Neural Gutzwiller-projected variational wave functions

Neural Gutzwiller-projected variational wave functions

Francesco Ferrari,[1,*] Federico Becca,[2] and Juan Carrasquilla[3,4]

[1] *SISSA-International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste, Italy*
[2] *Dipartimento di Fisica, Università di Trieste, Strada Costiera 11, I-34151 Trieste, Italy*
[3] *Vector Institute, MaRS Centre, Toronto, Ontario, M5G 1M1, Canada*
[4] *Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada*

(Dated: September 16, 2019)

Variational wave functions have enabled exceptional scientific breakthroughs related to the understanding of novel phases of matter. Examples include the Bardeen-Cooper-Schrieffer theory of superconductivity, the description of the fractional quantum Hall effect through the Laughlin state, and Feynman's variational understanding of large-scale quantum effects in liquid Helium. More recently, Gutzwiller-projected wave functions, typically constructed from fermionic degrees of freedom, have been employed to examine quantum spin models in the presence of competing interactions, where exotic phases with no spontaneous symmetry breaking and fractional excitations may exist. In this work, we investigate the aforementioned fermionic wave functions supplemented with neural networks, specifically with the so-called restricted Boltzmann machine (RBM), to boost their accuracy and obtain reliable approximations to the ground state of generic spin models. In particular, we apply our neural augmented fermionic construction to the description of both magnetically ordered and disordered phases of increasing complexity, including cases where the ground state displays a non-trivial sign structure. Even though the RBM state is by far more effective for Néel states endowed with a particularly simple sign structure, it provides a significant improvement over the original fermionic state in highly frustrated regimes where a complex sign structure is anticipated, thus marking the path to an understanding of strongly-correlated spin models on the lattice via neural Gutzwiller-projected variational wave functions.

I. INTRODUCTION

Fractionalization refers to a set of phenomena in quantum many-body physics where a collection of strongly interacting microscopic degrees of freedom, such as electrons or spins on a lattice, break down into multiple quasiparticles. One of the most prominent examples is the fractional quantum Hall effect [1], where the constituent electrons decay into quasiparticles carrying fractions of the electron charge. Another prime example of fractionalization arises in the study of spin liquids [2, 3], which are highly-entangled states of matter that host a wide array of exotic phenomena such as emergent gauge structures, non-local excitations, and fractional quasiparticles. These quasiparticles, often termed "spinons", behave as fractions of ordinary magnons and spin waves.

Traditionally, gauge theories have played a central role in the study of quantum spin liquids since their particular structure can capture the highly peculiar non-local excitations and entanglement properties present in these exotic phases [2–4]. To construct such gauge theories, physicists rely on representations of the original degrees of freedom in terms of pairs of bosons or fermions, whose local Hilbert space is constrained to a subspace corresponding to the original Hilbert space of the spin system. This reformulation results in a rich theory of quasiparticles coupled to gauge fields with the potential to shed light onto the properties of the original system.

Since a gauge theoretical reformulation of a quantum spin system is formally exact, an explanation of all physically conceivable phases in the original models is always possible. Whereas conventional phases of matter, such as Néel antiferromagnets and valence bond solids, are associated with the confining phases of the gauge theory, fractionalization is associated with the deconfined phases where the spinons materialize as legitimate excitations of the system. Thus, the hope is that this reformulation may help us understand exotic phases which naturally realize the quasiparticles and the emerging gauge fields of the theory. Originally suggested by Baskaran and Anderson in the context of high temperature superconductivity [5], fractionalization and emergent gauge structures have been demonstrated in several exactly solvable systems such as Kitaev's toric code [6] and honeycomb model [7], string-net condensates [8], as well confirmed in microscopic models [9–11], variational wave functions [12], and also experiments [1].

In practice, however, gauge theories of quantum spin systems are typically intractable. In such cases, the generic procedure, also known as the "parton" construction, is to rewrite the original model in terms of canonical bosons or fermions, perform a mean-field decoupling of the gauge theory, and then study the implications of the mean-field theory and its stability against fluctuations [4]. Alternatively, a numerical approach can be introduced, where wave functions from a mean-field treatment of the gauge theory are numerically projected back to the original Hilbert space through the variational Monte Carlo method (VMC) [13].

The accuracy and degree of applicability of these

* frferra@sissa.it

states, here referred to as Gutzwiller projected wave functions, have been demonstrated in several prototypical models of frustrated magnetism [14–19]. The main advantage of this strategy is that the states it produces can lead to very accurate estimates of the ground-state energy but their most important characteristic is the interpretability, as these states often provide us with a physically transparent description of the phases of the model under consideration. On the other hand, the main disadvantage of this approach is the lack of systematic ways of improving the quality of the approximations the construction introduces. While the parton construction remains a powerful approach to quantum spin liquids, it is clear that the Gutzwiller projector only partially reintroduces the missing gauge fluctuations of the mean-field treatment, which may not be sufficient to capture important long wavelength properties of the system. This effect is particularly important in special cases, e.g., in $U(1)$ spin liquids, since Gutzwiller projected states do not account for crucial spatial gauge fluctuations [20, 21].

Here, we explore the possibility of improving the accuracy of the parton construction through neural networks. In particular, because of their extensive use in condensed matter physics and quantum information [22–30], we consider restricted Boltzmann machines (RBM), which, due to the nature of their non-local structure, have been shown to represent some highly-entangled many-body states using a relatively small number of parameters [23, 26, 27, 31]. The family of states we consider can be thought of as quantum mechanical version of the so-called product of experts idea used in probabilistic models in machine learning [32], where the core strategy is to combine several simpler wave functions (the "experts") by multiplying their amplitudes in a certain computational basis [13, 29, 31, 33, 34]. Concretely, the family of states we consider is written as a product of a Gutzwiller-projected fermionic state and a complex-valued RBM. By virtue of the high representation power of the RBM [27, 35], the expectation is that the RBM may help us improve the accuracy of the Gutzwiller projected states and in part alleviate the problem of the missing gauge fluctuations in these family of states.

The paper is organized as follow: In Sec. II, we introduce the spin models that are considered in this study; in Sec. III, we define the variational wave functions; in Sec. IV, we briefly describe the numerical optimization of the wave functions; in Sec. V, we discuss the results; finally, in Sec. VI, we draw our conclusions.

II. SPIN MODELS

We consider two-dimensional (2D) quantum lattice models that are traditionally studied in condensed matter physics. These models host a wide array of effects that are relevant to the understanding of physical phenomena such as low-temperature experiments with superfluid helium [36], Fe-based superconducting materials [37], triangular lattice compounds [38], fractional excitations in the quantum antiferromagnets [39], among many others. We begin by introducing the $J_1 - J_2$ Heisenberg model, an archetypical model of frustration in quantum antiferromagnets. Its Hamiltonian is given by

$$\mathcal{H} = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{1}$$

where $\mathbf{S}_i = (S_i^x, S_i^y, S_i^z)$ are spin-1/2 degrees of freedom sitting on the sites of a square lattice and $J_1, J_2 \geq 0$ characterize the antiferromagnetic couplings between the magnetic degrees of freedom. Here, $\langle \cdots \rangle$ and $\langle\langle \cdots \rangle\rangle$ restrict the summations over first- and second-nearest neighboring sites, respectively. Throughout this paper, we focus on two different values of the *frustrating ratio*, namely $J_2/J_1 = 0$ and $J_2/J_1 = 0.5$, to assess the accuracy of the aforementioned variational wave functions in two different regimes. For $J_2 = 0$ the model reduces to the (unfrustrated) Heisenberg model, whose ground state displays long-range Néel magnetic order [40, 41], At finite J_2, the energetic competition between the antiferromagnetic couplings at first- and second-nearest distances introduces frustration, resulting in a reduction of the magnetic order. In particular, the system exhibits a high degree of frustration around $J_2/J_1 \approx 0.5$, where the nature of the ground state of the model is a long standing problem [42, 43]. Different scenarios have been suggested, such as the onset of a valence-bond solid phase, with columnar [44, 45] or plaquette [46, 47] order, and/or the existence of a spin-liquid phase [48–50].

Due to its relevance to physical situations ranging from low-temperature experiments with superfluid helium [36], magnetic insulators [51], and ultracold gases in optical lattices [52, 53], we also consider the XY model on the square lattice, whose Hamiltonian is given by

$$\mathcal{H} = J \sum_{\langle i,j \rangle} (S_i^x S_j^x + S_i^y S_j^y). \tag{2}$$

In this case, the ground state exhibits Néel order with the spins aligned in the xy plane. A fundamental difference between the XY and the unfrustrated Heinseberg models comes from their symmetries: while the Heisenberg exchange interaction is invariant under global spin rotations around any axis (i.e., it exhibits $SU(2)$ symmetry), the XY coupling is invariant only under global rotations around S_z (leading to a $U(1)$ symmetry).

Finally, we consider the Heisenberg model on the triangular lattice, a model of geometric frustration relevant to the understanding of organic Mott insulators displaying spin liquid behavior [54]. Its Hamiltonian is given by Eq. (1) with $J_2 = 0$. Due to frustration, this system exhibits strong quantum fluctuations. Even though this model was initially proposed as a candidate host for the resonating valence-bond spin liquid [55], which motivated the entire field of spin liquids, numerical evidence [56, 57] suggested that the ground state of the system is magnetically ordered with a 120° pattern, which is the accepted consensus.

