

Article

Approximating Ground States by Neural Network Quantum States

Ying Yang^{1,2}, Chengyang Zhang¹ and Huaixin Cao^{1,*} 

¹ School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710119, China; yangyingyy@snnu.edu.cn (Y.Y.); zhangcy@snnu.edu.cn (C.Z.)

² School of Mathematics and Information Technology, Yuncheng University, Yuncheng 044000, China

* Correspondence: caohx@snnu.edu.cn

Received: 16 December 2018; Accepted: 16 January 2019; Published: 17 January 2019



Abstract: Motivated by the Carleo's work (Science, 2017, 355: 602), we focus on finding the neural network quantum states approximation of the unknown ground state of a given Hamiltonian H in terms of the best relative error and explore the influences of sum, tensor product, local unitarity of Hamiltonians on the best relative error. Besides, we illustrate our method with some examples.

Keywords: approximation; ground state; neural network quantum state

1. Introduction

The quantum many-body problem is a general name for a vast category of physical problems pertaining to the properties of microscopic systems made of a large number of interacting particles. In such a quantum system, the repeated interactions between particles create quantum correlations [1–3], quantum entanglement [4–6], Bell nonlocality [7–9], Einstein-Podolsky-Rosen (EPR) steering [10–12]. As a consequence, the wave function of the system is a complicated object holding a large amount of information, which usually makes exact or analytical calculations impractical or even impossible. Thus, many-body theoretical physics most often relies on a set of approximations specific to the problem at hand, and ranks among the most computationally intensive fields of science. Gordon [13] indicated that if one knows an accurate energy for the ground state (say from experiment), then one can construct a sequence of upper and lower bounds to the overlap between the approximate function and the true (but unknown) ground-state wave function. Wang [14] constructed a set of intermediate resolvents from the intermediate Hamiltonians introduced by Weinstein. From these intermediate resolvents they obtained a new formula for the lower bound of the overlap between the approximate and exact wave functions of a quantum-mechanical system. Merkel [15] proposed a method and tested it for approximating the integrals over H^2 and H^3 required by Weinhold's technique as products of integrals involving only H . Cioslowski [16] constructed a connected-moments expansion for the overlap between the approximate and the exact (but unknown) wave function of the ground state. Hornik [17] shown that LRV's algorithm, with some relatively trivial modifications, can estimate the overlap of an approximate with the exact wave function. Marmorino [18] derived two methods from the t expansion of Horn and Weinstein to bound from above the magnitude of the overlap of an approximate wavefunction with the ground state. Nomura [19] developed a machine learning method to construct accurate ground-state wave functions of strongly interacting and entangled quantum spin as well as fermionic models on lattices.

Artificial neural networks are important tools in machine learning due to their efficient approximation ability [20–24]. Especially, Roux [25] proved that restricted Boltzmann machines are universal approximators of discrete distributions.

Applying neural networks in solving the quantum many-body problem, Carleo and Troyer in [26] demonstrated the remarkable power of a reinforcement learning approach in calculating the ground state or simulating the unitary time evolution of complex quantum systems with strong interactions. Their idea consists in using neural networks as variational wave functions to approximate ground states of many-body quantum systems. In this direction, the networks are trained or optimized by the standard variational Monte Carlo method while a few different neural-network architectures were tested [26–29], and the most promising results so far have been achieved with Boltzmann machines [29]. In particular, state-of-the-art numerical results have been obtained on popular models with restricted Boltzmann machines (RBM), and recent effort has demonstrated the power of deep Boltzmann machines to represent ground states of many-body Hamiltonians with polynomial-size gap and quantum states generated by any polynomial size quantum circuits [30,31]. Deng et al. [32] show that the RBM can be used to describe topological states and constructed exact representations for symmetry-protected topological states and intrinsic topologically ordered states. Glasser et al. [33] show that there are strong connections between neural network quantum states in the form of RBM and some classes of tensor-network states in arbitrary dimensions and obtain that neural network quantum states and their string-bond-state extension can describe a lattice fractional quantum Hall state exactly. Gardas, Rams and Dziarmaga [34] show that the approach of Carleo and Troyer [26] can be realized experimentally using quantum annealers and they conducted experimental simulations of many-body quantum systems using a hybrid classical-quantum algorithm. Cai and Liu in [35] demonstrated the expressibility of artificial neural networks in quantum many-body physics by showing that a feed-forward neural network with a small number of hidden layers can be trained to approximate with high precision the ground states of some notable quantum many-body systems. In [36], Saito and Kato developed a variational method to obtain many-body ground states of the Bose-Hubbard model using feed forward artificial neural networks and proved that many-body ground states with different numbers of atoms can be generated by a single network. By employing the formalism of tensor networks, Clark [37] show that neural network quantum states given in [26] are a special form of correlator product states.

Despite such exciting developments, it is unknown whether a general state can be expressed by neural networks efficiently. Recently, by generalizing the idea of [26], we introduced in [38] neural networks quantum states (NNQSs) based on general input observables and explored some related properties about NNQSs. Secondly, we established some necessary and sufficient conditions for a general graph state to be represented by an NNQS.

In this paper, based on the NNQSs introduced in [38], we focus on finding the NNQS approximation of the unknown ground state of a given Hamiltonian H . The remaining part of this paper is organized as follows. In Section 2, we recall the concept and the related properties of NNQSs introduced in [38]. In Section 3, we explore the NNQS approximation of the unknown ground state of a given Hamiltonian H in terms of the best relative error and consider the influence of sum, tensor product, local unitary of Hamiltonian on the best relative error. Besides, we illustrate our method with some examples.

2. Neural Network Quantum States

To start with, let us recall the concept and the related properties of NNQSs introduced in [38]. Let Q_1, Q_2, \dots, Q_N be N quantum systems with state spaces $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_N$ of dimensions d_1, d_2, \dots, d_N , respectively. We consider the composite system Q of Q_1, Q_2, \dots, Q_N with state space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_N$.

