# Approximating power of machine-learning ansatz for quantum many-body states

Artem Borin and Dmitry A. Abanin
*Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland*

(Received 16 May 2019; revised manuscript received 29 February 2020; accepted 3 March 2020; published 26 May 2020)

An artificial neural network (ANN) with the restricted Boltzmann machine (RBM) architecture was recently proposed as a versatile variational quantum many-body wave function. In this work we provide physical insights into the performance of this ansatz. We uncover the connection between the structure of the RBM and perturbation series, which explains the excellent precision achieved by the RBM ansatz in certain simple models, demonstrated in G. Carleo and M. Troyer [Science **355**, 602 (2017)]. Based on this relation, we improve the numerical algorithm to achieve a better performance of the RBM in cases where local minima complicate the convergence to the global one. Furthermore, we study the performance of a sparse RBM for approximating ground states of random, translationally invariant models in one dimension, as well as random matrix-product states. We find that the error in approximating such states exhibits a broad distribution and shows a positive correlation with the entanglement properties of the targeted state.

DOI: 10.1103/PhysRevB.101.195141

## I. INTRODUCTION

Variational methods play an invaluable role in quantum many-body physics because they allow one to represent exponentially many amplitudes of a many-body wave function using a small number of variational parameters. The choice of the variational ansatz is often motivated by the underlying physics of the system of interest, with notable examples being product states, the BCS wave function [1], and Laughlin states [2]. Broad classes of variational wave functions, such as tensor networks [3–5], which include matrix product states (MPSs) [4], rely on the low amount of quantum entanglement in ground states of physical systems.

Even though tensor network methods proved remarkably successful [5–9], it is important to investigate other classes of variational wave functions, including the ones that can capture states with higher entanglement. Recent proposals [10–20] for such ansätze considered variational wave functions inspired by machine learning (ML). More generally, ML is finding an increasing number of diverse applications in physics, including the detection of phase transitions [21–30], extraction of relevant degrees of freedom [31,32], and improvement of existing techniques [33–35]. In this paper we consider the variational ansatz inspired by the restricted Boltzmann machine (RBM) [10]. The architecture of the RBM—a neural network with a wide range of applications outside quantum physics [36–38]—is illustrated in Fig. 1.

The RBM ansatz for the system of $N$ physical spins $\{\sigma_i^z\} = \pm 1$ is constructed by introducing $M$ hidden spins $\{s_i\} = \pm 1$, which gives $\alpha = M/N$ hidden spins per each physical spin. Then, the amplitudes of the variational wave function in the $\{\sigma_i^z\}$ eigenbasis are obtained by summing over the states of hidden degrees of freedom:

$$\Psi_{\mathrm{RBM}}(\{\sigma_j\}) = \sum_{\{s_i=\pm 1\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i s_i + \sum_{ij} W_{ij} s_i \sigma_j^z}. \qquad (1)$$

This wave function is optimized over the variational parameters $\{a_j\}$, $\{b_i\}$, and $\{W_{ij}\}$ to yield the lowest-energy wave function for a given Hamiltonian.

The summation over the hidden-spin states in Eq. (1) can be performed explicitly, giving up to a normalization factor:

$$\Psi_{\mathrm{RBM}}(\{\sigma_j\}) \propto \exp\left(\sum_j a_j \sigma_j^z\right) \prod_{i=1}^{M} S_i,$$

$$S_i = \cosh\left(b_i + \sum_j W_{ij} \sigma_j^z\right). \qquad (2)$$

This property facilitates Monte Carlo sampling since any quantum amplitude can be computed by evaluating the function (2).

The utility of the RBM ansatz (2) is being actively explored [39–46]. Among its advantages are applicability in any number of spatial dimensions and the ability to describe states with a high degree of entanglement. It was found to approximate the ground states (GSs) of certain Hamiltonians [one-dimensional (1D) Ising and 1D and two-dimensional (2D) Heisenberg] [10] with a remarkably high precision, given relatively few variational parameters. Moreover, it can exactly represent certain topological states (1D cluster state and 2D and three-dimensional toric codes) [40].

An important question regarding the representation of tensor-network states by the RBM ansatz is intimately related to the entanglement properties of the latter. The entanglement entropy of a general RBM state obeys volume-law scaling [39], but finite-range RBM states (that is, states in which a hidden spin can have nonzero interaction $W_{ij}$ only with $d$ contiguous physical spins) obey the area law in any number of dimensions [39] and can be represented [41] by an MPS with a finite bond dimension $D$. However, some MPSs, the Affleck-Kennedy-Lieb-Tasaki (AKLT) state being an example, cannot

 

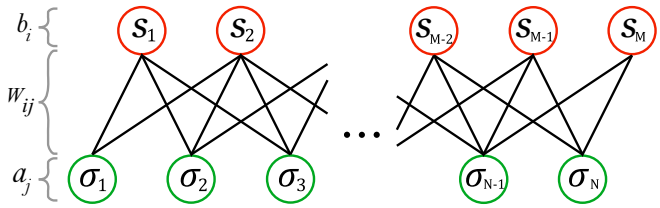


FIG. 1. Illustration of the RBM architecture: $\{\sigma_j\}$ denote physical spins, and $\{s_i\}$ are the hidden spins. The variational parameters $\{a_j\}$ and $\{b_i\}$ represent biases applied to physical and hidden spins, and parameters $\{W_{ij}\}$ correspond to couplings between the two flavors of spins.

be approximated by a finite-range RBM [41]. The infinite-range RBM can approximate any MPS arbitrarily well, but the number of required hidden units is exponential in the bond dimension of MPSs [47], which is impractical for numerical applications. Thus, it is important to understand how well sparse RBM states, with a relatively low density of hidden spins $\alpha$, which can be efficiently optimized, can approximate generic MPSs.

Here, we provide several results regarding the performance and representation power of the RBM ansatz. First, by connecting the RBM to perturbation theory, we explain the surprisingly good performance of the RBM with few variational parameters for certain models, as reported in Ref. [10]. To that end, we show that the RBM with $\alpha = 1$ already captures several orders of the perturbation series for those models. The connection to perturbation theory naturally suggests an improvement for the optimization algorithm, which we explore. Second, we investigate the performance of the RBM for *general* local Hamiltonians, picked at random, as well as its ability to approximate random MPSs in the practically interesting case of $\alpha = 1$. By applying the ansatz to various random realizations of these systems, we show that the performance of the RBM is positively correlated with the entanglement entropy and range of correlations.

## II. RBM AND PERTURBATION THEORY

To illustrate the connection of the structure of the RBM to perturbation series we first focus on a solvable, 1D transverse-field Ising (TFI) model:

$$H_{\mathrm{TFI}} = -\sum_i \sigma_i^x - J\sum_i \sigma_i^z\sigma_{i+1}^z, \tag{3}$$

where $\sigma_i^\beta$, $\beta = x, y, z$ are the standard Pauli operators.