III. CONSTRUCTION OF THE WAVE FUNCTIONS

In general, the variational *Ansätze* employed in this work can be written as a product of a Gutzwiller-projected fermionic wave function, $|\Psi_f\rangle$, and a many-body correlator, $\hat{\mathcal{C}}$:

$$|\Psi_{\mathcal{C}}\rangle = \hat{\mathcal{C}}|\Psi_f\rangle = \sum_{\sigma} \mathcal{C}(\sigma)\langle\sigma|\Psi_f\rangle|\sigma\rangle. \tag{3}$$

In the above formula we have inserted a resolution of the identity ($\hat{\mathbb{1}} = \sum_\sigma |\sigma\rangle\langle\sigma|$) and exploited the fact that $\hat{\mathcal{C}}$ is diagonal in the many-body computational basis $\{|\sigma\rangle\}$.

In order to construct the fermionic part of the *Ansatz* we turn to a parton construction, representing the spin degrees of freedom in terms of Abrikosov fermions [4]:

$$\mathbf{S}_i = \frac{1}{2}\sum_{\alpha,\beta} f^\dagger_{i,\alpha}\boldsymbol{\sigma}_{\alpha,\beta}f_{i,\beta}. \tag{4}$$

Here $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is the vector of Pauli matrices and the operator $f^\dagger_{i,\alpha}$ ($f_{i,\alpha}$) creates (annihilates) a fermion with spin α at site i. Within the Abrikosov picture, the commutation relations among spins are a consequence of the fermionic anticommutation relations. However, the fermionic representation of spins enlarges the Hilbert space of the system, allowing each lattice site to be in four different states, namely $|0\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, and $|\uparrow\downarrow\rangle$. Therefore, the mapping of Eq. (4) is a faithful representation of spin degrees of freedom only if the fermions are restricted to the subspace of configurations with one fermion per site. This constraint can be enforced by the Gutzwiller projection operator $\mathcal{P}_G = \prod_i (n_{i,\uparrow} - n_{i,\downarrow})^2$ (where $n_{i,\sigma} = f^\dagger_{i,\sigma}f_{i,\sigma}$). The wave function $|\Psi_f\rangle$ is obtained by Gutzwiller-projecting a certain fermionic state $|\Phi_0\rangle$ (see below):

$$|\Psi_f\rangle = \mathcal{P}_{S^z_{tot}=0}\mathcal{P}_G|\Phi_0\rangle. \tag{5}$$

In addition to the Gutzwiller projector, here we consider a second projector, $\mathcal{P}_{S^z_{tot}=0}$, which restricts the wave function to the subspace of configurations with $S^z_{tot} = \sum_i S^z_i = 0$. We emphasize the fact that both projections are treated exactly by sampling only the fermionic configurations which satisfy the desired constraints. Finally, $|\Phi_0\rangle$ is obtained by computing the ground state of a quadratic (i.e., a mean-field) Hamiltonian ($\mathcal{H}_0$) of Abrikosov fermions,

$$\mathcal{H}_0 = \mathcal{H}_{\rm BCS} + \mathcal{H}_{\rm AF}, \tag{6}$$

which we have split into two parts for clarity. The first part is a Bardeen-Cooper-Schrieffer (BCS) Hamiltonian containing a (complex) hopping term ($t_{i,j}$) and a singlet pairing term ($\Delta_{i,j} = \Delta_{j,i}$):

$$\mathcal{H}_{\rm BCS} = \sum_{i,j,\sigma} t_{i,j}f^\dagger_{i,\sigma}f_{j,\sigma} + \sum_{i,j}\Delta_{i,j}f^\dagger_{i,\uparrow}f^\dagger_{j,\downarrow} + h.c.. \tag{7}$$

The above Hamiltonian is invariant under global spin rotations and, therefore, the spin wave functions obtained by the Gutzwiller projection of its ground states are $SU(2)$ symmetric. *Ansätze* of these form are particularly suited to describe magnetically disordered phases of matter, such as quantum spin liquids and valence-bond solids [3]. The Gutzwiller-projected BCS wave functions fall back into the class of resonating valence-bond states, first introduced by Anderson in the context of high-Tc superconductivity [58]. The second term of Eq. (6), $\mathcal{H}_{\rm AF}$, contains a magnetic field in the xy-plane which induces magnetic order in the variational *Ansatz*:

$$\mathcal{H}_{\rm AF} = \Delta_{\rm AF}\sum_i \left(e^{iQ\cdot R_i}f^\dagger_{i,\uparrow}f_{i,\downarrow} + e^{-iQ\cdot R_i}f^\dagger_{i,\downarrow}f_{i,\uparrow}\right). \tag{8}$$

The vector Q determines the periodicity of the magnetic order: the Néel phase on the square lattice is obtained by setting $Q = (\pi, \pi)$, while the 120° phase on the triangular lattice corresponds to $Q = (\frac{4\pi}{3}, 0)$ [or, equivalently, $Q = (\frac{2\pi}{3}, \frac{2\pi}{\sqrt{3}})$]. Typically, the wave function obtained by Gutzwiller-projecting the ground state of $\mathcal{H}_{\rm AF}$ overestimates the magnetic order parameter [59]. To improve this construction, further quantum fluctuations can be added by including hopping terms in the BCS Hamiltonian and by applying a two-body Jastrow factor correlator to the fermionic wave function [as in Eq. (3)]:

$$\mathcal{C}_{\rm Jastrow}(\sigma) = \exp\left(\frac{1}{2}\sum_{i,j} v_{i,j}\sigma^z_i\sigma^z_j\right). \tag{9}$$

The Jastrow factor has the effect of adding correlations which are perpendicular to the in-plane magnetic field $\Delta_{\rm AF}$. In this work, we consider Jastrow factors with translationally invariant long-range *pseudopotentials* $v_{i,j} = v(|R_i - R_j|) \in \mathbb{R}$. We emphasize the fact that all the coupling constants defining $\mathcal{H}_0$ (i.e. $t_{i,j}, \Delta_{i,j}, \Delta_{\rm AF}$) and the Jastrow pseudopotentials play the role of variational parameters, which are optimized in order to find the best variational energy. The optimal parametrization of the auxiliary Hamiltonian $\mathcal{H}_0$ depends on the model under investigation.

The central aim of this work is to explore the possibility of improving Gutzwiller-projected fermionic wave functions by applying a stronger many-body correlator than the two-body Jastrow factor. For this purpose, a *neural network* is employed in the form of a restricted Boltzmann machine. This network is defined by introducing a set of auxiliary Ising variables, $\{h^\alpha\}_{\alpha=1,\dots,N_\alpha}$, which form the so-called *hidden* layer. These variables are coupled to the z-components of the spins of the lattice (dubbed as *visible* layer, $\{\sigma^z_i\}_{i=1,\dots,N}$) through a classical energy functional of the form:

$$E_{\rm RBM} = \sum_{i=1}^{N}\sum_{\alpha=1}^{N_\alpha} h^\alpha W^\alpha_i\sigma^z_i + \sum_{\alpha=1}^{N_\alpha} b^\alpha h^\alpha + \sum_{i=1}^{N} a_i\sigma^z_i. \tag{10}$$

The RBM correlator is then obtained by computing the Boltzmann factor $e^{E_{\rm RBM}}$ and taking its trace over the

hidden variables degrees of freedom. This operation can be performed exactly due to the particular form of the classical energy functional of Eq. (10), which only contains interactions between variables belonging to the two different layers (i.e., no intralayer couplings are allowed). Since we consider translationally invariant states with conserved magnetization $S^z_{tot} = 0$, we can set $a_i = 0$ and obtain the final form of the RBM correlator [22]:

$$\mathcal{C}_{\mathrm{RBM}}(\sigma) = \exp\left[\sum_\alpha \log\cosh\left(b^\alpha + \sum_i W_i^\alpha \sigma_i^z\right)\right]. \quad (11)$$

The parameters of $\hat{\mathcal{C}}_{\mathrm{RBM}}$ are called *biases* (b^α) and *weights* (W_i^α), and are (in general) assumed to be complex numbers in this work. The complex parametrization of the RBM allows the correlator to change both the amplitudes and the phases of the fermionic wave function to which it is applied. The expression of Eq. (11) can be regarded as a sort of many-body Jastrow factor, since a series expansion of the $\log\cosh(\dots)$ function contains the n-body terms of the σ^z variables. Unlike the Jastrow factor, the RBM correlator not only breaks the $SU(2)$ symmetry of spin, but also the Z_2 symmetry $\sigma^z \mapsto -\sigma^z$. This happens if the biases b^α are nonzero, since the aforementioned expansion can contain products of odd numbers of spins.

An important question to address is the implementation of lattice symmetries in the correlator $\hat{\mathcal{C}}$. As already pointed out, a symmetric two-body Jastrow factor can be obtained by simply taking a symmetric pseudopotential, e.g., $v_{i,j} = v(|R_i - R_j|)$. This procedure, however, cannot be applied to the RBM correlator, since its parameters depend upon the index α (labelling the hidden units), which does not have any physical meaning. Therefore, the most straightforward way of including symmetries is to implement them *a posteriori*. Concretely, if we want to enforce translational symmetry, we symmetrize the RBM correlator through a product over all possibile Bravais lattice translations $\{\mathcal{T}_R\}$:

$$\begin{aligned}\mathcal{C}_{\mathrm{tRBM}}(\sigma) &= \prod_R C_{\mathrm{RBM}}[\mathcal{T}_R(\sigma)] \\ &= \exp\left[\sum_R \sum_\alpha \log\cosh\left(b^\alpha + \sum_i W_i^\alpha \sigma^z_{i+R}\right)\right]. \quad (12)\end{aligned}$$

The above expression is translationally invariant with momentum $K = (0,0)$. In addition, the point group symmetries $\{\Sigma\}$ of the lattice can be implemented on top of the translationally invariant correlator. The procedure is the same as the one employed in Eq. (12):

$$\begin{aligned}\mathcal{C}_{\mathrm{sRBM}}(\sigma) &= \prod_\Sigma \prod_R C_{\mathrm{RBM}}[\Sigma\mathcal{T}_R(\sigma)] \\ &= \exp\left[\sum_\Sigma \sum_R \sum_\alpha \log\cosh\left(b^\alpha + \sum_i W_i^\alpha \sigma^z_{\Sigma(i+R)}\right)\right], \quad (13)\end{aligned}$$

where $\Sigma(j)$ indicates the position of the site obtained by applying the symmetry Σ to the site j. The quantum numbers of the above correlators, which are associated to the different point group symmetries, are all zero by construction. A more general strategy to implement symmetries with the desired quantum numbers is outlined in Ref. 60.