Let S_1, S_2, \dots, S_N be non-degenerate observables of systems Q_1, Q_2, \dots, Q_N , respectively. Then $S = S_1 \otimes S_2 \otimes \dots \otimes S_N$ is an observable of the composite system Q . Use $\{|\psi_{k_j}\rangle\}_{k_j=0}^{d_j-1}$ to denote the eigenbasis of S_j corresponding to eigenvalues $\{\lambda_{k_j}\}_{k_j=0}^{d_j-1}$. Thus,

$$S_j|\psi_{k_j}\rangle = \lambda_{k_j}|\psi_{k_j}\rangle (k_j = 0, 1, \dots, d_j - 1). \tag{1}$$

It is easy to check that the eigenvalues and corresponding eigenbasis of $S = S_1 \otimes S_2 \otimes \dots \otimes S_N$ are

$$\lambda_{k_1}\lambda_{k_2}\dots\lambda_{k_N} \text{ and } |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \dots \otimes |\psi_{k_N}\rangle (k_j = 0, 1, \dots, d_j - 1), \tag{2}$$

respectively. Put

$$V(S) = \left\{ \Lambda_{k_1 k_2 \dots k_N} \equiv (\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N})^T : k_j = 0, 1, \dots, d_j - 1 \right\},$$

called an input space. For parameters

$$a = (a_1, a_2, \dots, a_N)^T \in \mathbb{C}^N, b = (b_1, b_2, \dots, b_M)^T \in \mathbb{C}^M, W = [W_{ij}] \in \mathbb{C}^{M \times N},$$

write $\Omega = (a, b, W)$ and put

$$\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) = \sum_{h_i = \pm 1} \exp \left(\sum_{j=1}^N a_j \lambda_{k_j} + \sum_{i=1}^M b_i h_i + \sum_{i=1}^M \sum_{j=1}^N W_{ij} h_i \lambda_{k_j} \right). \tag{3}$$

Then we obtain a complex-valued function $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N})$ of the input variable $\Lambda_{k_1 k_2 \dots k_N}$. We call it a *neural network quantum wave function* (NNQWF). Then we define

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1 k_2 \dots k_N} \in V(S)} \Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \dots \otimes |\psi_{k_N}\rangle, \tag{4}$$

which is a nonzero vector (not necessarily normalized) of the Hilbert space \mathcal{H} . We call it a *neural network quantum state* (NNQS) induced by the parameter $\Omega = (a, b, W)$ and the input observable $S = S_1 \otimes S_2 \otimes \dots \otimes S_N$ (Figure 1).

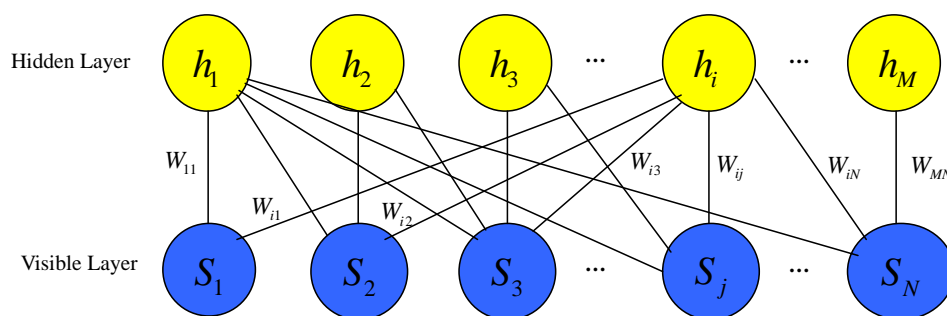


Figure 1. Artificial neural network encoding an NNQS. It is a restricted Boltzmann machine architecture that features a set of N visible artificial neurons (blue disks) and a set of M hidden neurons (yellow disks). For each value $\Lambda_{k_1 k_2 \dots k_N}$ of the input observable S , the neural network computes the value of the $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N})$.

Note that we do not assume that an NNQWF satisfies the normalization condition:

$$\sum_{\Lambda_{k_1 k_2 \dots k_N}} |\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N})|^2 = 1.$$

Indeed, when a state can be written as NNQS, this normalization condition is automatically satisfied.

By definition, an NNQWF can be reduced to

$$\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) = \prod_{j=1}^N e^{a_j \lambda_{k_j}} \cdot \prod_{i=1}^M 2 \cosh \left(b_i + \sum_{j=1}^N W_{ij} \lambda_{k_j} \right). \tag{5}$$

It can be described by the following “quantum artificial neural network” (Figure 2) where $a = 0$, $2 \cosh(z) = e^z + e^{-z}$, \sum_{b_i} and Π are functions such that

$$\sum_{b_i}(x_1, x_2, \dots, x_N) = b_i + \sum_{j=1}^N x_j, \quad \Pi(y_1, y_2, \dots, y_M) = \prod_{i=1}^M y_i.$$

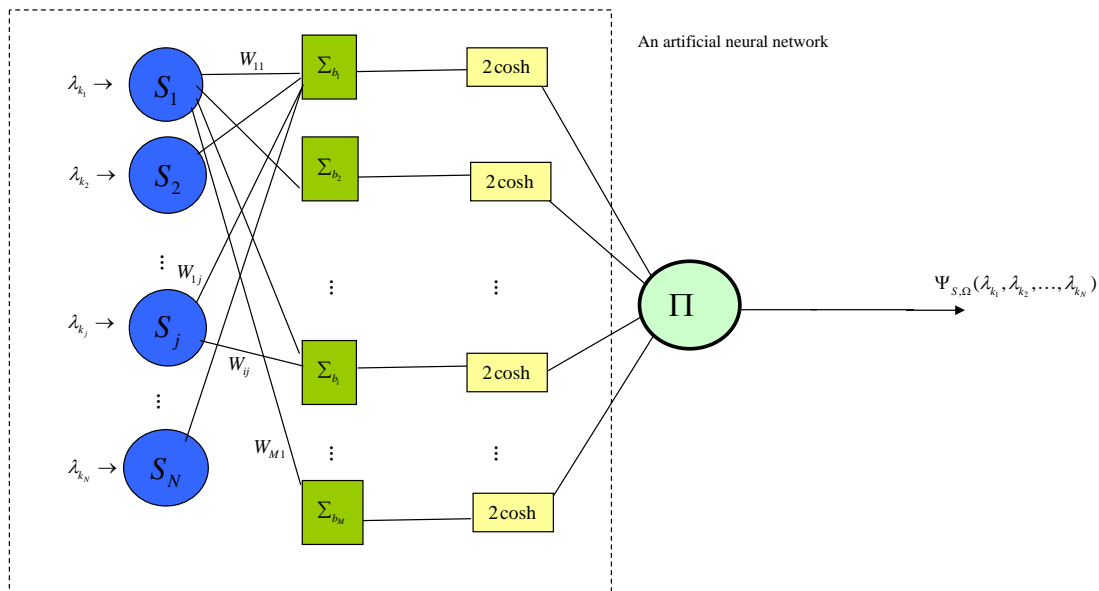


Figure 2. Quantum artificial neural network with parameter $\Omega = (0, b, W)$.

We call this network a quantum artificial neural network because its input eigenvalues of quantum observables and the outcomes are values of an NNQWF, while it has a network structure similar to a usual artificial neural network.

Next, let us consider the tensor product of the two NNQSs. We have proved the following.