We consider two points in the phase diagram: $J = 0.5$ (paramagnetic phase) and $J = 2$ (ferromagnetic phase). The GS wave function at these points can be presented in the following form, using perturbation theory:

$$|\psi\rangle = \frac{e^{\hat{W}}|\psi_0\rangle}{\langle\psi_0|e^{\hat{W}^\dagger}e^{\hat{W}}|\psi_0\rangle^{\frac{1}{2}}}, \tag{4}$$

where $|\psi_0\rangle$ is the unperturbed wave function, polarized along the $x$ axis and the $z$ axis for the paramagnetic and ferromagnetic cases, respectively. The $n$th term of the *perturbing operator* $\hat{W} = \sum_n \hat{W}_n$ accounts for the $n$th-order perturbative correction to the wave function. The terms $\hat{W}_n$ become less and less local as $n$ is increased. The explicit form of the operator $\hat{W}$ is obtained by iteratively solving the Schrödinger equation

$$(He^{\hat{W}} - e^{\hat{W}}E)|\psi_0\rangle = 0. \tag{5}$$

The RBM wave function (2) can be rewritten in a form similar to Eq. (4):

$$\Psi_{\mathrm{RBM}}(\{\sigma_j\}) \propto e^{\tilde{W}}|\tilde{\psi}_0\rangle, \tag{6}$$

where we introduce the *RBM operator* $\tilde{W} = \sum_j a_j\sigma_j^z + \sum_{i=1}^M \ln S_i$ and the RBM reference state $|\tilde{\psi}_0\rangle$, which corresponds to the state polarized along the $x$ axis.

The case of the RBM reference state coinciding with the unperturbed GS $|\psi_0\rangle = |\tilde{\psi}_0\rangle$ is the simplest example, where the RBM ansatz can approximate the perturbative wave function (4). For such an unperturbed GS the action of the perturbation can result only in spin flips. Thus, the perturbing operator $\hat{W}$ can be expressed solely in terms of $\{\sigma_i^z\}$, which makes it possible to establish the equivalence of the perturbing and RBM operators $\hat{W} = \tilde{W}$. We can see this equivalence by rewriting the expression for the RBM (2). Given a finite range $d$ of the RBM weights $W_{ij}$, the cosine terms $S_i$ of the RBM wave function can be written as

$$S_i = \cosh b_i \left(\prod_{j=i}^{i+d-1} \cosh W_{ij}\right) \times \sum_{\substack{\{n_k=\{0,1\}\}_{k=0}^d \\ (\sum_k n_k)\bmod 2 = 0}} \tanh^{n_d} b \prod_{j=i}^{i+d-1} (\tanh W_{ij}\sigma_j)^{n_{j-i}}. \tag{7}$$

We see that a cosine $S_i$ and, consequently, its logarithm in the RBM operator contain all possible spin-flip terms at the range $d$. At the same time we can always limit the perturbative expansion of the perturbing operator $\hat{W}$ to the finite order in such a way that the support of each individual term in the expansion does not exceed $d$. Then, the perturbing operator contains no more than $N_d = 2^{d-1}$ terms per spin, which is the number of all possible distinct spin-flip operators with a given support $d$. Further, the RBM ansatz with a hidden-unit density $\alpha$ has $\alpha(d+1)$ parameters per spin coming from the cosines $S_i$ [48]. Therefore, such a perturbative wave function can be approximated by the RBM ansatz if the following inequality holds:

$$N_d \leqslant \alpha(d+1). \tag{8}$$

To illustrate our argument we explicitly build a map between the perturbation series and the RBM ansatz at the paramagnetic point of the TFI model and for the XXZ model (Appendix B).

As shown above, the capability of the RBM ansatz to approximate a perturbative series is limited given a fixed hidden-unit density $\alpha$. However, being able to reproduce a finite order of perturbative expansion is sufficient to approximate GS energy with high precision. This is illustrated in Fig. 2, where at low orders ($\leqslant 8$) the curve describing the RBM precision of the paramagnetic GS energy approximation (blue) follows closely the expansion in powers of $J$ of the exact energy (purple), computed using Jordan-Wigner transformation.

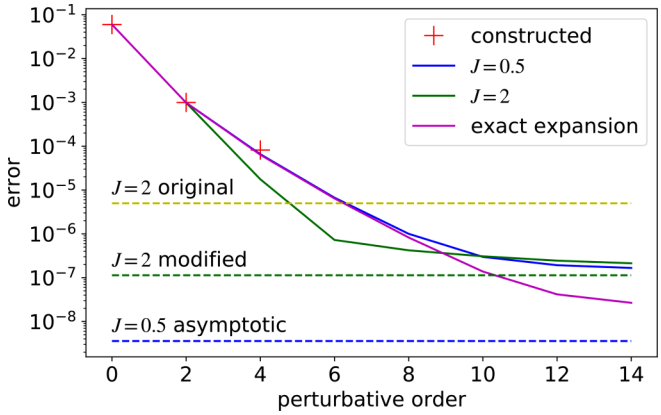


FIG. 2. Comparison of the energy accuracy obtained by the variational wave function with finite-range, translationally invariant couplings $W_{ij}$ (blue for the paramagnetic case, green for the ferromagnetic) to the perturbative expansion of the energy computed exactly (purple), as a function of the perturbative order. The results for the analytically constructed RBM are denoted by crosses. The asymptotic values of error for the paramagnetic case (blue dashed line), the asymptotic value for the ferromagnetic point with a gradually increasing RBM range $d$ (green dashed line), and the results of the original algorithm (yellow dashed line) are shown. The precision of the analytically constructed RBM wave function falls onto the curve for the exact expansion, indicating that the RBM ansatz captures several orders of perturbation theory. At higher orders the performance gain for RBM slows down and eventually saturates. The modified algorithm based on the perturbation theory argument outperforms the original one by 2 orders of magnitude for $J = 2$. The results are obtained for a system of $N = 18$ spins.

In Fig. 2 we relate a given RBM ansatz to a certain order of perturbation theory. We achieve this by taking the range $d$ of interactions $W_{ij}$ to be equal to the largest support among the irreducible terms that contribute to the corresponding order of the perturbative expansion of the energy. For example, the second-order correction to the energy is determined by the first-order correction to the GS wave function in the case of the TFI model, which gives $d = 2$ ($d = 1$) for the paramagnetic (ferromagnetic) point.

At this point we have established the mapping between the perturbation expansion and the RBM ansatz only in the case of the unperturbed GS being polarized along the $x$ axis. However, our argument is more general; that is, in Appendix A we show that in the case of an arbitrary product state the first-order perturbative wave function can be approximated by the RBM. Even though we cannot extend the perturbative argument to an arbitrary order of the perturbation theory rigorously, we still expect that given enough variational parameters higher orders can be captured in agreement with Eq. (8). We test our expectations numerically using the RBM as a variational wave function at the ferromagnetic point of the TFI. Interestingly, we observe that the restriction of the structure of the ansatz according to perturbation theory improves its convergence properties. We enforce this structure by modifying the original algorithm [10] in the following way: we gradually increase the range of variational couplings $W_{ij}$ at each step starting with the optimal parameters obtained in the previous step. We found that this modified procedure yields a lower error than the original procedure already at range $d = 3$.

