackets denote the anti-commutator. The unitary part of the dynamics is generated by the term depending on the Hamiltonian $\hat{H}$, while $\hat{F}_i$ are the jump operators associated to the dissipative processes induced by the environment. The equation is typically expressed in terms of the Liouvillian superoperator as $d\hat{\rho}/dt = \mathcal{L}(\hat{\rho})$, whose formal solution is $\hat{\rho}(t) = e^{\mathcal{L}t}\hat{\rho}(0)$ ($t > 0$). The existence and uniqueness of a NESS – defined as $\hat{\rho}_{ss} = \lim_{t\to\infty} \hat{\rho}(t)$ – satisfying

$$\mathcal{L}(\hat{\rho}_{ss}) = 0, \quad (2)$$

has been demonstrated under quite general assumptions [48, 49], in particular for finite-size spin and boson lattices [48].

The steady state can be computed as the long-time limit of the solution of the quantum master equation, or by directly solving the homogeoeous linear system (2) with an additional condition on the trace of the density matrix. In both cases, the size of the problem is quantified by the square of the Hilbert space dimension, thus becoming computationally prohibitive already for a modest number of degrees of freedom. A promising route to the numerical computation of the NESS is provided by the variational principle. In cases where a unique steady state exists [48], the NESS corresponds to the eigen-matrix of the Liouvillian super-operator $\mathcal{L}$ with zero eigenvalue [49]. As all other eigenvalues have strictly negative real part, the NESS can be formally derived as the matrix that maximizes the real part of the expectation value (computed in matrix space) of the Liouvillian.

Neural Network Density Matrix – We assume that the Hilbert space of the system is spanned by the computational basis $|\boldsymbol{\sigma}\rangle$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ labels the states of N degrees of freedom that compose the system. Here and in what follows we will assume binary local degrees of freedom, with $\sigma_i = \{-1, 1\}$, which applies to the broad class of interacting spin-1/2 or qubit models. The density matrix in this basis is formally expressed as $\rho(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \langle \boldsymbol{\sigma}|\hat{\rho}|\boldsymbol{\eta}\rangle$ in terms of the density operator $\hat{\rho}$. We denote a specific variational ansatz for the density matrix as $\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})$, where $\chi = (\chi_1, \chi_2, \ldots, \chi_{N_p})$ is a set of variational parameters.

A neural network ansatz for a self-adjoint, positive semi-definite density matrix was recently introduced [44] in the specific form of a Restricted Boltzmann Machine (RBM). In a variational approach, RBMs present the significant advantage that the sum over the hidden-spin configurations can be carried out analytically, and the logarithmic derivatives with respect to the variational parameters admit simple expressions [39]. Here we briefly describe how this ansatz can be derived from simple considerations on the density matrix. A self-adjoint, positive semi-definite expression for the density matrix is

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \sum_{j=1}^{J} p_j(\chi) \cdot \psi_j(\boldsymbol{\sigma}, \chi)\psi_j^*(\boldsymbol{\eta}, \chi) \qquad (3)$$

The states $\psi_j(\boldsymbol{\sigma}, \chi)$ are not necessarily mutually orthogonal and the sum extends over J states, with $J \leq d$ and $d = 2^N$ is the dimension of the Hilbert space under study.

We start by introducing a RBM ansatz for each state $\psi_j(\boldsymbol{\sigma}, \chi)$ entering expression (3). A RBM is composed of two layers of binary valued nodes (see Fig. 1): a visible layer for encoding the physical state and a hidden layer. Each node is associated with a bias (a- and b-parameters) and nodes in the different layers are connected via a set of weighted edges (X-parameters). For a large number of

hidden nodes, this structure is known to describe quantum correlations efficiently [38, 41].

In order to express the mixed structure in eq. (3) as a single RBM, we embed an intermediate set of L hidden nodes that are used to express the probabilities $p_j(\chi)$ in RBM form as $p_j(\chi) = \exp(\sum_l c_l h_l)$, with $h_l = \pm 1$ and $c_l \in \mathbb{R}$. To index the different states in the mixture accordingly, this new set of hidden nodes must also enter the RBM expression of the wave functions. When carrying out the sums over configurations of hidden nodes, the final expression for the RBM density matrix is [50]

$$\begin{aligned}
\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = 8 \exp\left(\sum_i a_i \sigma_i\right) \exp\left(\sum_i a_i^* \eta_i\right) \\
\times \prod_{l=1}^{L} \cosh\left(c_l + \sum_i W_{li}\sigma_i + \sum_i W_{li}^*\eta_i\right) \\
\times \prod_{m=1}^{M} \cosh\left(b_m + \sum_i X_{mi}\sigma_i\right) \\
\times \prod_{n=1}^{M} \cosh\left(b_n^* + \sum_i X_{ni}^*\eta_i\right).
\end{aligned} \qquad (4)$$

The RBM is sketched in Fig. 1, and $\chi = \{a_i, b_m, X_{mi}, c_l, W_{li}\}$ is the final set of parameters, which are assumed as complex valued with the exception of c_l that must take real values. The representational power of the RBM is determined by the number of hidden nodes [53]. Here we set the densities of hidden nodes through the parameters $\alpha = M/N$, $\beta = L/N$, which measure the representational power of the RBM ansatz independently of the size of the spin lattice. When separately accounting for the real and imaginary parts of complex-valued parameters, the total number of computational parameters in the RBM ansatz is $N_p = N[(\alpha + \beta)(2N + 1) + \alpha + 2]$. In what follows, we will always assume $\alpha = \beta$ for simplicity.