The main advantage of the RBM correlator with respect to the Jastrow factor comes from the fact that its accuracy can be, in principle, systematically improved by increasing the number of hidden variables N_α. This is due to the fact that a RBM is a universal function approximator [35], which means that it can approximate any function with arbitrary accuracy if the number of variables in the hidden layer is allowed to grow arbitarily large. In addition, the non-local structure of this neural network makes it capable of capturing highly-entangled phases of matter. On the other hand, the disadvantage of employing the RBM correlator in relation to the simpler two-body Jastrow factor mainly resides in its higher computational cost. Crucial for an efficient variational Monte Carlo, computing ratios of translationally invariant RBMs has a cost which scales linearly with the number of sites and the number of hidden units ($O(N \times N_\alpha)$), while computing ratios of two-body Jastrow factors simply scales as $O(1)$ [13]. Finally, another disadvantage of the RBM is the lack of straightforward physical interpretability of its variational parameters, which are associated to many-body spin-spin correlations at all distances. Instead, the pseudopotential $v_{i,j}$ of the Jastrow factor clearly accounts for the two-body correlation of spins i and j, and typically shows a clear physical behavior [61], decaying with the distance $|R_i - R_j|$.

IV. OPTIMIZATION OF THE VARIATIONAL WAVE FUNCTION

In order to minimize the variational energy of the wave function (3), namely

$$E_{\mathcal{C}} = \langle\mathcal{H}\rangle_{\mathcal{C}} = \frac{\langle\Psi_{\mathcal{C}}|\mathcal{H}|\Psi_{\mathcal{C}}\rangle}{\langle\Psi_{\mathcal{C}}|\Psi_{\mathcal{C}}\rangle}, \quad (14)$$

we employ the stochastic reconfiguration (SR) technique [13, 62], which is briefly summarized below.

Let us denote by $\omega = \{\omega_k\}$ the set of all the variational parameters of the RBM-fermionic state, which is formed by the couplings included in the auxiliary Hamiltonian $\mathcal{H}_0$ and by the weights and biases of the RBM correlator. For each parameter ω_k, we can define a corresponding operator $\hat{\mathcal{O}}_k$ that is diagonal in the basis of spin configurations, i.e., $\langle\sigma|\hat{\mathcal{O}}_k|\sigma'\rangle = \mathcal{O}_k(\sigma)\delta_{\sigma,\sigma'}$, and yields the logarithmic derivative of the amplitudes of $|\Psi_{\mathcal{C}}\rangle$:

$$\mathcal{O}_k(\sigma) = \frac{\partial \log\left[\langle\sigma|\Psi_{\mathcal{C}}\rangle\right]}{\partial\omega_k} = \frac{1}{\langle\sigma|\Psi_{\mathcal{C}}\rangle}\frac{\partial\langle\sigma|\Psi_{\mathcal{C}}\rangle}{\partial\omega_k}. \quad (15)$$

Let us also introduce another diagonal operator, the so-

called *local energy* $\hat{E}_{\rm loc}$, whose matrix elements read

$$\langle\sigma|\hat{E}_{\rm loc}|\sigma'\rangle = E_{\rm loc}(\sigma)\delta_{\sigma,\sigma'} = \frac{\langle\sigma|\mathcal{H}|\Psi_{\mathcal{C}}\rangle}{\langle\sigma|\Psi_{\mathcal{C}}\rangle}\delta_{\sigma,\sigma'}. \quad (16)$$

At each step of the SR algorithm, the parameters of the variational wave functions are updated according to

$$\omega_k' = \omega_k + \eta\sum_{k'}\mathcal{S}^{-1}_{k,k'}f_{k'}. \quad (17)$$

Here η is an arbitrary hyperparameter dubbed as the *learning rate*, $\mathcal{S}^{-1}$ is the inverse of the *covariance matrix*

$$\mathcal{S}_{k,k'} = \Re\left[\langle\hat{\mathcal{O}}_k^\dagger\hat{\mathcal{O}}_{k'}\rangle_{\mathcal{C}} - \langle\hat{\mathcal{O}}_k^\dagger\rangle_{\mathcal{C}}\langle\hat{\mathcal{O}}_{k'}\rangle_{\mathcal{C}}\right], \quad (18)$$

and f_k are the *forces* [13]

$$f_k = -\frac{\partial E_{\mathcal{C}}}{\partial\omega_k} = -2\Re\left[\langle\mathcal{H}\hat{\mathcal{O}}_k\rangle_{\mathcal{C}} - \langle\mathcal{H}\rangle_{\mathcal{C}}\langle\hat{\mathcal{O}}_k\rangle_{\mathcal{C}}\right]. \quad (19)$$

Both the elements of the covariance matrix and the forces are computed stochastically by Monte Carlo sampling (through the Metropolis algorithm [63]), and then employed to update the variational parameters as in Eq. (17). We note that a redundant parametrization of the variational wave function causes the covariance matrix to be non-invertible. This situation is commonly encountered when neural network correlators with a large number of variational parameters are considered. To avoid numerical instabilities connected to the singularity of the covariance matrix, we apply an explicit regularization of the form $\mathcal{S}_{k,k'} \mapsto (1+\epsilon\delta_{k,k'})\mathcal{S}_{k,k'}$, where ϵ is an arbitary small parameter [13].

For most of the variational results presented in this work, we performed few distinct optimizations starting from different set of initial parameters, and we selected as optimal wave function the one which provided the lowest variational energy. In general, we used $\approx 2\times10^4$ up to $\approx 2\times10^5$ Monte Carlo samples to evaluate the forces and the covariance matrix of the SR method. We typically chose a learning rate η in the range $[0.02, 0.05]$ and a regularization parameter $\epsilon = 10^{-3}$. For what concerns the latter, in the case of variational wave functions with a large number of variational parameters (e.g. $N_\alpha = 8$ on a 10×10 lattice), we observed that starting the optimization with a larger value of ϵ, e.g. 10^{-2}, helps reducing some numerical instabilities and yields a smoother decrease of the variational energy; then, after few hundreds SR steps, the value of ϵ can be safely changed to 10^{-3} for the rest of the optimization.

V. RESULTS

A. The $J_1 - J_2$ model on the square lattice

We first discuss the variational Monte Carlo results for the $J_1 - J_2$ model on a 6×6 square lattice. For

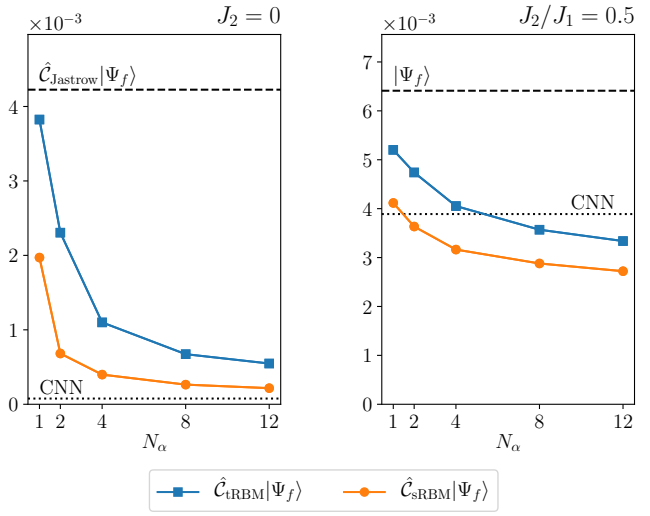

FIG. 1. Relative error of the VMC energies ΔE [see Eq. (20)] with respect to the exact ones for the $J_1 - J_2$ model on the 6×6 square lattice. The results for the unfrustrated case ($J_2 = 0$) and the frustrated one ($J_2/J_1 = 0.5$) are shown on the left and on the right panel, respectively. The relative error of the RBM-fermionic wave function is plotted as a function of the number of hidden units: blue squares refer to the case of translationally invariant RBM correlator $\hat{\mathcal{C}}_{\rm tRBM}$, while orange circles correspond to the fully symmetric RBM correlator $\hat{\mathcal{C}}_{\rm sRBM}$. The error-bars are smaller than the size of the dots. The dashed line represents the relative error of the fermionic wave function of reference, which includes a Jastrow factor in unfrustrated case ($J_2 = 0$). The dotted line refers to the relative error of the CNN quantum state of Ref. 64.

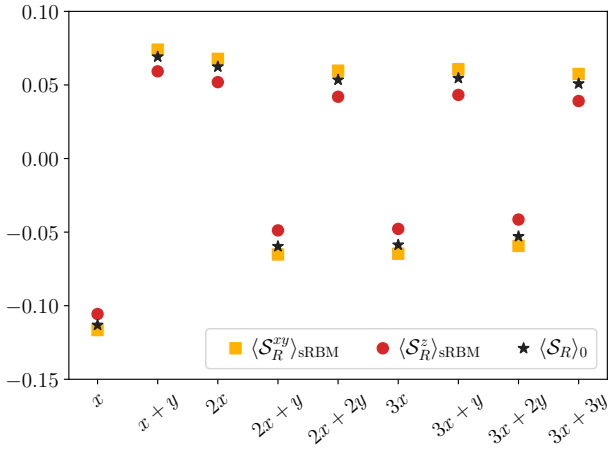

FIG. 2. Spin-spin correlations for the Heisenberg model ($J_2 = 0$) on the 6×6 square lattice, as a function of distance R. Here, $x = (1,0)$ and $y = (0,1)$ are the lattice unit vectors. The set of data represented with yellow squares (red circles) corresponds to the expectation value of $\mathcal{S}_R^{xy}$ ($\mathcal{S}_R^z$) on the RBM-fermionic wave function with the fully symmetric correlator $\hat{\mathcal{C}}_{\rm sRBM}$ ($N_\alpha = 1$). The error-bars are smaller than the size of the dots. The black stars represent the exact value of the spin-spin correlation, i.e., $\langle\mathcal{S}_R\rangle_0 = 1/3\langle\mathbf{S}_0\cdot\mathbf{S}_R\rangle_0$.