Proposition 1. [38] Suppose that $|\Psi'_{S',\Omega'}\rangle$ and $|\Psi''_{S'',\Omega''}\rangle$ are two NNQSs with parameters

$$S' = S'_1 \otimes \dots \otimes S'_{N'}, S'' = S''_1 \otimes \dots \otimes S''_{N''}, \Omega' = (a', b', W'), \Omega'' = (a'', b'', W''),$$

respectively. Then $|\Psi'_{S',\Omega'}\rangle \otimes |\Psi''_{S'',\Omega''}\rangle$ is also an NNQS $|\Phi_{S,\Omega}\rangle$ with parameters

$$S = S' \otimes S'', \Omega = (a, b, W), N = N' + N'', M = M' + M'',$$

$$a = \begin{pmatrix} a' \\ a'' \end{pmatrix}, b = \begin{pmatrix} b' \\ b'' \end{pmatrix}, W = [W_{ij}] = \begin{pmatrix} W'_{M' \times N'} & 0 \\ 0 & W''_{M'' \times N''} \end{pmatrix}.$$

Now, we discuss the influence of local unitary operation (LUO) on an NNQS. We conclude this as follows.

Proposition 2. [38] Suppose that $|\Psi_{S,\Omega}\rangle$ is an NNQS and $U = U_1 \otimes U_2 \otimes \dots \otimes U_N$ is a local unitary operator on \mathcal{H} . Then $U|\Psi_{S,\Omega}\rangle = |\Psi_{USU^\dagger,\Omega}\rangle$, which is also an NNQS with the input observable USU^\dagger and the parameter Ω , and has the same NNQWF as $|\Psi_{S,\Omega}\rangle$.

Remark 1. It can be seen from Proposition 2 that if two pure states are LU-equivalent and an NNQS representation of one of the two states is easily given, then that of another state can be obtained from that of the former.

To conclude this section, we discuss a special class of NNQSSs.

When $S = \sigma_1^z \otimes \sigma_2^z \otimes \dots \otimes \sigma_N^z$, we have

$$\lambda_{k_j} = \begin{cases} 1, & k_j = 0 \\ -1, & k_j = 1 \end{cases}, |\psi_{k_j}\rangle = \begin{cases} |0\rangle, & k_j = 0 \\ |1\rangle, & k_j = 1 \end{cases} \quad (1 \leq j \leq N),$$

and $V(S) = \{1, -1\}^N$.

In this case, the NNQS (4) becomes

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1 k_2 \dots k_N} \in \{1, -1\}^N} \Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \dots \otimes |\psi_{k_N}\rangle. \quad (6)$$

This leads to the NNQS induced in [26] and discussed in [32]. We call such an NNQS a *spin-z NNQS*.

3. Approximating Ground States by Neural Network Quantum States

In this section, we try to find approximate solution to the static Schrödinger equation $H|\psi\rangle = E|\psi\rangle$ for a given Hamiltonian H . For example, to find approximation of ground states by neural network quantum states.

Let $|\Psi_{S,\Omega}\rangle$ be an NNQS given by Equation (4) and let H be a Hamiltonian whose smallest eigenvalue E_{exact} is not zero. Put

$$E_H(S, \Omega) = \frac{\langle \Psi_{S,\Omega} | H | \Psi_{S,\Omega} \rangle}{\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle}.$$

We seek the minimum relative error between $E_H(S, \Omega)$ and E_{exact} over Ω ,

$$\epsilon = \min_{\Omega} \frac{|E_H(S, \Omega) - E_{exact}|}{|E_{exact}|}. \quad (7)$$

We call ϵ the *best relative error* between $E_H(S, \Omega)$ and E_{exact} .

Obviously, when the minimum in Equation (7) is attained at the parameter Ω , we can use the normalized NNQS $\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle^{-\frac{1}{2}} |\Psi_{S,\Omega}\rangle$ as an approximation of the ground state of H with the best relative error ϵ .

Generally, $E_H(S, \Omega) \geq E_{exact}$. Hence, ϵ can also be expressed as

$$\epsilon = \min_{\Omega} \frac{E_H(S, \Omega) - E_{exact}}{|E_{exact}|}.$$

Next, we discuss the influence of the sum of Hamiltonians on the best relative error. We obtain the following conclusion.

Proposition 3. Suppose that H_1 and H_2 are two Hamiltonians, E'_{exact} , E''_{exact} and E_{exact} are the smallest eigenvalue of H_1 , H_2 and $H_1 + H_2$, respectively, $|\Psi_{S,\Omega}\rangle$ is an NNQS. Then

$$E_{H_1+H_2}(S, \Omega) = E_{H_1}(S, \Omega) + E_{H_2}(S, \Omega).$$

Furthermore, if $\min_{\Omega}(E_{H_1}(S, \Omega) + E_{H_2}(S, \Omega)) = \min_{\Omega} E_{H_1}(S, \Omega) + \min_{\Omega} E_{H_2}(S, \Omega)$, then

$$0 \leq \epsilon \leq \epsilon_1 + \epsilon_2,$$

where

$$\epsilon_1 = \min_{\Omega} \frac{|E_{H_1}(S, \Omega) - E'_{exact}|}{|E'_{exact}|}, \epsilon_2 = \min_{\Omega} \frac{|E_{H_2}(S, \Omega) - E''_{exact}|}{|E''_{exact}|}, \epsilon = \min_{\Omega} \frac{|E_{H_1+H_2}(S, \Omega) - E_{exact}|}{|E_{exact}|}.$$

Proof. We can easily compute that

$$\begin{aligned} E_{H_1+H_2}(S, \Omega) &= \frac{\langle \Psi_{S,\Omega} | H_1 + H_2 | \Psi_{S,\Omega} \rangle}{\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle} \\ &= \frac{\langle \Psi_{S,\Omega} | H_1 | \Psi_{S,\Omega} \rangle}{\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle} + \frac{\langle \Psi_{S,\Omega} | H_2 | \Psi_{S,\Omega} \rangle}{\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle} \\ &= E_{H_1}(S, \Omega) + E_{H_2}(S, \Omega). \end{aligned}$$

It is easily see that $\epsilon \geq 0$. Generally,

$$\min_{\Omega} E_{H_1}(S, \Omega) \geq E'_{exact}, \min_{\Omega} E_{H_2}(S, \Omega) \geq E''_{exact}, \min_{\Omega} E_{H_1+H_2}(S, \Omega) \geq E_{exact}.$$

Besides, when $\min_{\Omega}(E_{H_1}(S, \Omega) + E_{H_2}(S, \Omega)) = \min_{\Omega} E_{H_1}(S, \Omega) + \min_{\Omega} E_{H_2}(S, \Omega)$, we see from $E_{exact} \geq E'_{exact} + E''_{exact}$ that

$$\begin{aligned} \epsilon &= \frac{|\min_{\Omega} E_{H_1+H_2}(S, \Omega) - E_{exact}|}{|E_{exact}|} \\ &= \frac{|\min_{\Omega}(E_{H_1}(S, \Omega) + E_{H_2}(S, \Omega)) - E_{exact}|}{|E_{exact}|} \\ &\leq \frac{|\min_{\Omega} E_{H_1}(S, \Omega) + \min_{\Omega} E_{H_2}(S, \Omega) - E'_{exact} - E''_{exact}|}{|E'_{exact} + E''_{exact}|} \\ &\leq \frac{|\min_{\Omega} E_{H_1}(S, \Omega) - E'_{exact}|}{|E'_{exact}|} + \frac{|\min_{\Omega} E_{H_2}(S, \Omega) - E''_{exact}|}{|E''_{exact}|} \\ &= \epsilon_1 + \epsilon_2. \end{aligned}$$

□

Now, we discuss the influence of tensor product of Hamiltonians on the best relative error. We get the following conclusion.