Optimization – It is convenient to rewrite eq. (2) in a vectorized form by reshaping $\hat{\rho}$ into a column vector $|\rho\rangle$. Following [45], $\mathcal{L}$ takes matrix form and the steady state density matrix fulfills $\langle \rho|\mathcal{L}|\rho\rangle = 0$. Therefore the expectation value over the variational density matrix $\langle\langle \mathcal{L}_\chi \rangle\rangle = \langle \rho_\chi|\mathcal{L}|\rho_\chi\rangle/\langle \rho_\chi|\rho_\chi\rangle$ is a function of the variational parameters χ. The parameter values that best approximate $\langle\langle \mathcal{L}_\chi \rangle\rangle = 0$ can be found by means of various optimization procedures [39, 42, 44, 54]. In this Letter, we choose to adopt the Stochastic Reconfiguration (SR) scheme by Sorella et al. [47] which we extend to open quantum systems. The parameters are initialized to a small random value and, at each iteration, they are updated as

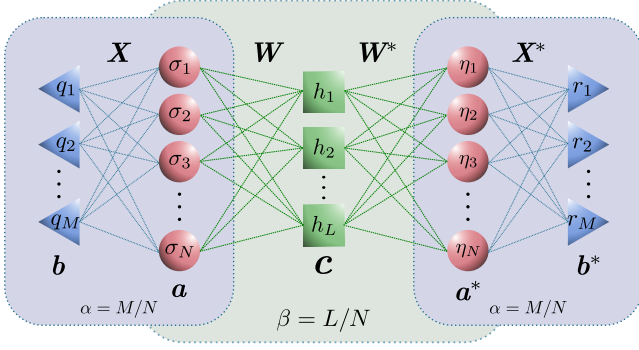$$\chi(n + 1) = \chi(n) + \nu \cdot S^{-1}(n)F(n), \qquad (5)$$

FIG. 1. Graphical representation of the neural network ansatz for the density matrix. The input states $|\boldsymbol{\sigma}\rangle, |\boldsymbol{\eta}\rangle$ are encoded in the visible layer, represented by circles. The hidden spins in the triangles encode the correlation between the physical spins in each state of the statistical mixture, while the hidden spins in the squares encode the mixture between the states. This structure is easily seen to coincide with a RBM, where the hidden layer is composed by the triangle and square nodes.
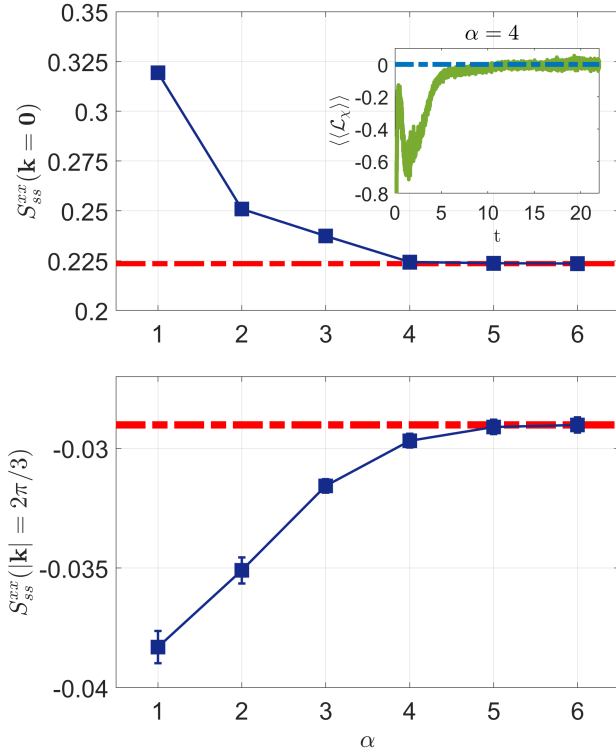

FIG. 2. The steady-state spin structure factor $S_{ss}^{xx}(\mathbf{k})$ computed as a function of $\alpha = \beta$ for a 3×3 lattice and $\mathbf{k} = \mathbf{0}$ (upper panel) and $\mathbf{k} = (2\pi/3, 0)$ (lower panel). The red dot-dashed line represents in both panels the exact result. The inset shows the evolution of $\langle\langle \mathcal{L}_\chi \rangle\rangle$ over the VMC run. Parameters: $J_x/\gamma = 0.9$, $J_y/\gamma = 1.2$, $J_z/\gamma = 1.0$.

where the learning rate ν is small enough to guarantee convergence. It can be shown [50] that the SR scheme

induces, at each iteration, a variation in the parameters that best approximates the time evolution of the density matrix over a time step ν. Here, we define the covariance matrix S, the vector of forces F, and the logarithmic derivatives O as

$$
\begin{aligned}
O_k(\boldsymbol{\sigma}, \boldsymbol{\eta}) &= \frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial \rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial \chi_k} \\
F_k(n) &= \langle\langle O_k^* \mathcal{L} \rangle\rangle - \langle\langle \mathcal{L} \rangle\rangle \langle\langle O_k^* \rangle\rangle \\
S_{kk'}(n) &= \langle\langle O_k^* O_{k'} \rangle\rangle - \langle\langle O_k^* \rangle\rangle \langle\langle O_{k'} \rangle\rangle,
\end{aligned}
\tag{6}
$$

where $k, k' = 1, 2, \ldots, N_p$. The notation $\langle\langle \cdot \rangle\rangle$ denotes the normalized expectation value taken over the variational density matrix $|\rho_\chi\rangle$, and the derivatives $O_k(\boldsymbol{\sigma}, \boldsymbol{\eta})$ are taken as diagonal operators in these expectation values. We point out that, while the expression for S in (6) results in the VMC iterations following the real time evolution, minimization can be achieved by using any positive-definite covariance matrix. In particular, setting S as the identity results in the steepest descent procedure. Since S can be non-invertible, we apply an explicit regularization scheme, as introduced in [39]: $S_{kk'}^{\text{reg}} = S_{kk'} + \lambda(n)\delta_{k,k'}S_{kk'}$, where $\lambda(n) = \max(\lambda_0 b^n, \lambda_{\min})$. For the present calculations, they were set to $\lambda_0 = 100$, $b = 0.998$ and $\lambda_{\min} = 10^{-2}$.