To conclude this part we note that it is a nonlinearity of the hyperbolic cosine and the logarithm in Eq. (6) that allows the RBM to approximate the action of all possible local operators. We expect other ansätze with these properties to perform in a similar way. We propose two alternative functions and verify that they have an approximative power comparable to that of RBM (Appendix C).

## III. GENERAL HAMILTONIANS

While the TFI model provides a good test for the variational wave functions, it is important to apply RBM to more general Hamiltonians. As an example, we consider a family of translationally invariant Hamiltonians with random nearest-neighbor spin-spin interactions:

$$H_{\text{rand}} = \sum_i \sum_{\alpha,\beta=\{0,x,y,z\}} J_{\alpha\beta}\sigma_i^\alpha \sigma_{i+1}^\beta, \tag{9}$$

with $J_{\alpha\beta}$ being independent random coefficients uniformly distributed in the range $[-0.5, 0.5]$. We use the RBM ansatz with $\alpha = 1$ and translationally invariant variational parameters to approximate the GS energy of these Hamiltonians. Therefore, we have to restrict our consideration only to those Hamiltonians that have translationally invariant GSs and do not exhibit a spontaneous breaking of translational symmetry (Appendix E).

The error in approximating the GS energy for different realizations of Hamiltonian (9) exhibits a broad distribution [49]. To analyze the nature of this spread, we also studied the properties of the exact GS of the corresponding Hamiltonians (computed using exact diagonalization of system with $N = 18$). Namely, we study how the RBM performance depends on the entanglement entropy and the correlation length [50] (Fig. 3). The positive correlation of the error with both quantities is evident, indicating that RBM performs best for states with short-range correlations and low entanglement.

Another thing that we analyze is how the value of the error scales with the hidden-unit density depending on the entanglement of the GS. For the system size $N = 16$ we compare the performance of RBM ansätze with $\alpha = 2, 4$ to the one with unit $\alpha = 1$ (Fig. 3). We observe that the increase in density generally improves the performance of the ansatz [51]. However, this improvement varies by an order of magnitude, being the most significant for the low-entanglement GSs.

## IV. RBM AND MPS

We can further understand the interplay of RBM and entanglement by studying the relation of the ansatz to MPSs. Let us first point out that any finite-range RBM can be represented as an MPS. In particular, in Appendix D we show that an RBM with the range $d$ can be represented as an MPS with a bond dimension $D$ that scales exponentially in $d$, compared to the exponential in $d^2$ scaling obtained in Ref. [41].

To study the inverse mapping, we use RBM with $\alpha = 1$ to approximate MPSs with a given bond dimension. We generate random translationally invariant MPSs $|\Psi_{MPS}\rangle$ with $D = 2, 4, 8$ for a system of $N = 18$ spins [52]. Then we minimize the quantity $1 - |\langle\Psi_{MPS}|\Psi_{\text{RBM}}\rangle|^2$, which is plotted against the entanglement entropy in Fig. 3. One can see a

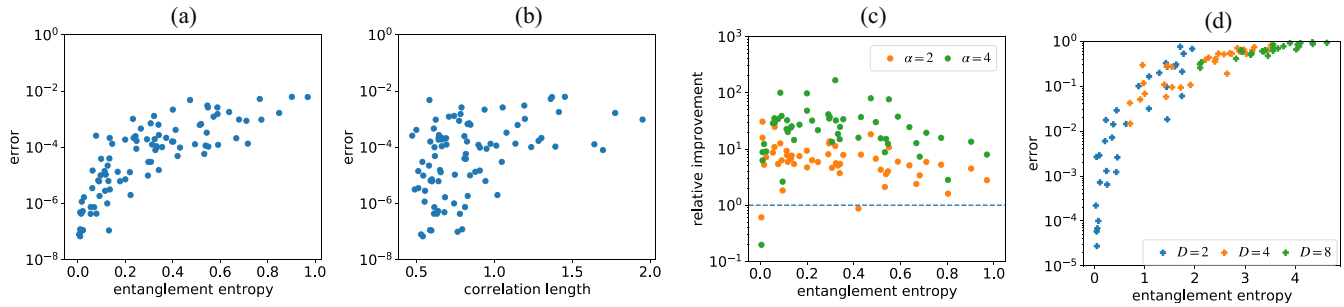


FIG. 3. The performance of the RBM with unit hidden spin density is analyzed in the context (a) of the entanglement entropy of the subsystem of six cites and (b) of the correlation length for the problem of the GS energy approximation of the random Hamiltonian. (c) The change in the RBM error with the increased hidden spin density compared to the base case of $\alpha = 1$ (dashed line) is summarized. (d) The RBM performance for states with higher entanglement is analyzed in the framework, where we generate random MPSs with different bond dimensions as target states. Generally, a larger bond dimension corresponds to larger entanglement. The numerical simulations are done for system sizes $N = 16$ in (c) and $N = 18$ in (a), (b), and (d).

clear dependence of the error on the entanglement entropy, which is in agreement with the results obtained for the random Hamiltonians above.

## V. CONCLUSIONS

In this paper we considered the RBM "black box" from the physical prospective. We pointed out an explicit connection between the RBM ansatz and the perturbative description of ground states of gapped models, which explains the remarkable accuracy of the sparse ($\alpha = 1$) RBM for simple models such as the TFI model [10]. In particular, we provided a mapping between RBM wave functions and perturbation theory for several concrete cases. This connection allowed us to introduce a simple modification of the optimization algorithm of RBM that leads to the improvement of the performance by several orders in certain cases.

Finally, we have investigated the performance of practical, sparse RBM states in approximating ground states of *random* (but translationally invariant) local spin Hamiltonians. We found that the approximation error has a broad distribution, and the sparse RBM performs very well for ground states with a low amount of entanglement. It is natural to expect that such states can be well approximated using just few orders of perturbation theory, which RBM can capture. We have also shown that for low-entanglement GSs the RBM performance is more significant when the number of hidden spins is increased.

We leave for future work the detailed investigation of how well infinite-range RBM can reproduce critical many-body states. Another open question is whether the volume-law entanglement that RBM states generally have may be useful for approximating, e.g., excited states and entanglement spreading in many-body systems.

The data for this work is available [53].

## ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation. We thank G. Carleo for very helpful correspondence and, in particular, for pointing out the influence of spontaneous symmetry breaking on the performance of a translationally invariant RBM. The numerical simulations were performed on the HPC cluster Baobab.

## APPENDIX A: A MAPPING BETWEEN THE FIRST-ORDER PERTURBATIVE EXPANSION OF A GENERAL PRODUCT STATE AND THE RBM

In the case of the finite-order perturbative expansion the perturbing operator (4) can be represented as the following sum:

$$\hat{W} = \sum_i \hat{w}_i, \tag{A1}$$

where operators $\hat{w}_i$ correspond to the local operators with a support localized between spin $i$ and $i + d - 1$, with $d$ being the range of the terms in a perturbation. In general, $w_i$ can be the sum of several local product operators. However, below, without a loss of generality, we assume that $w_i$ is just one product operator.