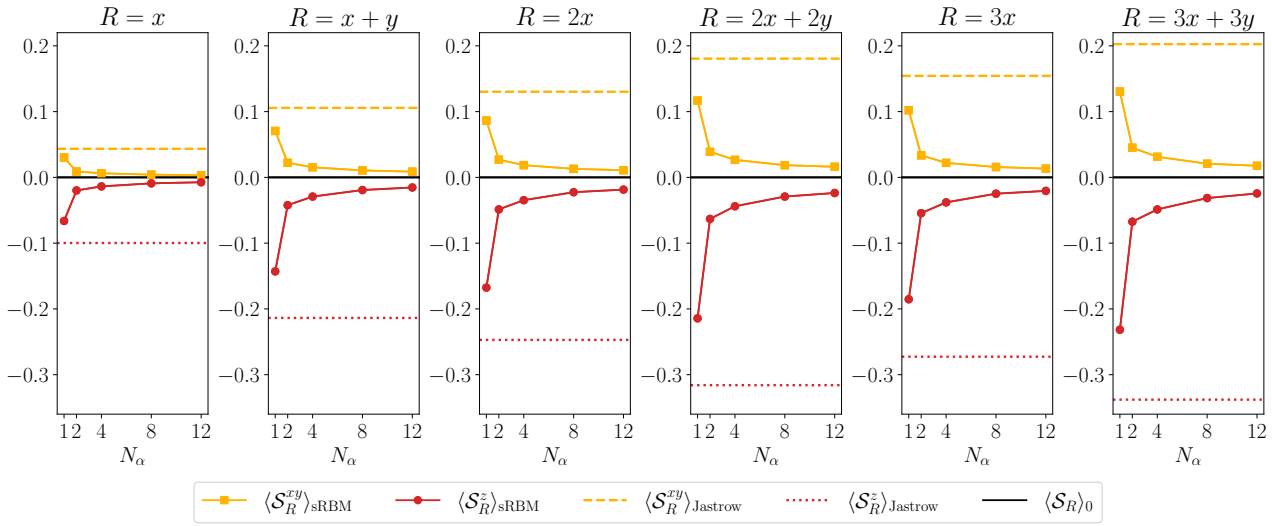

FIG. 3. Relative error of the spin-spin correlations [see Eq. (23)] for the Heinseberg model ($J_2 = 0$) on the 6×6 square lattice, as a function of the number of hidden units. The correlations are evaluated at different distances R, which are expressed in terms of the lattice unit vectors $x = (1, 0)$ and $y = (0, 1)$. The set of data represented with yellow squares (red circles) corresponds to the relative error of the expectation value of $\mathcal{S}_R^{xy}$ ($\mathcal{S}_R^{z}$), computed by employing the RBM-fermionic wave function with the fully symmetric correlator $\hat{\mathcal{C}}_{\text{sRBM}}$. The yellow dashed (red dotted) line, instead, refers to the relative error of the expectation value of $\mathcal{S}_R^{xy}$ ($\mathcal{S}_R^{z}$) computed by employing the Jastrow correlator instead of the RBM. Finally, the black line indicates the zero of the vertical axis, i.e. the position of the exact value in the relative error scale.

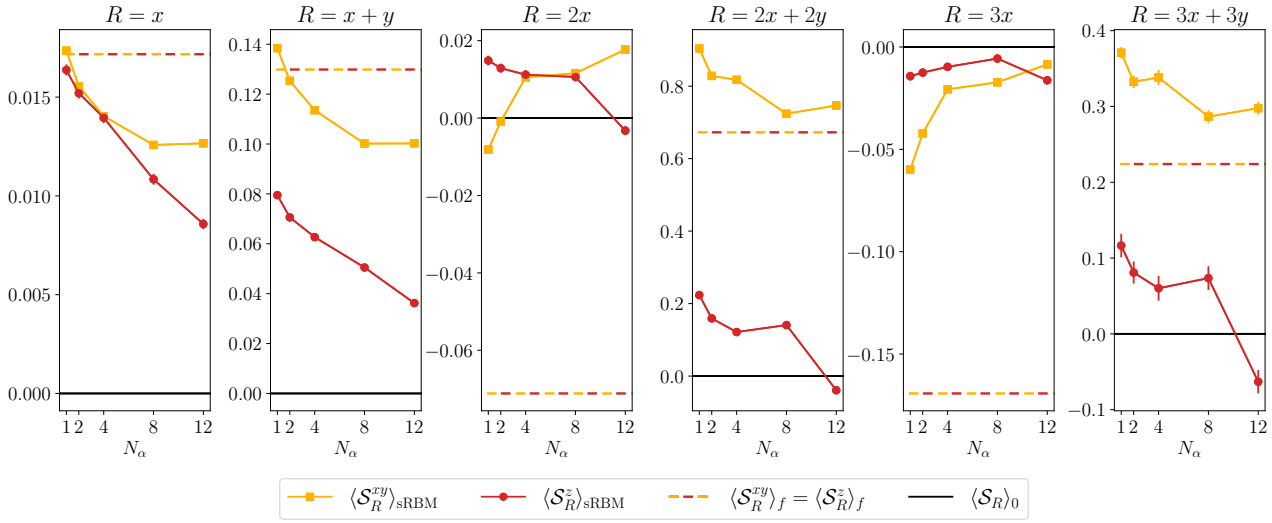

FIG. 4. The same as in Fig. 3 for $J_2/J_1 = 0.5$. The only difference is given by the fact that the fermionic state (without Jastrow factor) is $SU(2)$ invariant and, therefore, in-plane and out-of-plane correlations are equal (and denoted by the bicolor dashed line).

$J_2 = 0$, the auxiliary fermionic Hamiltonian $\mathcal{H}_0$ which defines the variational *Ansatz* contains a Néel magnetic field (with $Q = (\pi, \pi)$) in Eq. (8) and a complex hopping term in Eq. (7), which induces a staggered flux through the square plaquettes [65]. This fermionic wave function possesses the sign structure of the exact ground state of the Heisenberg model, i.e., it follows the so-called Marshall-Peierls sign rule [59, 66]. Therefore, it is sufficient to use an RBM correlator with real parameters, so that $C_{\text{RBM}}(\sigma) \geq 0$. On the contrary, in the frustrated regime, the sign structure of the exact ground state is unknown. For $J_2/J_1 = 0.5$ we combine a complex-valued RBM correlator and a spin liquid fermionic wave function. The latter is obtained by considering an auxiliary Hamiltonian with a uniform real hopping and a $d_{x^2-y^2}$ pairing at first neighbors, and a d_{xy} pairing at fifth neigh-

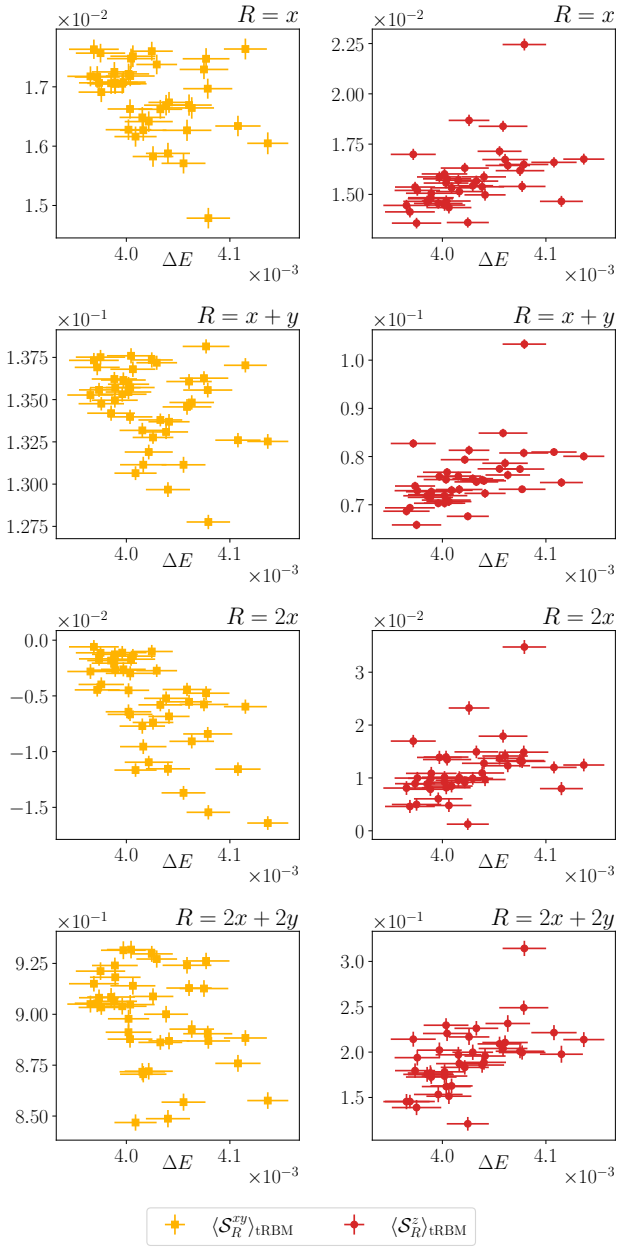

FIG. 5. Relative error of the spin-spin correlations [see Eq. (23)] as a function of the relative error of the variational energy for the $J_1 - J_2$ model on the 6×6 square lattice in the frustrated regime, $J_2/J_1 = 0.5$. The correlations are computed at different distances R, which are expressed in terms of the lattice unit vectors $x = (1, 0)$ and $y = (0, 1)$. The wave function employed in the calculations is a RBM-fermionic *Ansatz* with a translational invariant correlator ($N_\alpha = 4$). Different points correspond to the results of different optimizations of the variational parameters. The set of data represented with yellow squares (red circles) corresponds to the relative error of the expectation value of the in-plane (out-of-plane) correlations.