Sampling – The various expectation values in (6) must be evaluated at each iteration step. We evaluate these quantities stochastically over a Markov-chain of N_{MH} configurations $(\boldsymbol{\sigma}, \boldsymbol{\eta})$ sampling the square modulus of the density matrix $|\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})|^2$. For this we adopt the Metropolis-Hastings algorithm [55]. In the limit of $N_{MH} \to \infty$, the statistical error decays as $1/\sqrt{N_{MH}}$. Choosing an appropriate set of rules for the random walk is key to an efficient Monte Carlo sampling. Here we randomly choose each move among those allowed by the Liouvillian superoperator [50].

Observables – Once the optimal parameter values have been determined, the expectation value of any quantum mechanical observable $\hat{\mathcal{O}}$ over the steady state can be expressed as

$$
\langle \hat{\mathcal{O}} \rangle = \text{Tr}(\hat{\mathcal{O}}\hat{\rho}_\chi) = \sum_{\boldsymbol{\sigma}, \boldsymbol{\eta}} |\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})|^2 \cdot \frac{\mathcal{O}(\boldsymbol{\eta}, \boldsymbol{\sigma})}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})^*},
\tag{7}
$$

which can also be evaluated using the Metropolis-Hastings algorithm. For all the quantities considered here, the expectation values were additionally averaged over 100 sets of parameter values $\chi(n)$ chosen in the asymptotic region of the SR interation, in order to improve the statistical accuracy. The overall error in the sampled observables has, in addition to the contribution from the Metropolis-Hastings algorithms, a contribution from the SR scheme and a systematic
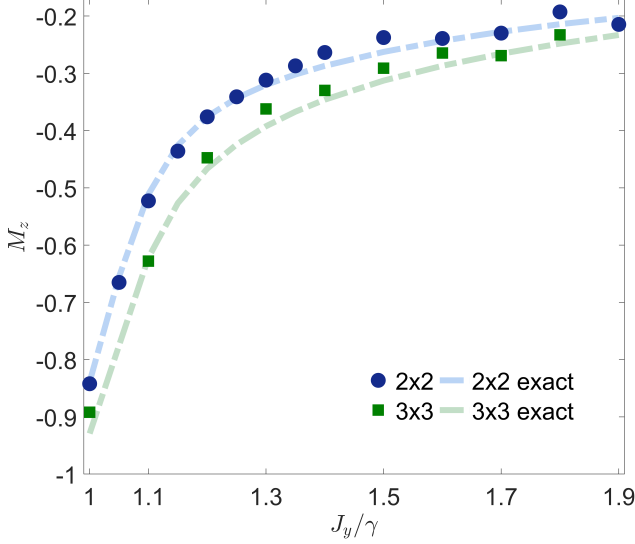
FIG. 3. The magnetization M_z computed as a function of the coupling J_y/γ. VMC and exact values are compared. Error bars, when not shown, are smaller than the symbol. Other parameters: $J_x/\gamma = 0.9$, $J_z/\gamma = 1.0$, $\alpha = \beta = 3$.

FIG. 4. The steady-state spin structure factor $S_{ss}^{xx}(\mathbf{k} = \mathbf{0})$ computed as a function of the coupling J_y/γ. VMC and exact values are compared. Other parameters: $J_x/\gamma = 0.9$, $J_z/\gamma = 1.0$, $\alpha = \beta = 3$.

contribution related to the representational power of the RBM ansatz, as measured by the α and β parameters.

Computational cost – The number of floating point operations to evaluate Eq. (6) scales as N_p^3, if we assume that the number of Metropolis-Hastings steps N_{MH} is set to roughly the number of parameters N_p, as in Ref. [39]. The Metropolis-Hastings procedure also scales with the number of connected states N_c, i.e. with the average number of nonzero elements in a column of the Liouvillian matrix. Finally, the efficiency of the whole procedure thus scales as $O(N_p^3 + N_p N_c)$.

Results – To assess the effectiveness of the method, we study a spin-1/2 XYZ model on a two-dimensional lattice with periodic boundary condition. Each spin is subject to a dissipation process into the $|\sigma^z = -1\rangle$ state. This model has been already widely investigated and is known to display a dissipative phase transition between a paramegnetic and a ferromagnetic phase [19–23]. The Hamiltonian and the quantum master equation read ($\hbar = 1$)

$$\hat{H} = \sum_{\langle i,j \rangle} \left(J_x \hat{\sigma}_i^x \hat{\sigma}_j^x + J_y \hat{\sigma}_i^y \hat{\sigma}_j^y + J_z \hat{\sigma}_i^z \hat{\sigma}_j^z \right) \quad (8)$$

$$\frac{d\hat{\rho}}{dt} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\{\hat{\sigma}_k^+ \hat{\sigma}_k^-, \hat{\rho}\} - 2\hat{\sigma}_k^- \hat{\rho}\hat{\sigma}_k^+ \right] \quad (9)$$