Proposition 4. Suppose that H_1 and H_2 are two Hamiltonians, E'_{exact} , E''_{exact} and E_{exact} are the smallest eigenvalue of H_1 , H_2 and $H_1 \otimes H_2$, respectively. $|\Psi'_{S',\Omega'}\rangle$ and $|\Psi''_{S'',\Omega''}\rangle$ are two NNQSSs with parameters

$$S' = S'_1 \otimes \dots \otimes S'_{N'}, S'' = S''_1 \otimes \dots \otimes S''_{N''}, \Omega' = (a', b', W'), \Omega'' = (a'', b'', W''),$$

respectively. Let

$$S_0 = S' \otimes S'', \Omega_0 = (a_0, b_0, W_0), N = N' + N'', M_0 = M' + M'',$$

$$a_0 = \begin{pmatrix} a' \\ a'' \end{pmatrix}, b_0 = \begin{pmatrix} b' \\ b'' \end{pmatrix}, W_0 = [W_{ij}] = \begin{pmatrix} W'_{M' \times N'} & 0 \\ 0 & W''_{M'' \times N''} \end{pmatrix}.$$

Then

$$E_{H_1 \otimes H_2}(S_0, \Omega_0) = E_{H_1}(S', \Omega') \cdot E_{H_2}(S'', \Omega'').$$

Furthermore, if H_1 and H_2 are positive definite, then $\epsilon' \epsilon'' \leq \epsilon^0$ where

$$\epsilon' = \min_{\Omega'} \frac{|E_{H_1}(S', \Omega') - E'_{exact}|}{|E'_{exact}|}, \epsilon'' = \min_{\Omega''} \frac{|E_{H_2}(S'', \Omega'') - E''_{exact}|}{|E''_{exact}|},$$

$$\epsilon^0 = \min_{\Omega_0} \frac{|E_{H_1 \otimes H_2}(S_0, \Omega_0) - E_{exact}|}{|E_{exact}|}.$$

Proof. Since $|\Psi'_{S', \Omega'}\rangle$ and $|\Psi''_{S'', \Omega''}\rangle$ are two NNQSs, we know from Proposition 1 that $|\Psi'_{S', \Omega'}\rangle \otimes |\Psi''_{S'', \Omega''}\rangle = |\Phi_{S_0, \Omega_0}\rangle$ is also an NNQS. Furthermore, we can compute

$$\begin{aligned} E_{H_1 \otimes H_2}(S_0, \Omega_0) &= \frac{\langle \Psi_{S_0, \Omega_0} | H_1 \otimes H_2 | \Psi_{S_0, \Omega_0} \rangle}{\langle \Psi_{S_0, \Omega_0} | \Psi_{S_0, \Omega_0} \rangle} \\ &= \frac{\langle \Psi'_{S', \Omega'} | H_1 | \Psi'_{S', \Omega'} \rangle}{\langle \Psi'_{S', \Omega'} | \Psi'_{S', \Omega'} \rangle} \cdot \frac{\langle \Psi''_{S'', \Omega''} | H_2 | \Psi''_{S'', \Omega''} \rangle}{\langle \Psi''_{S'', \Omega''} | \Psi''_{S'', \Omega''} \rangle} \\ &= E_{H_1}(S', \Omega') \cdot E_{H_2}(S'', \Omega''). \end{aligned}$$

Since H_1 and H_2 are positive, $E_{exact} = E'_{exact} E''_{exact}$. Observe that

$$\min_{\Omega'} E_{H_1}(S', \Omega') \geq E'_{exact} > 0, \min_{\Omega''} E_{H_2}(S'', \Omega'') \geq E''_{exact} > 0, \min_{\Omega_0} E_{H_1 \otimes H_2}(S_0, \Omega_0) \geq E_{exact} > 0.$$

Thus, we have

$$\begin{aligned} \epsilon^0 &= \frac{|\min_{\Omega_0} E_{H_1 \otimes H_2}(S_0, \Omega_0) - E_{exact}|}{|E_{exact}|} \\ &= \frac{|\min_{\Omega'} E_{H_1}(S', \Omega') \cdot \min_{\Omega''} E_{H_2}(S'', \Omega'') - E'_{exact} E''_{exact}|}{|E'_{exact}| \cdot |E''_{exact}|} \\ &\geq \frac{|\min_{\Omega'} E_{H_1}(S', \Omega') - E'_{exact}|}{|E'_{exact}|} \cdot \frac{|\min_{\Omega''} E_{H_2}(S'', \Omega'') - E''_{exact}|}{|E''_{exact}|} \\ &= \epsilon' \epsilon''. \end{aligned}$$

□

Now, we discuss the influence of local unitary operation on the best relative error. We conclude this conclusion as follows.

Proposition 5. Suppose that H is a Hamiltonian, $|\Psi_{S, \Omega}\rangle$ is an NNQS and $U = U_1 \otimes U_2 \otimes \dots \otimes U_N$ is a local unitary operator on \mathcal{H} . E_{exact}, E'_{exact} are the smallest eigenvalue of H and UHU^\dagger , respectively. Then

$$E_{UHU^\dagger}(S, \Omega) = E_H(U^\dagger S U, \Omega),$$

and $\epsilon = \epsilon'$ where

$$\epsilon = \min_{\Omega} \frac{|E_H(U^\dagger S U, \Omega) - E_{exact}|}{|E_{exact}|}, \epsilon' = \min_{\Omega} \frac{|E_{UHU^\dagger}(S, \Omega) - E'_{exact}|}{|E'_{exact}|}.$$

Proof. We can obtain from Proposition 2 that $U^\dagger |\Psi_{S,\Omega}\rangle = |\Psi_{U^\dagger S U, \Omega}\rangle$, which is also an NNQS. Therefore

$$\begin{aligned} E_{UHU^\dagger}(S, \Omega) &= \frac{\langle \Psi_{S,\Omega} | UHU^\dagger | \Psi_{S,\Omega} \rangle}{\langle \Psi_{S,\Omega} | \Psi_{S,\Omega} \rangle} \\ &= \frac{\langle \Psi_{U^\dagger S U, \Omega} | H | \Psi_{U^\dagger S U, \Omega} \rangle}{\langle \Psi_{U^\dagger S U, \Omega} | \Psi_{U^\dagger S U, \Omega} \rangle} \\ &= E_H(U^\dagger S U, \Omega). \end{aligned}$$

Since U is a local unitary operator, $E_{exact} = E'_{exact}$. We can easily obtain that $\epsilon = \epsilon'$. \square

Lastly, we give two examples in order to illustrate our method.