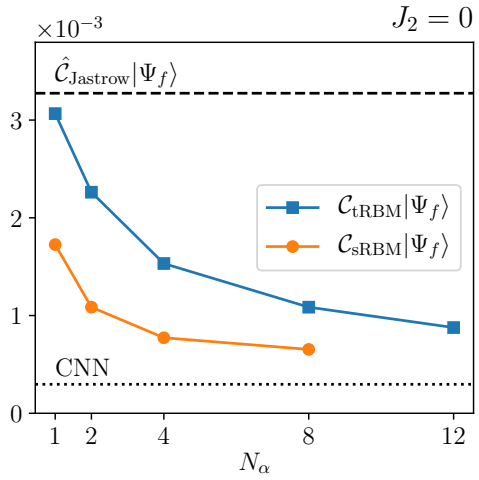

FIG. 6. Relative error of the VMC energies for the Heisenberg model on the 10×10 square lattice, computed with respect to the exact result of quantum Monte Carlo [40, 41]. The relative error of the RBM-fermionic wave function is plotted as a function of the number of hidden units: blue squares refer to the case of translationally invariant RBM correlator $\hat{\mathcal{C}}_{\text{tRBM}}$, while orange circles correspond to the fully symmetric RBM correlator $\hat{\mathcal{C}}_{\text{sRBM}}$. The error bars are smaller than the size of the dots. The dotted line refers to the relative error of the CNN quantum state of Ref. 64.

bors [48]. This wave function corresponds to a Z_2 spin liquid state, which satisfies all the symmetries of the model after the Gutzwiller projection [67]. Once more, we point out that all the parameters in the variational state, i.e., the weights and biases of the RBM and the couplings included in $\mathcal{H}_0$, are fully optimized through the stochastic reconfiguration technique [13, 62].

One of the purposes of this work is evaluating the accuracy gain provided by the use of the RBM correlator. For magnetically ordered phases (e.g., $J_2 = 0$), the best fermionic state breaks the $SU(2)$ symmetry [since Δ_{AF} is finite in Eq. (8)]; in this case, the reference state also contains the Jastrow factor, since it plays an important role to include quantum fluctuations. Instead, within the non-magnetic phase (e.g., $J_2/J_1 = 0.5$), the best fermionic state does not break the $SU(2)$ symmetry (i.e., $\Delta_{\text{AF}} = 0$) and the Jastrow factor typically gives a negligible contribution. In this case, we choose to compare the accuracy of the combined RBM-fermionic wave function to the pure Gutzwiller-projected BCS state. In both cases, the RBM correlator is expected to yield an improvement of the accuracy when all the parameters are properly optimized, since it can be shown that the Jastrow factor can be represented exactly as a RBM [29, 31]. The underlying question we address is understanding to which extent the application of the RBM correlator improves the accuracy and the physical content of the variational wave function. For the $J_1 - J_2$ model, we consider both the translationally invariant [$\hat{\mathcal{C}}_{\text{tRBM}}$, Eq. (12)] and

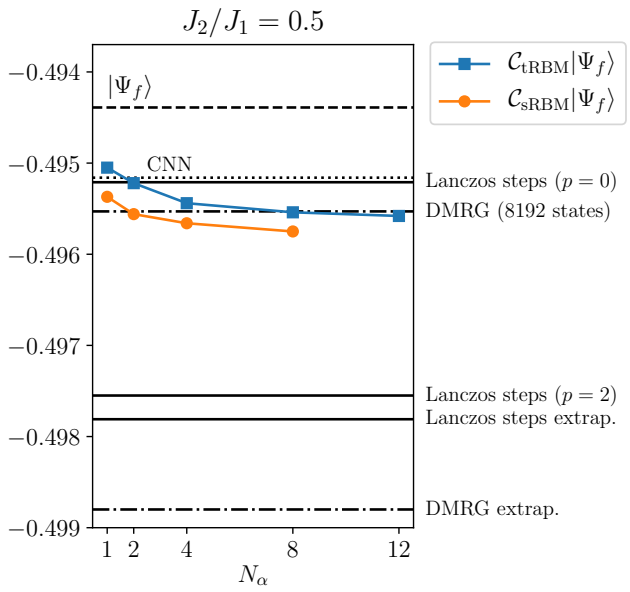

FIG. 7. VMC energies for the $J_1 - J_2$ model on the 10×10 square lattice in the frustrated regime ($J_2/J_1 = 0.5$). The variational energies of the RBM-fermionic wave function are plotted as a function of the number of hidden units: blue squares refer to the case of translationally invariant RBM correlator $\hat{\mathcal{C}}_{\text{tRBM}}$, while orange circles correspond to the fully symmetric RBM correlator $\hat{\mathcal{C}}_{\text{sRBM}}$. The error-bars are smaller than the size of the dots. As a comparison we report several different results. The dashed line indicates the energy of the fermionic wave function of reference. Density-matrix renormalization group (DMRG) energies from Ref. 47 are plotted with a dotted-dashed line: the highest energy corresponds to the most accurate result obtained by a DMRG calculation (using 8192 $SU(2)$ states), while the lowest energy line corresponds to the value which was obtained by extrapolating DMRG data with respect to the truncation error. Full lines represent the results of Ref. 48, in which Lanczos steps were applied to a fermionic wave function in order to improve its accuracy. Three values are reported here: the highest energy is obtained with the pure fermionic wave function ($p = 0$, i.e., no Lanczos steps), the middle one by the application of two Lanczos steps ($p = 2$), while the lowest one is the result of the variance extrapolation. Finally, the variational energy obtained with the CNN quantum state of Ref. 64 is depicted with a dotted line.

the fully symmetric [$\hat{\mathcal{C}}_{\text{sRBM}}$, Eq. (13)] RBM correlators.

In Fig. 1 we report the relative error of the variational energy of the $J_1 - J_2$ model on the 6×6 lattice with respect to the exact value, obtained by Lanczos diagonalization. This quantity is defined as

$$\Delta E = \left| \frac{E_{\mathcal{C}} - E_0}{E_0} \right|, \tag{20}$$

where $E_{\mathcal{C}}$ is the energy of a given variational *Ansatz* and E_0 is the exact ground-state energy. The results clearly show that the variational wave function is overall more accurate in the unfrustrated regime with respect to the frustrated one. In particular, at $J_2 = 0$ the inclusion of the RBM provides a large energy gain with respect to the fermionic reference state ($\Delta E \approx 0.4\%$): the relative error of the energy improves by a factor ≈ 8 when applying the translationally invariant RBM ($\Delta E \approx 0.05\%$ for $N_\alpha = 12$) and a factor of ≈ 20 when applying the fully symmetric RBM ($\Delta E \approx 0.02\%$ for $N_\alpha = 12$). In general, for each value of N_α we observe that the inclusion of the point group symmetries in the RBM correlator halves the relative error of the variational energy. By contrast, at $J_2/J_1 = 0.5$ the accuracy gain is considerably smaller: the relative error of the energy changes from $\Delta E \approx 0.6\%$ to $\Delta E \approx 0.3\%$ by applying a translationally invariant RBM correlator, and the addition of point group symmetries is far less effective than what we observe for the unfrustrated case. In Fig. 1 the variational energies obtained with the RBM-fermionic wave functions are compared to the ones of Ref. 64, in which a convolutional neural network (CNN) quantum state is employed: the CNN wave function is more accurate in the unfrustrated phase, while the RBM-fermionic *Ansatz* gives better energies in the frustrated regime.

To further elucidate the ability of the RBM-fermionic wave functions in capturing the ground state properties of the model, we compute the spin-spin correlations at different distances on the lattice. Whereas the exact ground state wave function of the $J_1 - J_2$ model on finite cluster possesses all the symmetries of its Hamiltonian, most of the variational *Ansätze* used in our calculations break the spin $SU(2)$ symmetry of the model due to the presence of the RBM correlator, the Jastrow factor, or the magnetic field Δ_{AF}. Therefore, to investigate the spin symmetry properties of the resulting states, we separate the computation of the spin-spin correlations at distance R into the in-plane part,

$$\mathcal{S}_R^{xy} \equiv \frac{1}{2N} \sum_{i=1}^{N} \left(S_i^x S_{i+R}^x + S_i^y S_{i+R}^y \right), \tag{21}$$

and the out-of-plane part,

$$\mathcal{S}_R^{z} \equiv \frac{1}{N} \sum_{i=1}^{N} S_i^z S_{i+R}^z. \tag{22}$$

We then compare the variational estimates of the two contributions separately with the exact value $\langle \mathcal{S}_R \rangle_0 \equiv 1/3 \langle \mathbf{S}_0 \cdot \mathbf{S}_R \rangle_0 = \langle S_0^a S_R^a \rangle_0$, $a = x, y, z$. Here $\langle \dots \rangle_0$ indicates the expectation value over the exact ground state $|\Psi_0\rangle$, obtained by Lanczos diagonalization. Thus, the relative error of the spin-spin correlations at distance R is computed as

$$\Delta \mathcal{S}_R^{\alpha} = \frac{\langle \mathcal{S}_R^{\alpha} \rangle_{\mathcal{C}} - \langle \mathcal{S}_R \rangle_0}{\langle \mathcal{S}_R \rangle_0}, \tag{23}$$

where $\langle \dots \rangle_{\mathcal{C}}$ indicates the expectation value over a given variational wave function, and $\alpha = xy, z$ refer to the in-plane, Eq. (21), and out-of plane, Eq. (22), estimators.