where $\hat{\sigma}_j^x$, $\hat{\sigma}_j^y$, $\hat{\sigma}_j^z$ are the Pauli matrices, $\hat{\sigma}_j^{\pm} = (\hat{\sigma}_j^x \pm i\hat{\sigma}_j^y)/2$, J_α are the coupling constants between nearest neighbour spins and γ is the dissipation rate. The excitations in the system – induced by the
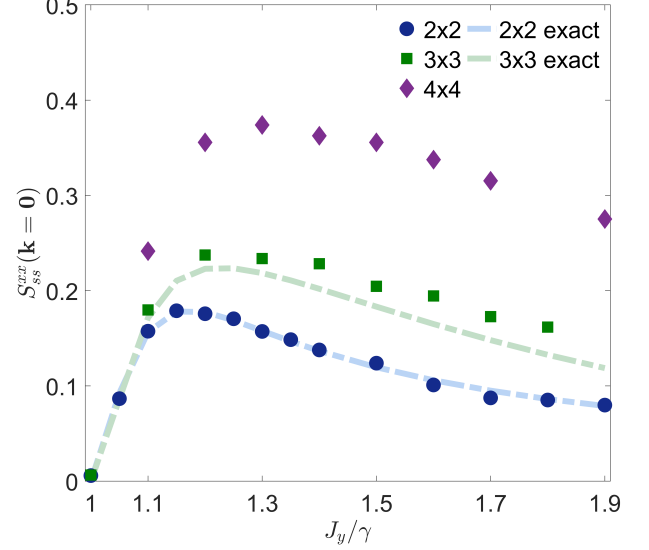
anisotropic spin coupling – compete with the isotropic dissipative process, and this competition is at the origin of the dissipative phase transition [19–23]. The effectiveness of the neural network ansatz is demonstrated by studying the system observables across a phase boundary.

In addition to the expectation value $\langle\langle \mathcal{L}_\chi \rangle\rangle$, we study the local magnetization

$$M_z = \frac{1}{N} \sum_{i=1}^{N} \text{Tr}(\hat{\rho}\hat{\sigma}_i^z), \quad (10)$$

and the steady-state structure factor

$$S_{ss}^{xx}(\mathbf{k}) = \frac{1}{N(N-1)} \sum_{\mathbf{j} \neq \mathbf{l}} e^{-i\mathbf{k}(\mathbf{j}-\mathbf{l})} \langle \hat{\sigma}_{\mathbf{j}}^x \hat{\sigma}_{\mathbf{l}}^x \rangle, \quad (11)$$

computed for the asymptotic steady state.

Fig. 2 shows the convergence of $S_{ss}^{xx}(\mathbf{k} = \mathbf{0})$ and $S_{ss}^{xx}(\mathbf{k} = (2\pi/3, 0))$ to the exact result for a 3×3 lattice, as $\alpha = \beta$ are increased. The parameters χ are initialized randomly and updated at each VMC step according to the SR scheme [50]. The parameters of the model are chosen to lie in the vicinity of the dissipative phase transition, i.e. $J_x/\gamma = 0.9$, $J_y/\gamma = 1.2$, $J_z/\gamma = 1.0$. A clear convergence towards the exact value upon increasing $\alpha = \beta$ is found. The inset in Fig. 2 shows the SR evolution of $\text{Re}(\langle\langle \mathcal{L}_\chi \rangle\rangle)$ over a typical VMC run. The oscillations at early times are a feature of the unitary part of the dynamics in the quantum master equation.

In Fig. 3 we display the magnetization as computed for different lattice sizes and as a function of the coupling parameter J_y/γ. For this choice of parameters, a para-to-ferromagnetic phase transition is expected to occur when increasing the coupling through the value $J_y \gtrsim 1.04$ [21, 22], while a second phase boundary between a ferromagnetic and a paramagnetic region has been predicted by cluster mean-field calculations at around $J_y \gtrsim 1.4$. For 2×2 and 3×3 lattices the VMC result agrees well with the exact calculation for a large enough number of variational parameters.

In Fig. 4 we display the spin structure factor $S_{ss}^{xx}(\mathbf{k} = \mathbf{0})$ for the same parameters as in Fig. 3. The quantity $S_{ss}^{xx}(\mathbf{k} = \mathbf{0})$ vanishes when in a paramagnetic phase, while it takes a finite value in the ferromagnetic region of the phase diagram. This behaviour is displayed both by the exact calculation for small lattices, and by the VMC data, in the vicinity of the phase boundary at $J_y \gtrsim 1.04$. For values $J_y > 1.4$ the system should become again paramagnetic in the thermodynamic limit of large lattices, but this feature was not displayed by the present data up to the largest lattice under study, in agreement also with recent stochastic Gutzwiller calculations [23].

In all the calculations, special care was devoted to the choice of the SR time step ν. The unitary part of the real-time dynamics generated by Eq. (1) makes the differential equation stiff, thus requiring to scale down ν appropriately as the system size – and thus the spectral width of the differential operator – is increased. A possible workaround would be to study an effective, purely dissipative dynamics using the super-operator $\mathcal{L}^\dagger \mathcal{L}$ as a generator. In the case of a unique steady state, this super-operator is self-adjoint and positive semi-defined, with the only null eigenvalue being associated to the steady-state solution. We argue that this effective dynamics would be more robust to the choice of the time step. The super-operator $\mathcal{L}^\dagger \mathcal{L}$ is however less sparse than $\mathcal{L}$ on the computational basis, calling for an efficient sampling scheme.