Example 1. Suppose that $H = |00\rangle\langle 00| + 2|01\rangle\langle 01| + 3|10\rangle\langle 10| + 4|11\rangle\langle 11|$. Then H can be represented under the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ by $H = \text{diag}(1, 2, 3, 4)$. It is easy to see that the minimum eigenvalue of H is 1, the ground state is $|00\rangle$.

Next we use spin-z NNQSs

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1 k_2} \in \{1, -1\}^2} \Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}) |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle.$$

to approximate the ground state $|00\rangle$ of H , where

$$\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}) = \prod_{j=1}^2 e^{a_j \lambda_{k_j}} \cdot \prod_{i=1}^M 2 \cosh \left(b_i + \sum_{j=1}^2 W_{ij} \lambda_{k_j} \right).$$

When $N = M = 2$, we have

$$\begin{aligned} |\Psi_{S,\Omega}\rangle &= 4e^{a_1} e^{a_2} \cosh(b_1 + W_{11} + W_{12}) \cosh(b_2 + W_{21} + W_{22}) |00\rangle \\ &+ 4e^{a_1} e^{-a_2} \cosh(b_1 + W_{11} - W_{12}) \cosh(b_2 + W_{21} - W_{22}) |01\rangle \\ &+ 4e^{-a_1} e^{a_2} \cosh(b_1 - W_{11} + W_{12}) \cosh(b_2 - W_{21} + W_{22}) |10\rangle \\ &+ 4e^{-a_1} e^{-a_2} \cosh(b_1 - W_{11} - W_{12}) \cosh(b_2 - W_{21} - W_{22}) |11\rangle. \end{aligned}$$

We can easily calculate that

$$\begin{aligned} E_H(S, \Omega) &= \left(|e^{a_1} e^{a_2} \cosh(b_1 + W_{11} + W_{12}) \cosh(b_2 + W_{21} + W_{22})|^2 \right. \\ &+ 2 |e^{a_1} e^{-a_2} \cosh(b_1 + W_{11} - W_{12}) \cosh(b_2 + W_{21} - W_{22})|^2 \\ &+ 3 |e^{-a_1} e^{a_2} \cosh(b_1 - W_{11} + W_{12}) \cosh(b_2 - W_{21} + W_{22})|^2 \\ &+ 4 |e^{-a_1} e^{-a_2} \cosh(b_1 - W_{11} - W_{12}) \cosh(b_2 - W_{21} - W_{22})|^2 \left. \right) \\ &/ \left(|e^{a_1} e^{a_2} \cosh(b_1 + W_{11} + W_{12}) \cosh(b_2 + W_{21} + W_{22})|^2 \right. \\ &+ |e^{a_1} e^{-a_2} \cosh(b_1 + W_{11} - W_{12}) \cosh(b_2 + W_{21} - W_{22})|^2 \\ &+ |e^{-a_1} e^{a_2} \cosh(b_1 - W_{11} + W_{12}) \cosh(b_2 - W_{21} + W_{22})|^2 \\ &+ |e^{-a_1} e^{-a_2} \cosh(b_1 - W_{11} - W_{12}) \cosh(b_2 - W_{21} - W_{22})|^2 \left. \right). \end{aligned}$$

Next we seek the minimum value of $E_H(S, \Omega)$ over Ω . By letting

$$b_1 = x_1, b_2 = x_2, W_{11} = x_3, W_{12} = x_4, W_{21} = x_5, W_{22} = x_6, a_1 = x_7, a_2 = x_8,$$

we define a function g by

$$\begin{aligned}
 &g(x_1, x_2, \dots, x_8) \\
 = & \left(|e^{x_7+x_8} \cdot \cosh(x_1 + x_3 + x_4) \cdot \cosh(x_2 + x_5 + x_6)|^2 \right. \\
 & + 2 |e^{x_7-x_8} \cdot \cosh(x_1 + x_3 - x_4) \cdot \cosh(x_2 + x_5 - x_6)|^2 \\
 & + 3 |e^{-x_7+x_8} \cdot \cosh(x_1 - x_3 + x_4) \cdot \cosh(x_2 - x_5 + x_6)|^2 \\
 & + 4 |e^{-x_7-x_8} \cdot \cosh(x_1 - x_3 - x_4) \cdot \cosh(x_2 - x_5 - x_6)|^2 \\
 & \left. / (|e^{x_7+x_8} \cdot \cosh(x_1 + x_3 + x_4) \cdot \cosh(x_2 + x_5 + x_6)|^2 \right. \\
 & + |e^{x_7-x_8} \cdot \cosh(x_1 + x_3 - x_4) \cdot \cosh(x_2 + x_5 - x_6)|^2 \\
 & + |e^{-x_7+x_8} \cdot \cosh(x_1 - x_3 + x_4) \cdot \cosh(x_2 - x_5 + x_6)|^2 \\
 & \left. + |e^{-x_7-x_8} \cdot \cosh(x_1 - x_3 - x_4) \cdot \cosh(x_2 - x_5 - x_6)|^2 \right)
 \end{aligned}$$

and then numerically minimize g over x_1, x_2, \dots, x_8 (see Figure 3).