In Fig. 2, we show the in-plane and out-of-plane variational correlations as a function of distance, in comparison to the exact value. The results are obtained with the aforementioned RBM-fermionic wave function $\hat{\mathcal{C}}_{\mathrm{sRBM}}|\Psi_f\rangle$ ($N_\alpha = 1$) for $J_2 = 0$. Due to the presence of the antiferromagnetic parameter Δ_{AF}, which induces magnetic ordering in the xy plane, the in-plane correlations overestimate the exact ones in absolute value, while the out-of-plane correlations are underestimated. This tendency is observed for the spin-spin correlations at any distance and for any value of N_α. The results obtained by increasing the number of hidden units are shown in Fig. 3, where the relative error of the variational estimates of the correlations with respect to the exact value is reported for some selected distances R. We observe a systematic improvement of the accuracy when the number of hidden units is increased, with both the in-plane and the out-of-plane correlations approaching the exact value. The fact that these two terms tend to get closer to each other when N_α is increased indicates that the RBM correlator tries to restore the anticipated spin $SU(2)$ symmetry in the wave function. As expected, since the optimal wave function is computed by minimizing the ground state energy, the most accurate values for the correlations are obtained at first-neighbors. We note also that, at any distance, the RBM correlator systematically provides a more accurate estimation of the spin-spin correlations than the simple Jastrow factor.

Before moving to the frustrated case, we explain the motivation of our choice to consider the accuracy of $\mathcal{S}_R^{xy}$ and $\mathcal{S}_R^z$ *separately*, instead of using the isotropic spin-spin correlation, $(\mathcal{S}_R^z + 2\mathcal{S}_R^{xy})/3$. When performing the optimization of the variational parameters of the wave function, we observed that the final set of weights and biases of the RBM can depend on their initial values. Indeed, the parametrization of the RBM displays a considerable degree of redundancy, as indicated by the presence of zero eigenvalues of the covariance matrix of Eq. (18) [13, 62]. Therefore, the optimization procedure can end into different local minima which can have equivalent energies and slightly different in-plane and out-of-plane spin-spin correlations. However, this small difference, which is typically negligible in the unfrustrated regime, can be enhanced or reduced in a random fashion by cancellation of error when the isotropic correlation is computed. Therefore, we argue that a study of the separate components of the correlation function, rather than the isotropic counterpart, provides a better characterization of the symmetry properties of our *Ansatz*.

At $J_2/J_1 = 0.5$, the situation is considerably different from the unfrustrated case, as demostrated in Fig. 4. The accuracy of the variational correlations does not show a systematic improvement with the number of hidden units. Indeed, even if in general the results are more accurate for $N_\alpha = 12$ than $N_\alpha = 1$, the behavior of the relative error is not as smooth as the respective one observed in Fig. 3 for $J_2 = 0$. Moreover, in some cases the relative error obtained by applying the RBM correlator is larger than the one obtained from the use of the simple Jastrow factor. A regular improvement is observed only at first and second neighbors, which are the correlations contributing to the value of the energy. Most importantly, the role of the RBM correlator regarding the expected $SU(2)$ symmetry is not clear; we find that the out-of-plane correlations display a better improvement than the in-plane ones when N_α is increased. We argue that this irregular behaviour of the accuracy of the spin-spin correlations is a consequence of the presence of several local minima in the optimization of the variational parameters of the RBM correlator, which lead to states with markedly different energies and correlation functions. While the effect is already present in the unfrustrated Heinseberg model, it is enhanced in the highly-frustrated regime of the $J_1 - J_2$ model.

To investigate this issue, we performed 40 distinct optimizations of the $\hat{\mathcal{C}}_{\mathrm{tRBM}}|\Psi_f\rangle$ *Ansatz* with $N_\alpha = 4$ at $J_2/J_1 = 0.5$, choosing different values of the RBM parameters as starting point. The relative error of the spin-spin correlations obtained by the 40 resulting wave functions is plotted as a function of the relative error of the variational energy in Fig. 5. We observe that the accuracy of the results show considerably large fluctuations. In particular, while the out-of-plane correlations seem to be more accurate when the variational energy is lower, an opposite effect is observed for the in-plane terms at some distances. The RBM correlator, which is a function of the σ_z degrees of freedom, tends to "sacrifice" the accuracy of the in-plane correlations for the sake of improving the variational energy. This numerical experiment suggests that the complexity of the optimization landscape of RBM in the presence of frustration is significantly different from the unfrustrated case: while in the unfrustrated case all the local minima display similar energies and correlation functions, the highly frustrated regime exhibits a wide array of local minima with similar variational energy but strinkingly different correlation functions. We speculate that these minima are due to the presence of a possible glassy phase in the optimization induced by frustration [68].

We conclude our analysis of the $J_1 - J_2$ model by presenting the variational energies for the 10×10 lattice. The wave functions we employ have the same form of the ones used for the 6×6 lattice. However, while the number of parameters entering $\mathcal{H}_0$ is the same, the number of weights of the RBM increases linearly with the size. For the Heinseberg model ($J_2 = 0$), in Fig. 6 we show the relative error of the variational energy with respect to the exact one, computed by quantum Monte Carlo [40, 41]. The relative error of the RBM-fermionic wave function is of the same order of magnitude of the one obtained on the 6×6 lattice and shows a remarkable energy gain with respect to the Jastrow-fermionic state. However, the accuracy gain provided by the inclusion of the point group symmetries is slightly smaller than the one observed for the 6×6 lattice. We note that the variational energy of the CNN quantum state of Ref. 64 is lower than the

best RBM-fermionic energy found in this work. Here, we emphasize the fact that the CNN state employs a larger number of variational parameters (3838 complex numbers [64]) than the RBM-fermionic wave function with $N_\alpha = 8$ (810 real numbers). However, the local structure of the CNN is advantageous in the process of optimization, since the optimal parameters obtained for a smaller lattice can be employed as a starting point for the optimization of the wave function on a larger lattice [64]. The same procedure cannot be applied in the case of the RBM correlator, due to its highly nonlocal structure, which implies that the optimization of the parameters of this network necessarily becomes harder when the size of the system increases.

In Fig. 7 we compare our variational energies in the frustrated phase, $J_2/J_1 = 0.5$, with several different results from literature. Here, at variance with the unfrustrated case, the variational energies obtained by using the RBM-fermionic wave function are better than the ones of the CNN of Ref. 64, and are very close to the best density matrix renormalization group (DMRG) estimates of Ref. 47. However, a considerably lower variational energy is obtained in Ref. 48, where a fermionic wave function, defined by a BCS Hamiltonian $\mathcal{H}_0$ which contains two additional $d_{x^2-y^2}$ pairings (at fourth- and sixth-neighbors) with respect to the one employed in this work, is improved by the application of few Lanczos steps. The relative energy gain provided by two Lanczos steps, which require the addition of only two variational parameter, is remarkably larger than the improvement which is obtained by the application of the RBM correlator, which contains more than 2000 parameters for $N_\alpha = 12$. In general, as already observed for the 6×6 lattice, in the frustrated regime the RBM correlator yields a much smaller energy gain with respect to the unfrustrated case.

In summary, these results suggest that the RBM correlator provides a systematic way of improving the description of magnetically ordered phases beyond the Jastrow factor, where the RBM effectivetively induces out-of-plane fluctuations that counterbalance the in-plane magnetic order induced by Δ_{AF}. In the frustrated regime, even though the application of the RBM leads to better variational energies, it generally does not improve the description of the correlation functions beyond nearest neighbors. Apart from their numerically expensive training procedure, the RBM's energetic enhancement comes at the price of breaking of the $SU(2)$ symmetry of the fermionic wave function. The symmetry breaking we observe is especially evident in the correlation functions beyond nearest neighbors, which do not directly affect the variational energy during the optimization procedure. Fortunately, all these aspects can be addressed in the future using a recently introduced parametrization of the RBM correlator, which by construction satisfies the $SU(2)$ symmetry of the models considered in our study [69].

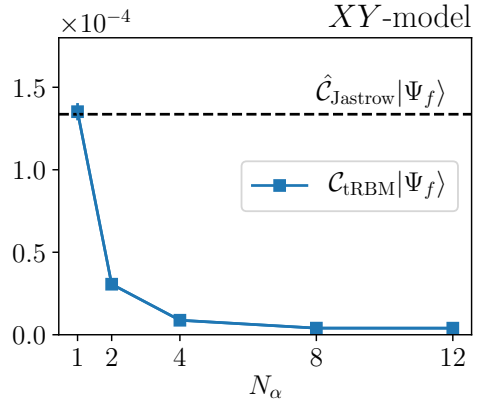

FIG. 8. Relative error of the VMC energies ΔE [see Eq. (20)] with respect to the exact ones for the XY model on the 6×6 square lattice. The blue squares correspond to the relative error of the RBM-fermionic wave function as a function of the number of hidden units (N_α). The error-bars are smaller than the size of the dots. The dashed line represents the relative error of the fermionic wave function of reference, which includes a Jastrow factor.