Existing numerical approaches to the simulation of the steady state of a Markovian open quantum system either require the full representation of the Hilbert space into memory, or rely on a properly chosen truncation of the Hilbert space to a relevant subspace. The present VMC approach is free of these two limitations, thanks to the stochastic evaluation of expectation values by means of the Metropolis-Hastings algorithm. The neural network ansatz in terms of a RBM is highly representative of quantum correlated statistical mixtures, while being simple to handle numerically. In cases with very strong quantum correlations, this ansatz could be extended to deep network representations, as was recently done in the case of Hamiltonian problems [37, 38, 43, 56]. For some of these networks [56], the hidden degrees of freedom can still be summed analytically, as for RBMs. Neural network representations are not restricted to spin degrees of freedom and have been successfully adopted to represent bosonic many-body states efficiently [57]. For these reasons, the present VMC approach may emerge as the election tool to numerically model open quantum systems, with considerable impact on the study of fundamental physics and on the modeling of near-term, noisy quantum information platforms [27].

While developing the present result we became aware of three related independent works that have been carried out in parallel [58–60].

[1] I. Carusotto and C. Ciuti, Reviews of Modern Physics **85**, 299 (2013).
[2] M. J. Hartmann, Journal of Optics **18**, 104005 (2016).
[3] C. Noh and D. G. Angelakis, Reports on Progress in Physics **80**, 016401 (2017).
[4] N. Bartolo, F. Minganti, W. Casteels, and C. Ciuti, Phys. Rev. A **94**, 033841 (2016).
[5] A. Biella, F. Storme, J. Lebreuilly, D. Rossini, R. Fazio, I. Carusotto, and C. Ciuti, Phys. Rev. A **96**, 023839 (2017).
[6] M. Biondi, G. Blatter, H. E. Türeci, and S. Schmidt, Phys. Rev. A **96**, 043809 (2017).
[7] H. J. Carmichael, Phys. Rev. X **5**, 031028 (2015).
[8] W. Casteels, R. Fazio, and C. Ciuti, Phys. Rev. A **95**, 012128 (2017).
[9] W. Casteels, F. Storme, A. Le Boité, and C. Ciuti, Phys. Rev. A **93**, 033824 (2016).
[10] J. M. Fink, A. Dombi, A. Vukics, A. Wallraff, and P. Domokos, Phys. Rev. X **7**, 011012 (2017).
[11] T. Fink, A. Schade, S. Höfling, C. Schneider, and A. Imamoglu, Nature Physics **14**, 365 (2018).
[12] M. Fitzpatrick, N. M. Sundaresan, A. C. Y. Li, J. Koch, and A. A. Houck, Phys. Rev. X **7**, 011016 (2017).
[13] 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).
[14] E. M. Kessler, G. Giedke, A. Imamoglu, S. F. Yelin, M. D. Lukin, and J. I. Cirac, Phys. Rev. A **86**, 012116 (2012).
[15] J. Marino and S. Diehl, Phys. Rev. Lett. **116**, 070407 (2016).
[16] V. Savona, Phys. Rev. A **96**, 033826 (2017).
[17] L. M. Sieberer, S. D. Huber, E. Altman, and S. Diehl, Phys. Rev. Lett. **110**, 195301 (2013).
[18] F. Vicentini, F. Minganti, R. Rota, G. Orso, and C. Ciuti, Phys. Rev. A **97**, 013853 (2018).
[19] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio, and D. Rossini, Physical Review X **6**, 031011 (2016).
[20] T. E. Lee, S. Gopalakrishnan, and M. D. Lukin, Physical Review Letters **110**, 257204 (2013).
[21] R. Rota, F. Minganti, A. Biella, and C. Ciuti, New Journal of Physics **20**, 045003 (2018).
[22] R. Rota, F. Storme, N. Bartolo, R. Fazio, and C. Ciuti,

Physical Review B **95**, 134431 (2017).

[23] W. Casteels, R. M. Wilson, and M. Wouters, Physical Review A **97**, 062107 (2018).

[24] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Quantum **1**, 8 (2016), arXiv:1610.01808.

[25] X. Gao and L. Duan, (2018), arXiv:1810.03176.

[26] A. W. Harrow and A. Montanaro, Nature **549**, 203 (2017).

[27] J. Preskill, Quantum **2**, 79 (2018), arXiv:1801.00862.

[28] H.-P. Breuer and F. F. Petruccione, *The theory of open quantum systems* (Oxford University Press, 2002) p. 625.

[29] T. c. v. Prosen, Phys. Rev. Lett. **107**, 137201 (2011).

[30] T. c. v. Prosen, Phys. Rev. Lett. **112**, 030603 (2014).

[31] J. Cui, J. I. Cirac, and M. C. Bañuls, Physical Review Letters **114**, 220601 (2015).

[32] A. Kshetrimayum, H. Weimer, and R. Orús, Nature Communications **8**, 1291 (2017).

[33] E. Mascarenhas, H. Flayac, and V. Savona, Physical Review A - Atomic, Molecular, and Optical Physics **92**, 56 (2015), arXiv:1504.06127.

[34] A. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Physical Review Letters **116**, 237201 (2016).

[35] S. Finazzi, A. Le Boité, F. Storme, A. Baksic, and C. Ciuti, Physical Review Letters **115**, 080604 (2015).

[36] A. Nagy and V. Savona, Phys. Rev. A **97**, 052129 (2018).

[37] Z. Cai and J. Liu, Physical Review B **97**, 035116 (2018).