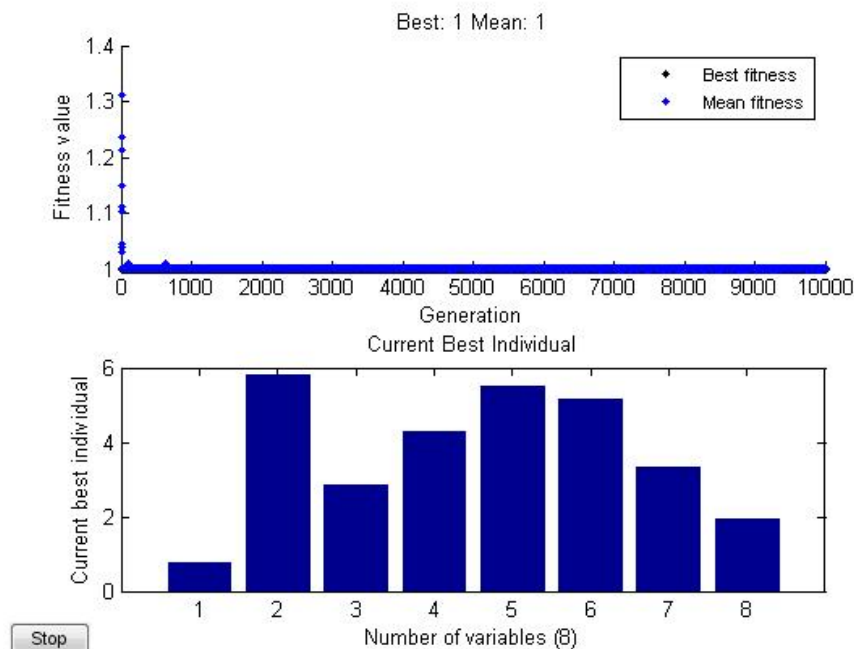


Figure 3. Numerically minimize g over x_1, x_2, \dots, x_8 by optimization.

By using Matlab, we find

$$\min_{x_i} g(x_1, x_2, \dots, x_8) = g(0.743, 5.788, 2.843, 4.274, 5.501, 5.148, 3.312, 1.916) = 1.$$

We obtain

$$\epsilon = \min_{\Omega} \frac{|E_H(S, \Omega) - E_{exact}|}{|E_{exact}|} = 0$$

Meanwhile, the corresponding NNQS is

$$|\Psi_{S, \Omega}\rangle = 6.6458 \times 10^{12} |00\rangle + 4.6761 \times 10^3 |01\rangle + 505.6622 |10\rangle + 406.2882 |11\rangle,$$

then the normalized NNQS is

$$|\Psi'_{S,\Omega}\rangle = \frac{|\Psi_{S,\Omega}\rangle}{\sqrt{\langle\Psi_{S,\Omega}|\Psi_{S,\Omega}\rangle}} \approx |00\rangle.$$

Besides, we can also calculate the distance between the actual ground state $|00\rangle$ and the approximate state $|\Psi'_{S,\Omega}\rangle$ to be

$$\text{dist}(|00\rangle, |\Psi'_{S,\Omega}\rangle) = \||00\rangle - |\Psi'_{S,\Omega}\rangle\| \approx 0.$$

Example 2. Suppose that

$$H_N^{\text{cluster}} = - \sum_{i=1}^N \sigma_{i-1}^z \sigma_i^x \sigma_{i+1}^z,$$

where $\sigma_0^z = I, \sigma_{N+1}^z = I$. It is easy to see that the minimum eigenvalue of H_N^{cluster} is $-N$, the ground state is cluster state $|C_N\rangle$. Hence, $E_{\text{exact}} = -N$.

Next we use spin-z NNQSs

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1 k_2 \dots k_N} \in \{1,-1\}^N} \Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \dots \otimes |\psi_{k_N}\rangle,$$

to approximate the ground state $|C_N\rangle$ of H_N^{cluster} , where

$$\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \dots, \lambda_{k_N}) = \prod_{j=1}^N e^{a_j \lambda_{k_j}} \cdot \prod_{i=1}^M 2 \cosh \left(b_i + \sum_{j=1}^N W_{ij} \lambda_{k_j} \right).$$

(i) When $N = M = 2$. By letting

$$a_1 = x_1 + x_2i, a_2 = x_3 + x_4i, b_1 = x_5 + x_6i, b_2 = x_7 + x_8i,$$

$$W_{11} = x_9 + x_{10}i, W_{12} = x_{11} + x_{12}i, W_{21} = x_{13} + x_{14}i, W_{22} = x_{15} + x_{16}i,$$

using Matlab(see Figure 4), we find

$$\epsilon = 1.438 \times 10^{-6},$$

where

$$a = \begin{pmatrix} 0.065 + 0.194i \\ 0.008 + 0.37i \end{pmatrix}, b = \begin{pmatrix} 0.022 + 0.693i \\ -0.431 - 0.056i \end{pmatrix}, W = \begin{pmatrix} 0.437 + 0.909i & 0.018 + 0.733i \\ -0.272 + 0.952i & 0.2 + 0.771i \end{pmatrix}.$$

Meanwhile, the corresponding NNQS is

$$|\Psi_{S,\Omega}\rangle = (3.5877 + 0.4407i)|00\rangle + (3.5755 + 0.5083i)|01\rangle + (3.5805 + 0.4372i)|10\rangle + (-3.5698 - 0.5169i)|11\rangle,$$

then normalized NNQS is

$$|\Psi'_{S,\Omega}\rangle = \frac{|\Psi_{S,\Omega}\rangle}{\sqrt{\langle\Psi_{S,\Omega}|\Psi_{S,\Omega}\rangle}} = (0.4969 + 0.0610i)|00\rangle + (0.4952 + 0.0704i)|01\rangle + (0.4959 + 0.0606i)|10\rangle + (-0.4944 - 0.0716i)|11\rangle.$$

Besides, we can also calculate the fidelity between the actual ground state

$$|C_2\rangle = \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle - \frac{1}{2}|11\rangle$$

and the approximate state $|\Psi'_{S,\Omega}\rangle$ to be

$$F(|C_2\rangle, |\Psi'_{S,\Omega}\rangle) = |\langle C_2 | \Psi'_{S,\Omega} \rangle| = 0.9999 \approx 1.$$

Hence, $|C_2\rangle \approx |\Psi'_{S,\Omega}\rangle$.

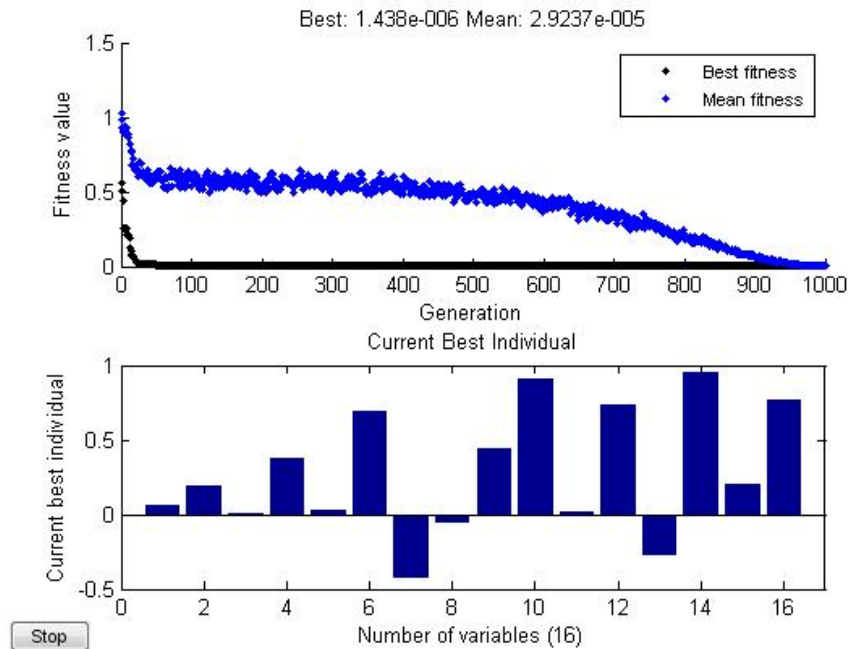


Figure 4. Numerically minimize ϵ by optimization.