The simplification that appears in the first-order expansion is the validity of the representation

$$e^{\hat{W}} \approx \prod_i (1 + \hat{w}_i) \tag{A2}$$

and the fact that all the terms corresponding to the product of neighboring $\hat{w}_i$ with the intersecting support can be neglected since they appear in the higher order of the perturbative expansion. Therefore, this product can be viewed as a product of the operators that are localized on distinct spins, and thus, we can match it and the RBM disregarding cross terms.

First of all, let us show how any unperturbed product GS can be represented as the RBM. Such a GS can always be related to the state $|\psi_0\rangle$ with all spins polarized along the $x$ axis as

$$|\mathrm{GS}\rangle = \prod_n \hat{R}(\mathbf{a}_n, \sigma_n)|\psi_0\rangle, \quad \hat{R}(\mathbf{a}, \sigma) = e^{i(\mathbf{a}\sigma)}. \tag{A3}$$

Let us now notice the following property of the action of the rotation operator $\hat{R}(\mathbf{a}, \sigma)$ on the spin polarized along the

$x$ axis $|\uparrow\rangle_x$: the action of such an operator can always be expressed in terms of only the $\sigma_z$ operator up to normalization:

$$\hat{R}(\mathbf{a},\sigma)|\uparrow\rangle_x = \left[\cos a + \frac{\sin a}{|a|}\sin a\left(a_x\sigma^x + a_y\sigma^y + a_z\sigma^z\right)\right]|\uparrow\rangle_x \tag{A4}$$

$$= \left[\left(\cos a + \frac{\sin a}{|a|}a_x\right) + \frac{\sin a}{|a|}(a_z - ia_y)\sigma_z\right]|\uparrow\rangle_x \propto e^{\gamma\sigma_z}|\uparrow\rangle_x \equiv \tilde{R}(\gamma)|\uparrow\rangle_x, \tag{A5}$$

where $\gamma$ is determined from the equation $\tanh\gamma = (a_z - ia_y)/(|a|\cot a + a_x)$. Therefore, the unperturbed GS may be represented as $|\mathrm{GS}\rangle \propto \prod_j \tilde{R}(\gamma_j)|\psi_0\rangle$. The next step is to include the perturbative corrections to the GS. Note that each term $\hat{w}_i$ has a representation $\hat{w}_i = v_i\prod_{j=i}^{i+d-1}\hat{R}(\mathbf{A}_j,\sigma_j)$. Let us now show that the action of this operator on the GS can be expressed in terms of $\{\sigma_i^z\}$

$$\hat{w}_i|\mathrm{GS}\rangle = v_i\prod_{j=i}^{i+d-1}\hat{R}(\mathbf{A}_j,\sigma_j)\prod_n\hat{R}(\mathbf{a}_n,\sigma_n)|\psi_0\rangle = v_i\prod_{n<i}^{n\geqslant i+d}\hat{R}(\mathbf{a}_n,\sigma_n)\prod_{j=i}^{i+d-1}\hat{R}(\tilde{\mathbf{A}}_j,\sigma_j)|\psi_0\rangle \tag{A6}$$

$$\propto \prod_n\tilde{R}_n(\gamma_n)\prod_{j=i}^{i+d-1}\tilde{R}_j(\Gamma_j-\gamma_j)|\psi_0\rangle \equiv \tilde{R}_n(\gamma_n)w_i^z|\psi_0\rangle, \tag{A7}$$

where $\tilde{R}_i(\gamma_i) \equiv e^{\gamma_i\sigma_i^z}$, $\hat{R}(\tilde{\mathbf{A}}_i,\sigma_i) = \hat{R}(\mathbf{a}_i,\sigma_i)\hat{R}(\mathbf{A}_i,\sigma_i)$, and $\hat{R}(\tilde{\mathbf{A}}_i,\sigma_i)|\psi_0\rangle \propto \tilde{R}(\Gamma_i)|\psi_0\rangle$. We see that the structure of the operator acting on the GS is preserved even after all operators are represented solely in terms of $\{\sigma_i^z\}$. Generalizing this result to the whole perturbative expansion, we obtain

$$e^{\hat{W}}\prod_n\hat{R}(\mathbf{a}_n,\sigma_n)|\psi_0\rangle \propto e^{\sum_i w_i^z}\prod_j\tilde{R}(\gamma_j)|\psi_0\rangle. \tag{A8}$$

Note that the discrepancy between the left-hand side and the right-hand side appears only in the higher-order terms. The relation above allows us to match the perturbative expansion and RBM using the parametric argument discussed in the main text. Moreover, it sheds light on the role of $\{a_j\}$ parameters in the RBM ansatz. These parameters allow transformation from the true unperturbed GS to the state with all spins polarized along the $x$ axis.

## APPENDIX B: AN EXACT MAPPING OF THE PERTURBATIVE WAVE FUNCTION TO THE RBM ANSATZ FOR TFI AND XXZ MODELS

### 1. Paramagnetic point of the TFI Hamiltonian (3)

As defined in Eq. (4), $\hat{W}$ is given by

$$(He^{\hat{W}} - e^{\hat{W}}E)|\psi_0\rangle = 0. \tag{B1}$$

Let us write $\hat{W}$ and the energy $E$ in the form $\hat{W} = \hat{W}_1 + \hat{W}_2 + \hat{W}_3 + O(J^4)$ and $E = E_0 + E_1 + E_2 + E_3 + O(J^4)$ to outline their perturbative structure.

To simplify the analysis we want to fix the freedom in choice of $\hat{W}$. It could be chosen to depend solely on $\{\sigma_i^z\}$. This is the case since an action of any operator on $|\psi_0\rangle$ that is the product of $\{\sigma_i^\alpha\}$ results only in spin flips at certain positions. Thus, to cancel such terms we need to put $\sigma_i^z$ in the corresponding positions of $\hat{W}$. When commuted with $H_0$, some $\sigma_i^z$ would be changed to $\sigma_i^y$, but the resulting action on $|\psi_0\rangle$ stays the same up to a phase ($\sigma_i^z|\psi_0\rangle = i\sigma_i^y|\psi_0\rangle$).