[38] G. Carleo, Y. Nomura, and M. Imada, Nature Communications **9**, 5322 (2018).

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

[40] N. Freitas, G. Morigi, and V. Dunjko, International Journal of Quantum Information **16**, 1840008 (2018).

[41] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Physical Review X **8**, 011006 (2018).

[42] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017).

[43] X. Gao and L.-M. Duan, Nature Communications **8**, 662 (2017).

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

[45] M. Jakob and S. Stenholm, Phys. Rev. A **67**, 032111 (2003).

[46] H. Weimer, Physical Review Letters **114**, 040402 (2015).

[47] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys **127**, 014105 (2007).

[48] D. Nigro, (2018), arXiv:1803.06279.

[49] F. Minganti, A. Biella, N. Bartolo, and C. Ciuti, Physical Review A **98**, 042118 (2018).

[50] See Supplemental Material for more information on the neural network ansatz, the stochastic reconfiguration method, the stochastic sampling procedure and the details of the numerical approach, which includes Refs. [39, 47, 51, 52, 55].

[51] S. Choi, C. Paige, and M. Saunders, SIAM J. Sci. Comput. **33**, 1810 (2011).

[52] Y. Liu, "Contribute to syangliu/MINRES-QLP development by creating an account on GitHub," (2019), original-date: 2018-06-08T04:18:21Z .

[53] N. Le Roux and Y. Bengio, Neural Computation **20**, 1631 (2008), https://doi.org/10.1162/neco.2008.04-07-510.

[54] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nature Physics **14**, 447 (2018).

[55] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys **21**, 1087 (1953).

[56] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Phys. Rev. Lett. **121**, 167204 (2018).

[57] H. Saito, Journal of the Physical Society of Japan **86**, 093001 (2017).

[58] M. J. Hartmann and G. Carleo, arXiv:1902.05131 [cond-mat, physics:quant-ph] (2019), arXiv: 1902.05131.

[59] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, arXiv:1902.10104 [cond-mat, physics:quant-ph] (2019), arXiv: 1902.10104.

[60] N. Yoshioka and R. Hamazaki, arXiv:1902.07006 [cond-mat, physics:quant-ph] (2019), arXiv: 1902.07006.

SUPPLEMENTAL MATERIAL

NEURAL NETWORK REPRESENTATION OF THE DENSITY MATRIX

In the main text we approximate the density matrix as a mapping ρ_χ with parameters χ which, given two input states $|\boldsymbol{\sigma}\rangle$ and $|\boldsymbol{\eta}\rangle$, returns the matrix element $\langle \boldsymbol{\sigma}|\hat{\rho}_\chi|\boldsymbol{\eta}\rangle = \rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})$. We describe quantum systems with N degrees of freedom on a computational basis $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ assuming two-dimensional local Hilbert-spaces $\sigma_i = \{-1, 1\}$. This choice generally applies to spin-1/2 and quantum bits. A general expression, ensuring the positive semi-definite property of the density matrix, is

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \sum_{j=1}^{J} p_j(\chi) \cdot \psi_j(\boldsymbol{\sigma}, \chi) \psi_j^*(\boldsymbol{\eta}, \chi) \tag{12}$$

where χ are the parameters. The states $\psi_j(\boldsymbol{\sigma}, \chi)$ are not necessarily mutually orthogonal and, together with the associated probabilities $1 \geq p_j(\chi) \geq 0$, are an expression of the statistical mixture described by $\hat{\rho}_\chi$. Here, the sum extends over J states, with $J \leq d$ and $d = 2^N$ is the dimension of the Hilbert-space under study.

Following Ref. [39], we may represent any wave function $\psi(\boldsymbol{\sigma})$ as a Restricted Boltzmann Machine (RBM),

$$\psi(\boldsymbol{\sigma}) = \sum_{\{q\}} \exp \left(\sum_i a_i \sigma_i + \sum_m b_m q_m + \sum_{m,i} q_m \sigma_i X_{mi} \right) \tag{13}$$

where $q_m = \pm 1$ are a set of $M = \alpha \times N$ hidden spin variables and the leftmost sum runs over all possible hidden spin configurations $\{q\}$. The coupling and bias parameters $\{a_i, b_m, X_{mi}\}$ are, in general, complex-valued. In order to express the mixed state (12) in terms of a single RBM, we introduce a second set of $L\beta \times N$ hidden spins that we use to represent the probabilities $p_j(\chi)$ in RBM form as $p_j(\chi) = \exp(\sum_l c_l h_l)$, with $h_l = \pm 1$ and $c_l \in \mathbb{R}$. This new set of hidden nodes should also enter in the expression of the wave functions, in order to index the different wave functions that make the mixed state. Hence, the full RBM form of the density matrix that we propose is

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \sum_{\{h\}} \sum_{\{q\}} \sum_{\{r\}} \exp \left(\sum_l c_l h_l \right)$$

$$\times \exp \left(\sum_i a_i \sigma_i + \sum_m b_m q_m + \sum_{m,i} q_m \sigma_i X_{mi} + \sum_{l,i} h_l \sigma_i W_{li} \right) \tag{14}$$

$$\times \exp \left(\sum_i a_i^* \eta_i + \sum_n b_n^* r_n + \sum_{n,i} r_n \eta_i X_{ni}^* + \sum_{l,i} h_l \eta_i W_{li}^* \right) .$$

Since no intra-layer connection is allowed, the hidden variables can be explicitly traced out and the neural network density matrix reads as