In addition, we find that when $N = 2$, ϵ gets smaller and smaller as M changes, see Table 1.

Table 1. The numerical simulation results of N, M .

N	M	ϵ
2	2	1.438×10^{-6}
2	4	1.0716×10^{-6}
2	6	6.7887×10^{-7}
2	8	4.987×10^{-7}

(ii) When $N = 3, M = 3$. By using Matlab (see Figure 5), we find

$$\epsilon = 2.981 \times 10^{-4}.$$

where

$$a = \begin{pmatrix} 0.956 + 1.669i \\ 1.309 - 0.255i \\ -0.148 - 0.152i \end{pmatrix}, b = \begin{pmatrix} 0.653 + 0.863i \\ 0.569 + 0.706i \\ -0.613 + 0.894i \end{pmatrix},$$

$$W = \begin{pmatrix} -0.066 + 0.969i & -1.213 + 2.029i & -0.354 - 0.647i \\ -0.233 + 3.12i & 0.986 + 0.198i & 0.438 + 0.16i \\ 0.74 + 1.206i & 0.749 - 0.985i & -0.445 + 0.8i \end{pmatrix}.$$

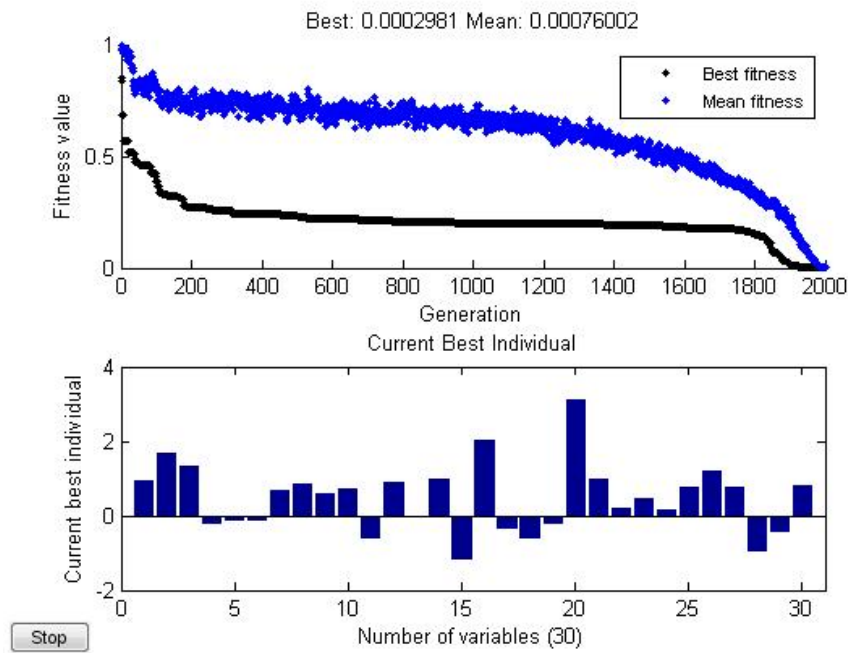


Figure 5. Numerically minimize ϵ by optimization.

Meanwhile, the corresponding NNQS is

$$|\Psi_{S,\Omega}\rangle = (-4.5329 - 9.8797i)|000\rangle + (-4.4661 - 9.6734i)|001\rangle + (-4.5709 - 9.9717i)|010\rangle \\ + (4.3557 + 9.7498i)|011\rangle + (-4.4258 - 9.8957i)|100\rangle + (-4.4603 - 9.6152i)|101\rangle \\ + (4.6706 + 9.8781i)|110\rangle + (-4.1489 - 9.7979i)|111\rangle,$$

then normalized NNQS is

$$|\Psi'_{S,\Omega}\rangle = \frac{|\Psi_{S,\Omega}\rangle}{\sqrt{\langle\Psi_{S,\Omega}|\Psi_{S,\Omega}\rangle}} = (-0.1488 - 0.3242i)|000\rangle + (-0.1466 - 0.3175i)|001\rangle \\ + (-0.1500 - 0.3272i)|010\rangle + (0.1429 + 0.32i)|011\rangle + (-0.1452 - 0.3248i)|100\rangle \\ + (-0.1464 - 0.3156i)|101\rangle + (0.1533 + 0.3242i)|110\rangle + (-0.1362 - 0.3215i)|111\rangle.$$

Besides, we can also calculate the fidelity between the actual ground state

$$|C_3\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |001\rangle + |010\rangle - |011\rangle + |100\rangle + |101\rangle - |110\rangle + |111\rangle)$$

and the approximate state $|\Psi'_{S,\Omega}\rangle$ to be

$$F(|C_3\rangle, |\Psi'_{S,\Omega}\rangle) = |\langle C_3|\Psi'_{S,\Omega}\rangle| = 0.9999 \approx 1.$$

Hence, $|C_3\rangle \approx |\Psi'_{S,\Omega}\rangle$.

4. Conclusions

In this paper, the question of approximating ground states by neural network quantum states has been discussed in terms of the best relative error (BRE). Some properties of the BREs have been obtained, including the BREs of sums, tensor products, and local unitary transformations of Hamiltonians. Besides, our method has been illustrated with two examples.

Author Contributions: The work of this paper was accomplished by Y.Y., C.Z. and H.C. Moreover, all authors have read the paper carefully and approved the research contents that were written in the final manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 11871318, 11771009, 11571213, 11601300), the Fundamental Research Funds for the Central Universities (GK201703093, GK201801011), Shaanxi Province Innovation Ability Support Program (2018KJXX-054), and Subject Research Project of Yuncheng University (XK-2018032).