Two important simplifications follow from this observation, namely, that $[V,\hat{W}] = 0$ at any order and $e^{\hat{W}_n+\hat{W}_m} = e^{\hat{W}_n}e^{\hat{W}_m}$. Also, a simple property of the TFI Hamiltonian is that all odd-energy corrections are zero. Having said that, the equations that determine the terms of interest in $\hat{W}$ can be written as

$$([H_0,\hat{W}_1] + V)|\psi_0\rangle = 0, \tag{B2}$$

$$\left([H_0,\hat{W}_2] + \tfrac{1}{2}[[H_0,\hat{W}_1],\hat{W}_1]\right)|\psi_0\rangle = E_2|\psi_0\rangle, \tag{B3}$$

$$\left([H_0,\hat{W}_3] + [[H_0,\hat{W}_1],\hat{W}_2] + \tfrac{1}{6}[[[H_0,\hat{W}_1],\hat{W}_1],\hat{W}_1]\right)|\psi_0\rangle = 0. \tag{B4}$$

To have more compact notation we denote $\lambda \equiv \frac{J}{4}$. If we take $\hat{W}_1 = \lambda\sum_i\sigma_i^z\sigma_{i+1}^z$, then $[H_0,\hat{W}_1] = 2i\lambda(\sum_i\sigma_i^y\sigma_{i+1}^z + \sum_i\sigma_i^z\sigma_{i+1}^y)$. Therefore, the term that should be canceled in the second order is

$$\frac{1}{2}[[H_0,\hat{W}_1],\hat{W}_1] = -4\lambda^2\sum_i\left(\sigma_i^x + \sigma_i^z\sigma_{i+1}^x\sigma_{i+2}^z\right). \tag{B5}$$

We choose $\hat{W}_2 = \lambda^2\sum_i\sigma_i^z\sigma_{i+2}^z$ and get an extensive energy correction $E_2 = -4\lambda^2\sum_i 1$; then $[H_0,\hat{W}_2] = 2i\lambda^2(\sum_i\sigma_i^y\sigma_{i+2}^z + \sum_i\sigma_i^z\sigma_{i+2}^y)$. The third-order interaction is then given by

$$\begin{aligned}[[H_0,\hat{W}_1],\hat{W}_2] + \frac{1}{6}[[[H_0,\hat{W}_1],\hat{W}_1],\hat{W}_1] &= \frac{16\lambda^3 i}{3}\sum_i\left(\sigma_i^y\sigma_{i+1}^z + \sigma_i^z\sigma_{i+1}^y\right)\\ &- 4\lambda^3\sum_i\left(\sigma_i^x\sigma_{i+1}^z\sigma_{i+2}^z + \sigma_i^z\sigma_{i+1}^z\sigma_{i+2}^y\right.\\ &\left.+ \sigma_i^z\sigma_{i+1}^x\sigma_{i+3}^z + \sigma_i^z\sigma_{i+2}^x\sigma_{i+3}^z\right).\end{aligned} \tag{B6}$$

This gives $\hat{W}_3 = 2\lambda^3\sum_i(\sigma_i^z\sigma_{i+3}^z + \sigma_i^z\sigma_{i+1}^z/3)$, and the total wave function up to normalization can be written as

$$|\psi\rangle = e^{\hat{W}}|\psi_0\rangle + O(\lambda^4), \tag{B7}$$

with

$$\hat{W} = \sum_i \left[ \left( \lambda - \frac{2\lambda^3}{3} \right) \sigma_i^z \sigma_{i+1}^z + \lambda^2 \sigma_i^z \sigma_{i+2}^z + 2\lambda^3 \sigma_i^z \sigma_{i+3}^z \right]. \tag{B8}$$

At this point it is clear how the RBM ansatz could be constructed in order to recover perturbative expansion; that is, the Taylor expansion of each cosine should reproduce local terms in the exponent of (B7). Since we need only terms even in operators $\sigma^z$, it is natural to take parameters $a_i$ and $b_j$ of the RBM ansatz to zero. The parameters $W_{ij}$ are chosen to be translationally invariant, i.e., dependent only on the difference between the position of a hidden unit and the physical spin $W_{ij} = \tilde{W}_{i-j}$. Since the largest support of the terms in $\hat{W}$ is 4, we restrict ourselves to four nonzero parameters $\tilde{W}_i$, $\{i = 0, \ldots, 3\}$. Applying all the assumptions to the RBM wave function (2), we may write the ansatz in the form

$$\Psi_{\mathrm{RBM}}(\{\sigma_j\}) = \left( \prod_{i=0}^{3} \cosh \tilde{W}_i \right)^N \prod_j \big(1 + w_0 w_1 \sigma_j^z \sigma_{j+1}^z + w_0 w_2 \sigma_j^z \sigma_{j+2}^z + w_0 w_3 \sigma_j^z \sigma_{j+3}^z + w_1 w_2 \sigma_{j+1}^z \sigma_{j+2}^z + w_1 w_3 \sigma_{j+1}^z \sigma_{j+3}^z + w_2 w_3 \sigma_{j+2}^z \sigma_{j+3}^z + w_0 w_1 w_2 w_3 \sigma_j^z \sigma_{j+1}^z \sigma_{j+2}^z \sigma_{j+3}^z \big), \tag{B9}$$

where $w_i = \tanh \tilde{W}_i$. Since we assume this expression has the perturbative structure, the terms containing Pauli matrices should be small. Thus, we can take the approximate logarithm of this expression to get the exponential form (B7) up to normalization. Then we could easily identify the value of the parameters $w_i$ that reproduce the perturbative expansion to the third order: $w_0 = 1$ (it is possible with an exponential precision for finite $\tilde{W}_0$), $w_1 = \lambda - 2\lambda^3$, $w_2 = \lambda^2 + \lambda^3$, and $w_3 = 2\lambda^3$.

To account for the first three orders of the perturbation theory we can set the parameters to the following values: $\tanh W_0 = 1$ (this could be done with exponential precision for finite $W_0$), $\tanh W_1 = \lambda - \lambda^3/2$, $\tanh W_2 = \lambda^2 + \lambda^4/2$, $\tanh W_3 = 2\lambda^3$. For these parameter values the RBM ansatz coincides with the perturbative expansion given the accuracy.

### 2. Ferromagnetic point

Here, we construct the RBM ansatz that captures the perturbation series at the ferromagnetic point up to the second order. The main challenge here is that in contrast to the paramagnetic case the unperturbed GS is a product state of spins polarized along the $z$ axis. However, as we have shown in Appendix A even in this case the first order is guaranteed to be captured by the RBM. Moreover, for this particular Hamiltonian we show how this could be extended to the second order.

The TFI Hamiltonian is a sum of a zero-order part $H_0 = -\sum_i \sigma_i^z \sigma_{i+1}^z$ and the perturbation $V = -4\lambda \sum_i \sigma_i^x$. The first two orders of the perturbing term $\hat{W}$ (4) are given by

$$\hat{W}_1 = \lambda \sum_i \sigma_i^x,$$

$$\hat{W}_2 = \lambda^2 \sum_i \sigma_i^x \sigma_{i+1}^x.$$

Our first goal is to express the second-order wave function in terms of $\{\sigma_i^z\}$ operators acting on the state $|\tilde{\psi}_0\rangle$ of all spins polarized along the $x$ axis. First of all, an unperturbed GS $|\psi_0\rangle$ can be represented as

$$|\psi_0\rangle \propto \prod_i e^{\Gamma \sigma_i^z} |\tilde{\psi}_0\rangle, \tag{B10}$$

with exponential precision for a large $\Gamma$. In a similar manner we express the second-order wave function as

$$e^{\hat{W}_1 + \hat{W}_2} |\psi_0\rangle \propto \prod_i e^{\Gamma \sigma_i^z} \big[ 2e^{-2\Gamma} + (\lambda + \lambda^2) - (\lambda + \lambda^2)\sigma_i^z - \lambda^2 \sigma_{i+1}^z + \lambda^2 \sigma_i^z \sigma_{i+1}^z \big] |\psi_0\rangle \tag{B11}$$

within the precision of the expansion. We can approximate this wave function by an RBM with a range $d = 2$ and translationally invariant weights satisfying conditions $a_i = \Gamma$ for all spins, $\tanh b_j = [(\lambda + \lambda^2)/(2e^{-2\Gamma} + \lambda + \lambda^2)]^{1/2}$ for all hidden spins, and $\tanh W_{ii} = -\tanh b_j$, $\tanh W_{ii+1} = -\lambda^2/(2e^{-2\Gamma} + \lambda + \lambda^2)/\tanh b_j$, $\tanh W_{ij} = 0$ for $j \neq i, i+1$.