B. The XY model on the square lattice

As already pointed out, one of the main drawbacks of the introduction of the RBM correlator is the breaking of the spin $SU(2)$ symmetry of the wave function. For this reason, we evaluate the accuracy of the RBM-fermionic construction for a model whose Hamiltonian has lower symmetry, i.e., XY model of Eq. (2). The exact ground state of the XY model has the same sign structure of the one of the Heinseberg model, i.e. it follows the Marshall-Peierls rule [66]. Thus, we employ an analogous RBM-fermionic wave function like the one used for the Heinseberg model (with real weights and biases). The variational results are reported in Fig. 8 for the translationally invariant correlator $\hat{\mathcal{C}}_{\mathrm{tRBM}}$. The relative error of the variational energy with respect to the exact one is at least a factor of 10 smaller than the one obtained for the Heisenberg model and the accuracy gain provided by the RBM correlator is remarkable ($\Delta E \approx 0.0004\%$ for $N_\alpha = 12$). The higher accuracy of the wave function is related to the fact that here the symmetry of the variational *Ansatz* is consistent with the spin symmetry of the model.

C. The Heisenberg model on the triangular lattice

Our previous results suggest the idea that the application of the RBM correlator is more effective for magnetically ordered phases, as exemplified in the Heisenberg and XY model, rather than for non-magnetic ones, as in the frustrated region of the $J_1 - J_2$ model. However, in the Néel phases considered above, the exact sign struc-

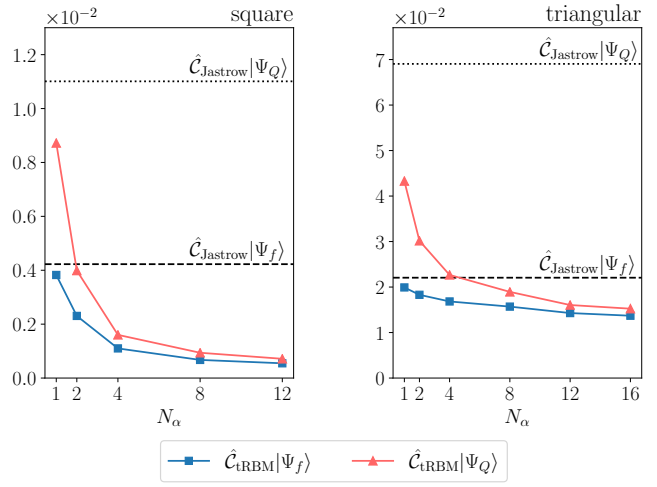

FIG. 9. Relative error of the VMC energies ΔE [see Eq. (20)] with respect to the exact ones for the Heisenberg model on the 6×6 square (left) and triangular (right) lattices. The blue squares (pink triangles) correspond to the relative error of the wave function obtained by applying a translationally invariant RBM correlator to $|\Psi_f\rangle$ ($|\Psi_Q\rangle$), as a function of the number of hidden units N_α. The error-bars are smaller than the size of the dots. The dashed lines represent the relative error of the Jastrow-fermionic wave function $\hat{\mathcal{C}}_{\text{Jastrow}}|\Psi_f\rangle$, while the dotted ones correspond to $\hat{\mathcal{C}}_{\text{Jastrow}}|\Psi_Q\rangle$.

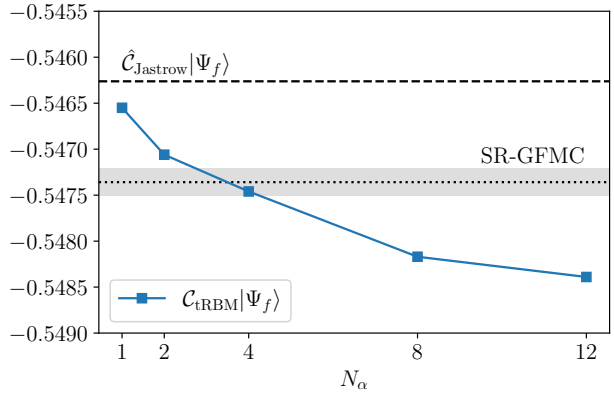

FIG. 10. VMC energies for the Heisenberg model on the 12×12 triangular lattice. The variational energies of the RBM-fermionic wave function are plotted as a function of the number of hidden units (blue squares). The error bars are smaller than the size of the dots. The dashed line indicates the energy of the Jastrow-fermionic wave function of reference. The dotted line corresponds to the Green's function Monte Carlo result of Ref. 56, whose error bar is represented by the grey shaded area.

ture of the wave function is particularly simple and exactly captured by the fermionic part of the variational *Ansatz*. To try to disentangle whether the successes observed in our simulations are related to the special structure of the sign or to the presence of magnetic order, we consider a model whose ground state is magnetically ordered but displays a non-trivial sign structure: the Heisenberg model on the triangular lattice. In this case, the fermionic part of the *Ansatz* is constructed via an auxiliary Hamiltonian $\mathcal{H}_0$ which features a magnetic field Δ_{AF} with pitch vector $Q=(\frac{4\pi}{3},0)$ [or, equivalently, $Q=(\frac{2\pi}{3},\frac{2\pi}{\sqrt{3}})$] and a real nearest-neighbor hopping t. The sign structure of the hopping generates an alternation of 0 and π fluxes threading the triangular plaquettes [19]. While in square lattice a real parametrization on top of the fermionic state already gives an accurate representation for both signs and amplitudes of the exact ground state, we anticipate that the triangular lattice Heisenberg model requires a complex-valued correlator to approximate the unknown sign structure of the wave function induced by the geometric frustration of the problem. Thus, on top of the fermionic state, we apply a translationally invariant RBM correlator with complex parameters ($\hat{\mathcal{C}}_{\text{tRBM}}$).

In Fig. 9, we compare the accuracy of the RBM-fermionic wave function for the Heisenberg model on the 6×6 square and triangular lattices. The energy gain provided by the application of the RBM correlator is considerably larger in the case of the square lattice, where the relative error of the energy decreases of a factor ≈ 8, with respect to the case of the triangular lattice, where it decreases of a factor ≈ 1.5 (from $\Delta E\approx 2.2\%$ to $\Delta E\approx 1.4\%$). Overall, the variational energy is more accurate on the square lattice than on the triangular lattice, and the relative errors differ by an order of magnitude.

In Fig. 9, we also compare the results of the aforementioned RBM-fermionic wave functions to the ones obtained by simpler *Ansätze*, which are constructed by setting the hopping terms to zero and considering an auxiliary fermionic Hamiltonian with only magnetic field ($\mathcal{H}_0=\mathcal{H}_{\text{AF}}$). In this way, the fermionic degrees of freedom are localized and $|\Psi_f\rangle$ reduces to a (projected) product state $|\Psi_Q\rangle=\mathcal{P}_{S^z_{tot}=0}\prod_{i=1}^{N}(|\uparrow\rangle_i+e^{iQR_i}|\downarrow\rangle_i)$, which displays "classical" order in the xy-plane. This wave function can be employed as a reference state for the application of a Jastrow factor [34] or the RBM correlator. Here, we apply a translationally invariant RBM correlator to $|\Psi_Q\rangle$. We observe that the energy gain provided by the presence of the hopping term in the fermionic *Ansatz* is remarkably large when the simple Jastrow factor is applied to the reference state. However, when the RBM correlator is employed, the contribution of the hopping term becomes less important, decreasing considerably with the number of hidden units, which suggests that the RBM replaces the effect of the fermionic hopping term in the state.

Finally, in Fig. 10 we present the variational energies obtained for the Heisenberg model on the 12×12 triangular lattice. For a large enough N_α, the variational energy of the RBM-fermionic wave function is more accurate than the Green's function Monte Carlo (GFMC)

result of Ref. 56. To summarize our numerical experiments, we surmise that the high accuracy of the results for the unfrustrated square lattice Heisenberg model is due to having an exact representation of the sign structure built in the Ansatz, which in turn, alleviates the energy optimization problem. On the other hand, the lower accuracy in the triangular lattice model is presumably due to a combination of the approximate nature of the sign structure imposed by our Ansatz and the energy optimization problem in the presence of complex numbers. Since magnetically ordered states are pervasive in frustrated magnetism, it remains an important issue to unequivocally establish whether the origin of the high accuracy of the results for the Heisenberg model on the square lattice is only due to the absence of frustration or to the fact that the sign structure is exactly known in the unfrustrated case.

VI. CONCLUSION

Inspired by Hinton's product of experts idea [32], we have studied a neural augmentation of the parton construction that combines a family of physically motivated Gutzwiller-projected variational states with a complex-valued RBM correlator. We focused our attention on various prototypical spin models traditionally used in condensed matter physics which exhibit a wide array of conventional ground states with magnetic order, as well as more exotic ones where a spin-liquid behavior has been anticipated.

In agreement with previous results based upon neural networks alone [22, 64], our calculations showed that RBMs are very effective in the unfrustrated Heisenberg and XY models on the square lattice, where the knowledge of the exact sign structure of the ground state allows us to use a real parametrization of the RBM. Here, a few hidden units in the neural network are sufficient to reach a striking accuracy, which for the Heisenberg case is comparable with the best variational wave functions defined within the bosonic resonating valence-bond picture [70]. Moreover, we emphasize the remarkable ability of the RBM to systematically recover the spin $SU(2)$ symmetry upon increasing the number of hidden units. This is clear not only by looking at first-neighbor correlations, but also at further distances: both out-of-plane correlations (i.e., along z) and in-plane ones (i.e., in the xy plane) exhibit a clear tendendcy to converge toward the exact (isotropic) result upon increasing N_α.