Acknowledgments: We thank Wenfei Cao of Shaanxi Normal University for his help in writing algorithms.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Luo, S.L. Using measurement-induced disturbance to characterize correlations as classical or quantum. *Phys. Rev. A* **2008**, *77*, 022301. [[CrossRef](#)]
2. Guo, Z.H.; Cao, H.X.; Qu, S.X. Partial correlations in multipartite quantum systems. *Inf. Sci.* **2014**, *289*, 262. [[CrossRef](#)]
3. Zhang, T.G.; Yang, H.; Lijost, X. Uniform Quantification of Correlations for Bipartite Systems. *Phys. Rev. A* **2017**, *95*, 042316. [[CrossRef](#)]
4. Vedral, V.; Plenio, M.B.; Rippin, M.A.; Knight, P.L. Quantifying entanglement. *Phys. Rev. Lett.* **1997**, *78*, 2275. [[CrossRef](#)]
5. Vidal, G.; Werner, R.F. A computable measure of entanglement. *Phys. Rev. A* **2002**, *65*, 032314. [[CrossRef](#)]
6. Amico, L.; Osterloh, A.; Vedral, V. Entanglement in many-body systems. *Rev. Mod. Phys.* **2007**, *80*, 517. [[CrossRef](#)]
7. Luo, S.L.; Fu, S.S. Measurement-induced nonlocality. *Phys. Rev. Lett.* **2011**, *106*, 120401. [[CrossRef](#)] [[PubMed](#)]
8. Guo, Y.; Hou, J.C. Local channels preserving the states without measurement-induced nonlocality. *J. Phys. A Math. Theor.* **2013**, *46*, 325301. [[CrossRef](#)]
9. Cao, H.X.; Guo, Z.H. Characterizing Bell nonlocality and EPR steering. *Sci. China-Phys. Mech. Astron.* **2019**, *62*, 030311. [[CrossRef](#)]
10. Skrzypczyk, P.; Navascues, M.; Cavalcanti, D. Quantifying Einstein-Podolsky-Rosen steering. *Phys. Rev. Lett.* **2014**, *112*, 180404. [[CrossRef](#)]
11. Sun, K.; Ye, X.J.; Xu, J.S.; Xu, X.Y.; Tang, J.S.; Wu, Y.C.; Chen, J.L.; Li, C.F.; Guo, G.C. Experimental quantification of asymmetric Einstein-Podolsky-Rosen steering. *Phys. Rev. Lett.* **2016**, *116*, 160404. [[CrossRef](#)]
12. Cavalcanti, D.; Skrzypczyk, P. Quantum steering: A review with focus on semidefinite programming. *Rep. Prog. Phys.* **2017**, *80*, 024001. [[CrossRef](#)] [[PubMed](#)]
13. Gordon, R.G. Error bounds for the energy and overlap of approximate wavefunctions. *J. Chem. Phys.* **1968**, *48*, 4984. [[CrossRef](#)]
14. Wang, P.S.C. Lower bounds for the overlap of approximate wave functions. *J. Chem. Phys.* **1970**, *52*, 4464. [[CrossRef](#)]
15. Merkel, R.R. A study of a lower bounding formula for the overlap between the exact and an approximate wavefunction. *J. Chem. Phys.* **1975**, *62*, 3198. [[CrossRef](#)]
16. Cioslowski, J. Estimation of the overlap between the approximate and exact wave function of the ground state from the connected-moments expansion. *Phys. Rev. A* **1987**, *36*, 3441. [[CrossRef](#)]
17. Hornik, M.; Snajdr, M.; Rothstein, S.M. Estimating the overlap of an approximate with the exact wave function by quantum Monte Carlo methods. *J. Chem. Phys.* **2000**, *113*, 3496. [[CrossRef](#)]
18. Marmorino, M.G. Upper bounds to the overlap of approximate and exact wavefunctions. *Int. J. Theor. Phys.* **2000**, *39*, 2439. [[CrossRef](#)]
19. Nomura, Y.; Darmawan, A.; Yamaji, Y.; Imada, M. Restricted-Boltzmann-machine learning for solving strongly correlated quantum systems. *Phys. Rev. B* **2006**, *96*, 205152. [[CrossRef](#)]
20. Kolmogorov, A.N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Am. Math. Soc. Transl.* **1963**, *28*, 55.
21. Cybenko, G. Approximation by superposition of a sigmoidal function. *Math. Control Signal* **1989**, *2*, 303. [[CrossRef](#)]
22. Funahashi, K. On the approximate realization of continuous mappings by neural networks. *Neural Netw.* **1989**, *2*, 183. [[CrossRef](#)]

23. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359. [[CrossRef](#)]
24. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251. [[CrossRef](#)]
25. Le Roux, N.; Bengio, Y. Representational power of restricted boltzmann machines and deep belief networks. *Neural Comput.* **2008**, *20*, 1631. [[CrossRef](#)] [[PubMed](#)]
26. Carleo, G.; Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **2017**, *355*, 602. [[CrossRef](#)] [[PubMed](#)]
27. Ackley, D.H.; Hinton, G.E.; Sejnowski, T.J. A learning algorithm for Boltzmann machines. *Cogn. Sci.* **1985**, *9*, 147. [[CrossRef](#)]
28. Cai, Z. Approximating quantum many-body wave functions using artificial neural networks. *arXiv* **2017**, arXiv:1704.05148.
29. Saito, H. Solving the bose-hubbard model with machine learning. *J. Phys. Soc. Jpn.* **2017**, *86*, 093001. [[CrossRef](#)]
30. Gao, X.; Duan, L.M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **2017**, *8*, 662. [[CrossRef](#)]
31. Huang, Y.; Moore, J.E. Neural network representation of tensor network and chiral states. *arXiv* **2017**, arXiv:1701.06246.
32. Deng, D.L.; Li, X.P.; Das Sarma, S. Machine learning topological states. *Phys. Rev. B* **2017**, *96*, 195145. [[CrossRef](#)]
33. Glasser, I.; Pancotti, N.; August, M.; Rodriguez, I.D.; Cirac, J.I. Neural-network quantum states, string-bond states, and Chiral topological states. *Phys. Rev. X* **2018**, *8*, 011006. [[CrossRef](#)]
34. Gardas, B.; Rams, M.M.; Dziarmaga, J. Quantum neural networks to simulate many-body quantum systems. *Phys. Rev. B* **2018**, *98*, 184304. [[CrossRef](#)]
35. Cai, Z.; Liu, J.G. Approximating quantum many-body wave-functions using artificial neural network. *Phys. Rev. B* **2018**, *97*, 035116. [[CrossRef](#)]
36. Saito, H.; Kato, M. Machine learning technique to find quantum many-body ground states of bosons on a lattice. *J. Phys. Soc. Jpn.* **2018**, *87*, 014001. [[CrossRef](#)]
37. Clark, S.R. Unifying neural-network quantum states and correlator product states via tensor networks. *arXiv* **2017**, arXiv:1710.03545.
38. Yang, Y.; Cao, H.X.; Zhang, Z.J. Neural network representations of ground states. *Sci. China-Phys. Mech. Astron.* **2019**, submitted.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).