### 3. XXZ Hamiltonian

Let us now map a perturbative expansion in $\lambda$ of the GS of the antiferromagnetic (AF) XXZ Hamiltonian to a RBM wave function. This Hamiltonian is given by

$$H_{\mathrm{AF}} = \sum_i \sigma_i^x \sigma_{i+1}^x + \lambda \sum_i \left( \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z \right). \tag{B12}$$

In the limit of $\lambda = 1$ it becomes an AF Heisenberg model, which was numerically studied using the RBM in Ref. [10]. Note that an unperturbed GS of this Hamiltonian is doubly degenerate with one state being $|\mathrm{GS}\rangle = |\uparrow\downarrow\cdots\rangle$ and another being translated by one index. Below we build an expansion over the state $|\mathrm{GS}\rangle$. Note that this state can be mapped onto the state with all spins polarized along the $x$ axis $|\psi_0\rangle$ if the following unitary transformation is applied:

$$U = \prod_{i\mathrm{even}} \sigma_i^z, \quad |\psi_0\rangle = U|\mathrm{GS}\rangle. \tag{B13}$$

And the Hamiltonian is transformed to

$$\tilde{H}_{\mathrm{AF}} \equiv U H_{\mathrm{AF}} U^\dagger = -\sum_i \sigma_i^x \sigma_{i+1}^x - \lambda \sum_i \left( \sigma_i^y \sigma_{i+1}^y - \sigma_i^z \sigma_{i+1}^z \right). \tag{B14}$$

Within the new basis one can use the translationally invariant RBM ansatz to approximate the perturbative function. The first-order correction to $|\psi_0\rangle$ is determined from Eq. (B2), where $H_0 = -\sum_i \sigma_i^x \sigma_{i+1}^x$ and $V = \tilde{H}_{\mathrm{AF}} - H_0$. The

second-order correction is determined from

$$\left([H_0,\hat{W}_2]+[V,\hat{W}_1]+\tfrac{1}{2}[[H_0,\hat{W}_1],\hat{W}_1]\right)|\psi_0\rangle = E_2|\psi_0\rangle. \tag{B15}$$

The corresponding perturbing terms are given by

$$\hat{W}_1 = -\frac{\lambda}{2}\sum_i \sigma_i^z\sigma_{i+1}^z, \tag{B16}$$

$$\hat{W}_2 = -\frac{\lambda^2}{4}\sum_i \left(\sigma_i^z\sigma_{i+2}^z - \sigma_i^z\sigma_{i+1}^z\sigma_{i+2}^z\sigma_{i+3}^z\right). \tag{B17}$$

To capture this perturbative structure we use the RBM with hidden-unit density $\alpha = 2$ and the range $d = 4$. As in the case of the TFI model we take parameters $a_i$ and $b_j$ to zero, and for the couplings $W_{ij}$, we assume translational invariance. Since there are two hidden spins for each real one, we have to introduce two sets of translationally invariant parameters $w^{(1)}_{i-j/2} \equiv \mathrm{arctanh} W_{ij}$ for even hidden spins $j$ and $w^{(2)}_{i-(j-1)/2} \equiv \mathrm{arctanh} W_{ij}$ for odd ones. To capture the second-order perturbative expansion of the GS we may choose the following values of the parameters:

$$w_0^{(2)} = \lambda^{\frac{1}{2}}x/t \approx 0.376\lambda^{\frac{1}{2}}e^{-i\pi/4},$$

$$w_0^{(1)} = w_0^{(2)} + \frac{\lambda^{\frac{3}{2}}tx}{4(t^4+1/16)} \approx 0.376\lambda^{\frac{1}{2}}e^{-i\pi/4} + 0.468\lambda^{\frac{3}{2}}e^{i\pi/4},$$

$$w_1^{(1)} = -w_1^{(2)} = w_3^{(2)} = \lambda^{\frac{1}{2}}t \approx 0.411\lambda^{\frac{1}{2}}e^{i\pi/4},$$

$$w_2^{(1)} = w_2^{(2)} = -\lambda^{\frac{1}{2}}y/t \approx 0.984\lambda^{\frac{1}{2}}e^{-i\pi/4},$$

$$w_3^{(1)} = w_3^{(2)} - \frac{t^3\lambda^{\frac{3}{2}}}{4(t^4+1/16)} = 0.411\lambda^{\frac{1}{2}}e^{i\pi/4} + 0.513\lambda^{\frac{3}{2}}e^{i\pi/4},$$

where $x = (-1+\sqrt{5})/8$, $y = (-1-\sqrt{5})/8$, and $t = \frac{e^{i\pi/4}}{2}(9/2+2\sqrt{5}-\sqrt{(9/2+2\sqrt{5})^2-8})^{\frac{1}{4}}$ are numerical constants that emerge in the process of mapping of the expansion to the RBM.

Another thing that is worth mentioning is that if we allow an absence of translational invariance in the RBM parameters, we can easily construct the second-order perturbative expansion of the GS in the original basis. It can be achieved if we include parameters $\{a_i\}$, which take the values $a_i = 0$ on odd sites $i$ and $a_i = i\pi/2$ on even sites.

## APPENDIX C: A FAMILY OF ANSÄTZE

Another interesting question concerns the uniqueness of the RBM ansatz. The connection of the RBM to perturbation theory suggests that this ansatz belongs to a whole class of variational wave functions which can reproduce perturbative expansions. To demonstrate this, we investigate two examples below. In the first example, each hidden spin generates a quadratic polynomial:

$$\Psi_{quad}(\{\sigma_j\}) = \exp\left(\sum_j a_j\sigma_j^z\right)\prod_{i=1}^{M}\left(1+\alpha_i\left[b_i+\sum_j W_{ij}\sigma_j^z\right]^2\right), \tag{C1}$$

and coefficients $\alpha_i$ are variational parameters. The second trial ansatz is a sixth-order polynomial with only even powers and fixed coefficients

$$\Psi_{fix}(\{\sigma_j\}) = \exp\left(\sum_j a_j\sigma_j^z\right)\prod_{i=1}^{M}\sum_{k=0}^{3}\left(b_i+\sum_j W_{ij}\sigma_j^z\right)^{2k}. \tag{C2}$$

We restrict ourselves to only even powers since we found such ansätze yield a better performance.