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = 8 \exp \left(\sum_i a_i \sigma_i \right) \exp \left(\sum_i a_i^* \eta_i \right)$$

$$\times \prod_{l=1}^{L} \cosh \left(c_l + \sum_i W_{li} \sigma_i + \sum_i W_{li}^* \eta_i \right)$$

$$\times \prod_{m=1}^{M} \cosh \left(b_m + \sum_i X_{mi} \sigma_i \right) \tag{15}$$

$$\times \prod_{n=1}^{M} \cosh \left(b_n^* + \sum_i X_{ni}^* \eta_i \right) .$$

The variational parameters in this RBM are $\chi = \{a_i, b_m, X_{mi}, c_l, W_{li}\}$ and, with the exception of the c_l's, are all complex valued. When separately accounting for the real and imaginary parts of each parameter, the total number

of real-valued computational parameters in the RBM ansatz is $N_p = N[(\alpha + \beta)(2N + 1) + \alpha + 2]$. We introduce the so called effective angles [39]

$$
\tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta}) = c_l + \sum_i W_{li}\sigma_i + \sum_i W_{li}^*\eta_i
$$

$$
\theta_m(\boldsymbol{\sigma}) = b_m + \sum_i X_{mi}\sigma_i \,,
$$

(16)

finally obtaining

$$
\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = 8\exp\left(\sum_i a_i\sigma_i\right)\exp\left(\sum_i a_i^*\eta_i\right)\prod_{l=1}^{L}\prod_{m=1}^{M}\prod_{n=1}^{M}\cosh\tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta})\cosh\theta_m(\boldsymbol{\sigma})\cosh\theta_n^*(\boldsymbol{\eta})\,.
$$

(17)

In order to find the best set of parameters to describe the steady state, we choose to use the Stochastic Reconfiguration (SR) algorithm. This requires the calculation of the logarithmic derivatives of $\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})$ with respect to the real and imaginary parts of all variational parameters. Given expression (17), these are written in a rather compact form as

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Re(a_k)} = \sigma_k + \eta_k
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Im(a_k)} = i(\sigma_k - \eta_k)
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Re(b_k)} = \tanh\left(\theta_k(\boldsymbol{\sigma})\right) + \tanh\left(\theta_k^*(\boldsymbol{\eta})\right)
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Im(b_k)} = i\left[\tanh\left(\theta_k(\boldsymbol{\sigma})\right) - \tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\right]
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Re(X_{kl})} = \tanh\left(\theta_k(\boldsymbol{\sigma})\right)\sigma_l + \tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\eta_l
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Im(X_{kl})} = i\left[\tanh\left(\theta_k(\boldsymbol{\sigma})\right)\sigma_l - \tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\eta_l\right]
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial c_k} = \tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta})\right)
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Re(W_{kl})} = \tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta})\right)(\sigma_l + \eta_l)
$$

$$
\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial\Im(W_{kl})} = i\tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta})\right)(\sigma_l - \eta_l)
$$

(18)

STOCHASTIC RECONFIGURATION FOR OPEN QUANTUM SYSTEMS

We approximate the density matrix with a neural network ansatz ρ_χ, in order to find the best representation of the steady state. In the following, we use the vectorized formalism of the Liouville-von-Neumann master equation by reshaping $\hat{\rho}$ into a column vector $|\rho\rangle$ and the Liouvillian superoperator into a matrix $\mathcal{L}$ of dimension $d^2 \times d^2$ where d is the dimension of the Hilbert-space under study. In order to find the optimal set of parameters, we choose to extend the SR method [47] to open quantum systems.

In the SR optimization the variational parameters are changed at each iteration step by

$$
\chi'_k = \chi_k + \delta\chi_k \,.
$$

(19)

After a small perturbation in linear approximation the density matrix reads as

$$
|\rho_{\chi'}\rangle \simeq \sum_k \delta\chi_k O_k|\rho_\chi\rangle\,,
$$

(20)

where the diagonal operators O_k are defined for any configuration $|\boldsymbol{\sigma}, \boldsymbol{\eta}\rangle = |x\rangle$ as the variational derivative with respect to the k-th parameter

$$O_k(x) = O_k(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial \rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial \chi_k} \,. \tag{21}$$

Open quantum systems evolve under a one-parameters semigroup dictated by the Liouvillian superoperator as

$$|\bar{\rho}_\chi\rangle = e^{\mathcal{L}t}|\rho_\chi\rangle \simeq (\mathbb{1} + \mathcal{L})|\rho_\chi\rangle \,. \tag{22}$$

Under general assumptions, this dynamics asymptotically converges to the non-equilibrium steady state. We can enforce the real-time dynamics generated by the Liouvillian within the SR scheme. The method can be then interpreted as a real time evolution in the variational subspace.

To achieve this goal, we equate eq. (20) and (22) in the subspace spanned by the vectors $\{O_k|\rho_\chi\rangle\}$, thus obtaining

$$\delta\chi_k = \nu \sum_{k'} S_{kk'}^{-1} F_{k'} \,, \tag{23}$$

where ν is small enough to guarantee convergence and we introduced the generalized forces and the covariance matrix for a given set of variational parameters at iteration step n

$$\begin{aligned}
F_k(n) &= \langle\langle O_k^* \mathcal{L} \rangle\rangle - \langle\langle \mathcal{L} \rangle\rangle \langle\langle O_k^* \rangle\rangle \\
S_{kk'}(n) &= \langle\langle O_k^* O_{k'} \rangle\rangle - \langle\langle O_k^* \rangle\rangle \langle\langle O_{k'}^* \rangle\rangle \,.
\end{aligned} \tag{24}$$

Here, $\langle\langle \cdot \rangle\rangle$ denotes the normalized expectational value taken over the variational density matrix $|\rho_\chi\rangle$. The covariance matrix can be any positive-definite matrix. Choosing it as the identity will correspond to the steepest descent optimization method. Since S can be non-invertible we apply an explicit regularization scheme, as introduced in [39]

$$S_{kk'}^{\text{reg}} = S_{kk'} + \lambda(n)\delta_{k,k'}S_{kk'} \,. \tag{25}$$

Here, the $\lambda(n)$ parameter is a function of the iteration step n and decays with $\lambda(n) = \max(\lambda_0 b^n, \lambda_{\min})$. For the present calculations, they were set to $\lambda_0 = 100$, $b = 0.998$ and $\lambda_{\min} = 10^{-2}$.