In the highly frustrated regime, the exact ground-state sign is not known *a priori*, even when the system displays magnetic order, like in the Heisenberg model on the triangular lattice. Here, we find a substantial energy gain with respect to the original parton wave function, although the variational procedure does not yield the same accuray as in the square lattice, even for a relatively large number of hidden units. A similar effect is also observed in the highly-frustated regime of the $J_1 - J_2$ Heisenberg model, where for $J_2/J_1 = 0.5$ our neural Gutzwiller-projected wave functions exhibit accuracies beyond recent neural network calculations based on CNNs [64], though with a lower accuracy than in the $J_2 = 0$ case. Based on the numerical experiments presented in Fig. 5, we surmise that the application of neural network variational states to frustrated quantum spin systems requires an extensive investigation of the intricate relation between the representation power of neural networks to capture highly-entangled states of matter with a complicated sign structure and the complexity of the optimization landscape of the problem induced by frustration. Motivated by Ref. 71, the disentangling of these factors could be approached through a clustering analysis of the trained RBM parameters and their associated spin-spin correlation functions, which may shed light onto the interplay between frustration, sign structure and entanglement [72], and the rough optimization landscape of the problem [68].

At variance with the unfrustrated limit, the $SU(2)$ symmetry is broken and hardly recovered when increasing N_α in the frustrated case: the out-of-plane correlations are clearly more accurate than in-plane ones, highligthing the "asymmetry" of the RBM form, which is defined in terms of the z-component of the local spin. In this regard, our calculations provide a first investigation into the *inductive bias* of the RBM applied to ground states of frustrated systems, where we observe a strong tendency to break $SU(2)$ symmetry in our setting. Thus, the recent proposal to generalize the RBM to fulfill the $SU(2)$ symmetry [69] may pave the way for studying fully symmetric neural Gutzwiller-projected wave functions, which will be especially relevant in the understanding of models for spin liquid phases and other states exhibiting fractionalized excitations and gauge structures.

ACKNOWLEDGMENTS

We would like to thank K. Choo, G. Carleo, G. Torlai, and R. Melko for useful discussions. F.F. acknowledges the kind hospitality and financial support of the Vector Institute for Artificial Intelligence in Toronto, where this project started. We also thank the Kavli Institute for Theoretical Physics (KITP) in Santa Barbara and the program "Machine Learning for Quantum Many-Body Physics". This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958. J.C. acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada CIFAR AI chair program.

[1] D. C. Tsui, H. L. Stormer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
[2] L. Balents, Nature **464**, 199 (2010).
[3] L. Savary and L. Balents, Rep. Prog. Phys. **80**, 016502 (2016).
[4] X. Wen, *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons*, Oxford Graduate Texts (Oxford University Press, 2004).
[5] G. Baskaran and P. W. Anderson, Phys. Rev. B **37**, 580 (1988).
[6] A. Y. Kitaev, Ann. Phys. **303**, 2 (2003).
[7] M. Hermanns, I. Kimchi, and J. Knolle, Ann. Rev. Condens. Mat. Phys. **9**, 17 (2018).
[8] M. A. Levin and X.-G. Wen, Phys. Rev. B **71**, 045110 (2005).
[9] L. Cincio and G. Vidal, Phys. Rev. Lett. **110**, 067208 (2013).
[10] S. Isakov, M. Hastings, and R. Melko, Nat. Phys. **7**, 772 (2011).
[11] C. Hickey and S. Trebst, Nat. Comm. **10**, 530 (2019).
[12] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath, Phys. Rev. B **85**, 235151 (2012).
[13] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017).
[14] D. Sheng, O. Motrunich, and M. Fisher, Phys. Rev. B **79**, 205112 (2009).
[15] Y. Iqbal, F. Becca, and D. Poilblanc, Phys. Rev. B **84**, 020407 (2011).
[16] Y. Iqbal, F. Becca, S. Sorella, and D. Poilblanc, Phys. Rev. B **87**, 060405 (2013).
[17] S.-S. Gong, D. Sheng, O. Motrunich, and M. Fisher, Phys. Rev. B **88**, 165138 (2013).
[18] W.-J. Hu, S.-S. Gong, W. Zhu, and D. Sheng, Phys. Rev. B **92**, 140403 (2015).
[19] Y. Iqbal, W.-J. Hu, R. Thomale, D. Poilblanc, and F. Becca, Phys. Rev. B **93**, 144411 (2016).
[20] T. Tay and O. Motrunich, Phys. Rev. B **83**, 235122 (2011).
[21] Z.-X. Liu, J.-W. Mei, P. Ye, and X.-G. Wen, Phys. Rev. B **90**, 235146 (2014).
[22] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[23] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).
[24] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).
[25] G. Torlai and R. Melko, ArXiv e-prints , arXiv:1801.09684 (2018).
[26] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).
[27] X. Gao and L.-M. Duan, Nat. Comm. **8**, 662 (2017).
[28] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B **96**, 195145 (2017).
[29] Y. Nomura, A. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017).
[30] S. Clark, J. Phys. A: Math. and Theor. **51**, 135301 (2018).
[31] I. Glasser, N. Pancotti, M. August, I. Rodriguez, and J. Cirac, Phys. Rev. X **8**, 011006 (2018).
[32] G. Hinton, Neural Comput. **14**, 1771 (2002).
[33] J. Toulouse and C. Umrigar, J. Chem. Phys. **128**, 174101 (2008).
[34] J. Carrasquilla, A. di Ciolo, F. Becca, V. Galitski, and M. Rigol, Phys. Rev. B **88**, 241109 (2013).
[35] N. Le Roux and Y. Bengio, Neural Comput. **20**, 1631 (2008).
[36] J. Daunt and R. Smith, Rev. Mod. Phys. **26**, 172 (1954).
[37] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, J. Am. Chem. Soc. **130**, 3296 (2008).
[38] J. Paddison, M. Daum, Z. Dun, G. Ehlers, Y. Liu, M. Stone, H. Zhou, and M. Mourigal, Nat. Phys. **13**, 117 (2017).
[39] B. Dalla Piazza, M. Mourigal, N. Christensen, G. Nilsen, P. Tregenna-Piggott, T. Perring, M. Enderle, D. McMorrow, D. Ivanov, and H. M. Ronnow, Nat. Phys. **11**, 62 (2015).
[40] A. Sandvik, Phys. Rev. B **56**, 11678 (1997).
[41] M. Calandra Buonaura and S. Sorella, Phys. Rev. B **57**, 11446 (1998).
[42] N. Read and S. Sachdev, Phys. Rev. Lett. **62**, 1694 (1989).
[43] H. Schulz, T. Ziman, and D. Poilblanc, J. Phys. I France **6**, 675 (1996).
[44] D. Poilblanc and M. Mambrini, Phys. Rev. B **96**, 014414 (2017).
[45] R. Haghshenas and D. Sheng, Phys. Rev. B **97**, 174408 (2018).
[46] M. Mambrini, A. Läuchli, D. Poilblanc, and F. Mila, Phys. Rev. B **74**, 144422 (2006).
[47] S.-S. Gong, W. Zhu, D. Sheng, O. Motrunich, and M. Fisher, Phys. Rev. Lett. **113**, 027201 (2014).
[48] W.-J. Hu, F. Becca, A. Parola, and S. Sorella, Phys. Rev. B **88**, 060402 (2013).
[49] S. Morita, R. Kaneko, and M. Imada, J. Phys. Soc. Jpn. **84**, 024720 (2015).
[50] L. Wang and A. Sandvik, Phys. Rev. Lett. **121**, 107202 (2018).
[51] T. Giamarchi, C. Rüegg, and O. Tchernyshyov, Nat. Phys. **4**, 198 (2008).
[52] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
[53] J. Carrasquilla and M. Rigol, Phys. Rev. A **86**, 043629 (2012).
[54] Y. Shimizu, K. Miyagawa, K. Kanoda, M. Maesato, and G. Saito, Phys. Rev. Lett. **91**, 107001 (2003).
[55] P. Anderson, Mat. Res. Bull. **8**, 153 (1973).
[56] L. Capriotti, A. Trumper, and S. Sorella, Phys. Rev. Lett. **82**, 3899 (1999).
[57] S. White and A. Chernyshev, Phys. Rev. Lett. **99**, 127004 (2007).
[58] P. Anderson, Science **235**, 1196 (1987).
[59] F. Becca, L. Capriotti, A. Parola, and S. Sorella, in *Introduction to Frustrated Magnetism: Materials, Experiments, Theory*, Springer Series in Solid-State Sciences, edited by C. Lacroix, P. Mendels, and F. Mila (Springer (Berlin), 2011) p. 379.
[60] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Phys. Rev. Lett. **121**, 167204 (2018).
[61] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, Phys. Rev. Lett. **94**, 026406 (2005).
[62] S. Sorella, Phys. Rev. B **71**, 241103 (2005).

[63] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, The Journal of Chemical Physics **21**, 1087 (1953), https://doi.org/10.1063/1.1699114.
[64] K. Choo, T. Neupert, and G. Carleo, ArXiv e-prints , arXiv:1903.06713 (2019).
[65] I. Affleck and J. Marston, Phys. Rev. B **37**, 3774 (1988).
[66] W. Marshall and R. Peierls, Proc. Royal Soc. Lon. Series A. Math. and Phys. Sciences **232**, 48 (1955).
[67] X.-G. Wen, Phys. Rev. B **65**, 165113 (2002).
[68] M. Baity-Jesi, L. Sagun, M. Geiger, S. Spigler, G. B. Arous, C. Cammarota, Y. LeCun, M. Wyart, and G. Biroli, in *Proceedings of the 35th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 80, edited by J. Dy and A. Krause (PMLR, Stockholmsmssan, Stockholm Sweden, 2018) pp. 314–323.
[69] T. Vieijra, C. Casert, J. Nys, W. De Neve, J. Haegeman, J. Ryckebusch, and F. Verstraete, ArXiv e-prints , arXiv:1905.06034 (2019).
[70] S. Liang, B. Doucot, and P. Anderson, Phys. Rev. Lett. **61**, 365 (1988).
[71] A. G. R. Day, M. Bukov, P. Weinberg, P. Mehta, and D. Sels, Physical Review Letters **122**, 020601 (2019).
[72] T. Grover and M. P. A. Fisher, Physical Review A **92**, 042308 (2015).