The approximative power of these ansätze is summarized in Fig. 4, where we plot the accuracy of the GS energy approximation as a function of the hidden-unit density, using either the original algorithm or the modified one. The numerical results are presented for the system of $N = 18$ spins and four Hamiltonians, namely, a paramagnetic one ($J = 0.5$), a ferromagnetic one ($J = 2$), a critical one ($J = 1$), and a nonintegrable one, defined as

$$H_{ni} = -\sum_i \sigma_i^x - h\sum_i \sigma_i^z - J\sum_i \sigma_i^z\sigma_{i+1}^z, \tag{C3}$$

where we choose $J = 1$ and $h = 0.5$. All three ansätze exhibit similar numerical performance, which indicates that the RBM ansatz belongs to a wider class of functions that are capable of capturing local correlations.

## APPENDIX D: EXACT MAPPING OF THE RBM TO MPSs

Let us start with constructing an exact and transparent mapping from the finite-range RBM to MPSs in one dimension. The RBM ansatz with coupling range $d$ can be written as

$$\Psi_{\mathrm{RBM}}(\{\sigma_j\}) = \prod_i f^{(i)}_{\sigma_i\cdots\sigma_{i+d-1}}, \tag{D1}$$

where $f^{(i)}_{\sigma_i\cdots\sigma_{i+d-1}}$ is a function acting on $d$ neighboring spins. For example, to write the RBM wave function with $\alpha = 1$ in this form, we can choose $f^{(i)}_{\sigma_i\cdots\sigma_{i+d-1}} = e^{a_i\sigma_i^z}\cosh(b_i + \sum_{j=0}^{d-1} W_{i,i+j}\sigma^z_{i+j})$. Next, we construct a mapping from this form to an MPS form. We can associate the tensor $A^{\sigma_i}_{i-d,\ldots,i+d-1}$ with each function $f^{(i)}_{\sigma_i\cdots\sigma_{i+d-1}}$ using the following rule:

$$A^{\sigma_i}_{j_1\cdots j_{d-1}k_1\cdots_{d-1}} = f^{(i)}_{\sigma_i k_1\cdots k_{d-1}}\delta_{\sigma_i j_1}\,\delta_{k_1 j_2}\times\cdots\times\delta_{k_{d-2}j_{d-1}}, \tag{D2}$$

where the first $d-1$ indices $(j_1\cdots j_{d-1})$ of the tensor should be considered one index taking $2^{d-1}$ values, which should be contracted with the $(i-1)$th tensor, and the last $d-1$ indices $(k_1\cdots k_{d-1})$ are contracted with the $(i+1)$th tensor. Then the RBM wave function can be written in the form

$$\Psi_{\mathrm{RBM}} = \sum_{\{\sigma_j\}}\mathrm{Tr}(A^{\sigma_1}\cdots A^{\sigma_N})|\{\sigma_j\}\rangle. \tag{D3}$$

Our argument makes use of the fact that the spin system either has periodic boundary conditions or is infinite; however, this argument can be effortlessly generalized to finite-size systems with boundaries. Also, the slightly modified argument relates the RBM to MPSs in higher dimensions.

In contrast to Ref. [41], where the mapping was made to an MPS with a bond dimension $D$ that is exponential in the number of pairs of the connected hidden unit and the physical spin that have a given spin in between ($\sim e^{d^2}$), we get an exponential ($D \sim e^d$) scaling with the range of the RBM.

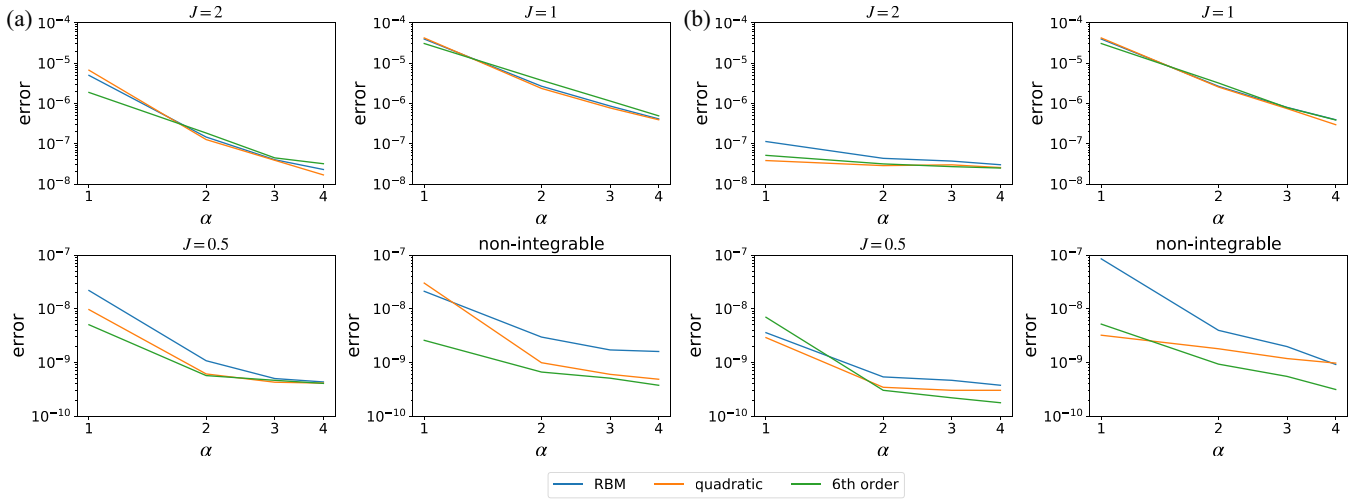


FIG. 4. Comparison of the performance of the RBM ansatz and two polynomial ansätze using (a) the original algorithm and (b) the one in which the range of the couplings $W_{ij}$ is gradually increased. Four Hamiltonians have been considered: Ferromagnetic, paramagnetic, critical, and nonintegrable (see the main text). In all cases, the modified algorithm allows one to achieve comparable or lower error. The system size is $N = 18$.

## APPENDIX E: PROCESSING OF DATA FOR THE RBM APPLIED TO RANDOM HAMILTONIANS AND MPSs

In this paper we are interested in the simplest, translationally invariant RBM wave functions. Therefore, we need to check whether the problem at hand (either the GS of a random Hamiltonian or a random MPS) is compatible with this symmetry of the ansatz. The most straightforward property to check is the translational symmetry of the target wave function. While this symmetry is present by construction for MPSs analyzed in this paper, GSs of a random translationally invariant Hamiltonians may lack it. We do not expect (and do not see) accidental degeneracy of the GSs related to the symmetry breaking. However, some GSs, even being unique, do not transform trivially under the action of the translation operator, acquiring a nontrivial phase. Removing such states from our sample, we found that there are 109 out of overall randomly generated 150 Hamiltonians that have translationally invariant GSs.

However, the translational symmetry of the GS is not sufficient to justify the use of the translationally invariant RBM ansatz, which was brought to our attention by G. Carleo. In particular, in the case of spontaneous breaking of translational symmetry there are two or more distinct ground states in the thermodynamic limit. The true eigenstates of a finite-size system, which are "Schrödinger's cat" superpositions of these broken-symmetry states, are translationally invariant. The sparse RBM ansatz has difficulty approximating such catlike states and, instead, may give a good approximation of the symmetry-broken states. Since we chose to consider only the translationally invariant RBM, we decided not to include realizations of random Hamiltonian ground states which exhibit spontaneous symmetry breaking.