STOCHASTIC SAMPLING

For the SR optimization scheme, one must evaluate (24) at each iteration step. In order to perform a stochastic sampling of these quantities, we write the general forces and the covariance matrix into a suitable form as follows

$$\begin{aligned}
F_k(n) &= \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \frac{\langle x|\mathcal{L}|\rho_\chi\rangle}{\langle x|\rho_\chi\rangle} \\
&\quad - \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \frac{\langle x|\mathcal{L}|\rho_\chi\rangle}{\langle x|\rho_\chi\rangle} \sum_{x'} |\langle x'|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x'|\rho_\chi\rangle}{\partial \chi_k}\right)^* \\
S_{kk'}(n) &= \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_{k'}}\right) \\
&\quad - \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \sum_{x'} |\langle x'|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x'|\rho_\chi\rangle}{\partial \chi_{k'}}\right) \,,
\end{aligned} \tag{26}$$

where $|x\rangle = |\boldsymbol{\sigma}, \boldsymbol{\eta}\rangle$ are the basis elements spanning the space of the vectorized density matrix. Therefore, for a given χ we can generate a Markov-chain of N_{MH} many-body configurations $|x_1\rangle \to |x_2\rangle \to \cdots \to |x_{N_{MH}}\rangle$ sampling the square modulus of the density matrix elements $|\langle x|\rho_\chi\rangle|^2 = |\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})|^2$ via a standard Metropolis-Hastings algorithm [55]. The configurations are generated based on a proposition scheme, and accepted according to the probability

$$\mathcal{A}(|x_y\rangle \to |x_z\rangle) = \min\left(1, \frac{|\langle x_z|\rho_\chi\rangle|^2 \cdot \mathcal{P}(y|z)}{|\langle x_y|\rho_\chi\rangle|^2 \cdot \mathcal{P}(z|y)}\right) \,. \tag{27}$$

where $\mathcal{P}(z|y)$ is the conditional probability of proposing a state z given y. Our move generation is based on the transitions dictated by the Liouvillian superoperator which occur by applying $\mathcal{L}|x\rangle$. Doing so decreases the required number of thermalization and sampling steps by promoting, in the random walk, the configurations with the highest importance. However, in case of a close-to-pure steady-state, a handful of density matrix elements have magnitude dominating the others. This causes the random walk to perform poorly, as some regions in configuration space become scarcely accessible. Hence, we introduced additional moves which occur with a low probability and reach a configuration by randomly flipping one spin in both $\boldsymbol{\sigma}$ and $\boldsymbol{\eta}$. The propositions always need to obey detailed balance. For sampling $|\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})|^2$ in the case of a two-dimensional spin lattice model, we introduce the following possible moves on the configuration space spanned by $|\boldsymbol{\sigma},\boldsymbol{\eta}\rangle$

1. *Column-hopping in $\boldsymbol{\sigma}$*: a site σ_j and its right neighbour are flipped. This move accounts for single spin hopping as well as double excitation or loss.

2. *Row-hopping in $\boldsymbol{\sigma}$*: a site σ_j and its down neighbour are flipped.

3. *Column-hopping in $\boldsymbol{\eta}$*: a site η_j and its right neighbour are flipped.

4. *Row-hopping in $\boldsymbol{\eta}$*: a site η_j and its down neighbour are flipped.

5. *Excitation in $\boldsymbol{\sigma}$*: a site σ_j is flipped. This move corresponds to the effect of having an external field in the direction perpendicular to the quantization axis.

6. *Excitation in $\boldsymbol{\eta}$*: a site η_j is flipped.

7. *Dissipator*: sites $\sigma_j = \eta_j$ are flipped with an asymmetrical transition ratio. The dissipative moves are always proposed, while excitation are only proposed with ten-percent probability.

8. *Jumper:* two randomly chosen spins σ_l, η_m are flipped.

For the move proposition we choose one of the actions with uniform probability.

DETAIL ON NUMERICAL APPROACH

The VMC code was written in Python. The Metropolis Hastings algorithm was efficiently parallelized by splitting the Markov chain into several independent chains that were run in parallel using MVAPICH2. Both a CPU and a GPU version of the code were developed. GPU calculations allowed to significantly speed up the update of the O_k expectation values, as well as the iterative solution of the linear system in the SR, using the MINRES-QLP algorithm [51, 52] which can correctly handle the case of a singular matrix. The GPU version brought considerable advantage over CPU when the number of variational parameters was larger than 1000. In the Metropolis-Hastings sampling we set the number of *accepted* moves instead of the total number of moves. This is necessary as, for some choices of the physical parameters, the density matrix is highly concentrated on a few matrix elements and consequently MH moves are seldom accepted. The number of accepted moves was set to be roughly twice the number of variational parameters in all VMC runs.