To identify such cases, we removed the states that have strong nonlocal correlations. To that end, we first define the correlation function

$$K_\alpha(n) = \langle \sigma_1^\alpha \sigma_{1+n}^\alpha \rangle - \langle \sigma_1^\alpha \rangle \langle \sigma_{1+n}^\alpha \rangle, \quad \alpha = x, y, z. \tag{E1}$$

To check that a given state does not exhibit nonlocal correlations, we then apply a heuristic condition that distant (we take $n = 8, 9$) correlations are smaller than $C \max_\alpha \{\text{mean}_n[|K_\alpha(n)|]\}$, where $C$ is a constant, chosen at $C = 0.5$. We found that this condition reduces the number of random Hamiltonians from 109 to 91 and the number of MPSs from 116 to 88. The postselected data are are presented in Fig. 3, and the raw data are presented in Fig. 5.

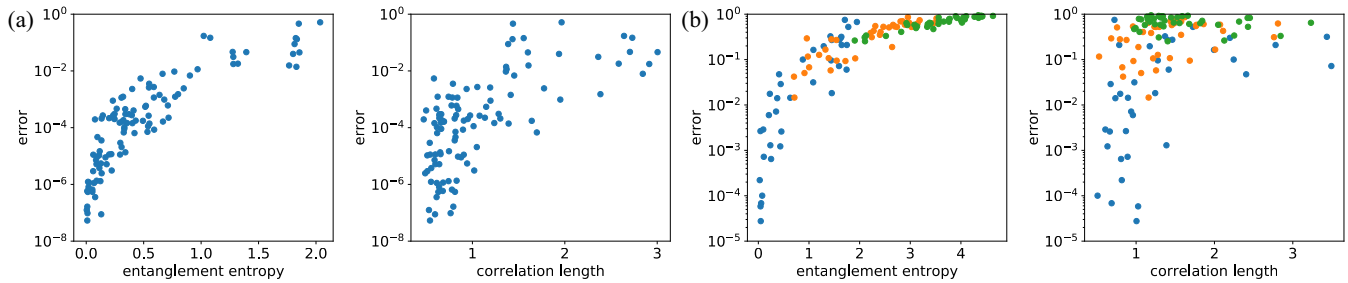


FIG. 5. The raw data for Fig. 3: random Hamiltonian (a) and random MPS (b).

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[2] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
[3] R. Orús, Ann. Phys. (N.Y.) **349**, 117 (2014).
[4] D. Perez-Garcia, F. Verstraete, M. Wolf, and J. Cirac, Quantum Inf. Comput. **7**, 401 (2007).
[5] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).
[6] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[7] U. Schollwöck, Ann. Phys. (N.Y.) **326**, 96 (2011).
[8] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).
[9] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
[10] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[11] H. Saito, J. Phys. Soc. Jpn. **86**, 093001 (2017).
[12] X. Gao and L.-M. Duan, Nat. Commun. **8**, 662 (2017).
[13] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017).
[14] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X **8**, 011006 (2018).
[15] Z. Cai and J. Liu, Phys. Rev. B **97**, 035116 (2018).
[16] R. Kaubruegger, L. Pastori, and J. C. Budich, Phys. Rev. B **97**, 195136 (2018).
[17] X. Liang, W.-Y. Liu, P.-Z. Lin, G.-C. Guo, Y.-S. Zhang, and L. He, Phys. Rev. B **98**, 104426 (2018).
[18] L. Pastori, R. Kaubruegger, and J. C. Budich, Phys. Rev. B **99**, 165123 (2019).
[19] J. Bausch and F. Leditzky, New J. Phys. **22**, 023005 (2020).
[20] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Phys. Rev. Lett. **121**, 167204 (2018).
[21] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
[22] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[23] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[24] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).
[25] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).
[26] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **86**, 044708 (2017).
[27] F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
[28] A. Tanaka and A. Tomiya, J. Phys. Soc. Jpn. **86**, 063001 (2017).
[29] J. Venderley, V. Khemani, and E.-A. Kim, Phys. Rev. Lett. **120**, 257204 (2018).
[30] W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017).
[31] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).
[32] M. Koch-Janusz and Z. Ringel, Nat. Phys. **14**, 578 (2018).
[33] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
[34] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 041101(R) (2017).
[35] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).
[36] G. E. Hinton and R. R. Salakhutdinov, Science **313**, 504 (2006).
[37] R. R. Salakhutdinov, A. Mnih, and G. E. Hinton, in *Proceedings of the 24th International Conference on Machine Learning, ICML '07* (ACM, New York, 2007), pp. 791–798.
[38] H. Larochelle and Y. Bengio, in *Proceedings of the 25th International Conference on Machine Learning, ICML '08* (ACM, New York, 2008), pp. 536–543.
[39] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).
[40] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B **96**, 195145 (2017).
[41] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).
[42] Z.-A. Jia, Y.-H. Zhang, Y.-C. Wu, L. Kong, G.-C. Guo, and G.-P. Guo, Phys. Rev. A **99**, 012307 (2019).
[43] Y.-H. Zhang, Z.-A. Jia, Y.-C. Wu, and G.-C. Guo, arXiv:1809.08631.
[44] S. Lu, X. Gao, and L.-M. Duan, Phys. Rev. B **99**, 155136 (2019).
[45] S. Czischek, M. Gärttner, and T. Gasenzer, Phys. Rev. B **98**, 024311 (2018).
[46] Y. Zheng, H. He, N. Regnault, and B. A. Bernevig, Phys. Rev. B **99**, 155129 (2019).
[47] Y. Huang and J. E. Moore, arXiv:1701.06246.
[48] Here we ignore the contribution of the parameters $\{a_j\}$. Their role is discussed in Appendix A.
[49] We verified that these values of errors can be reproduced using the NETKET package [54] for a random sample of 15 points .
[50] If $K_i(n) = \langle \sigma_1^i \sigma_{1+n}^i \rangle - \langle \sigma_1^i \rangle \langle \sigma_{1+n}^i \rangle$, then the correlation length $l$ is defined by $l = \max_{i=\{x,y,z\}} \sum_n \mathrm{dist}(1, 1+n)|K_i(n)|/\sum_n |K_i(n)|$, where $\mathrm{dist}(i, j)$ is the distance between two spins at sites $i, j$ with the periodic boundary conditions taken into account .
[51] A few points for which the performance decreased can be explained by a slower convergence of the RBM optimization, which is expected due to the larger number of parameters.
[52] We construct MPSs as products of identical random matrices that have independent random entries with both real and imaginary parts being uniformly distributed in the interval $[-0.5, 0.5]$. We postselect these states according to the procedure described in Appendix E.
[53] DOI: 10.26037/yareta:pen5stdvz5bp3fgjhyyf3l2jiq.
[54] NETKET, https://www.netket.org.