Article

Approximating ground states by neural network quantum states

Ying Yang [1,2], Chengyang Zhang [1], Huaixin Cao [1,*]

1 School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710119, China
2 School of Mathematics and Information Technology, Yuncheng University, Yuncheng 044000, China
* Correspondence: caohx@snnu.edu.cn; Tel.: 18691958211

Abstract: The many-body problem in quantum physics originates from the difficulty of describing the non-trivial correlations encoded in the exponential complexity of the many-body wave function. Motivated by the Giuseppe Carleo's work titled solving the quantum many-body problem with artificial neural networks [Science, 2017, 355: 602], we focus on finding the NNQS 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 unitary of Hamiltonians on the best relative error. Besides, we illustrate our method with some examples.

Keywords: Approximation; ground state; neural network quantum state

---

An artificial neural network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. Up to now, there is a lot of research on the approximation ability of neural network architectures, such as Kolmogorov [1], Hornik [2], Cybenko [3], Funahashi [4], Hornik [5], Roux [6].

The 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, or entanglement. 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. Recently, an idea that received a lot of attention from the scientific community 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 (VMC) method while a few different neural-network architectures were tested [7–10], and the most promising results so far have been achieved with Boltzmann machines [10]. 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 [11,12]. Carleo and Troyer in [7] 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. Deng et al [13,14] show that this representation can be used to describe topological states. Besides, they have constructed exact representations for SPT states and intrinsic topologically ordered states. Very recently, Glasser et al [15] 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. In addition, there are a lot of related studies, such as [16–19].

Despite such exciting development, but it is unknown whether a general state can be expressed by neural networks efficiently. Generalizing the idea of [7], we introduced in [20] first neural networks

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

quantum states (NNQSs) baced on general input observables and explored some related properties about NNQSs, such as tensor product, local unitary operation and so on. Secondly, based on the construction of neural network representations for the cluster state in $1D$, we proved necessary and sufficient conditions for a general graph state to be represented by an NNQS. We illustrated our method with some examples and observed that some N-qubit states can be represented by a normalized NNQS, such as any separable pure state, every Bell state, GHZ states and so on.

In this paper, based on the NNQSs introduced in [20], we focus on finding the NNQS approximation of the unknown ground state of a given Hamiltonian H. The remainder part of this paper are organized as follows. In Section 2, we recall the concept and the related properties of NNQSs introduced in [20]. 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 [20]. Let $Q_1, Q_2, \ldots, Q_N$ be n quantum systems with state spaces $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_N$ of dimensions $d_1, d_2, \ldots, d_N$, respectively. We consider the composite system Q of $Q_1, Q_2, \ldots, Q_n$ with state space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_N$.

Let $S_1, S_2, \ldots, S_N$ are non-degenerate observables of systems $Q_1, Q_2, \ldots, Q_N$, respectively. Then $S = S_1 \otimes S_2 \otimes \ldots \otimes S_N$ 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,

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

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

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

respectively. Put

$$V(S) = \left\{\Lambda_{k_1k_2\ldots k_N} \equiv (\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N})^T : k_j = 0, 1, \ldots, d_j - 1\right\},$$

called an input space. For parameters

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

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

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

Then we obtain a complex-valued function $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N})$ of the input variable $\Lambda_{k_1k_2\ldots k_N}$. We call it a *neural network quantum wave function* (NNQWF). It may be identically zero. In what follows, we assume that this is not the case, that is, assume that $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N}) \neq 0$ for some input variable $\Lambda_{k_1k_2\ldots k_N}$. Then we define

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1k_2\ldots k_N}\in V(S)} \Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N})|\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \ldots \otimes |\psi_{k_N}\rangle, \tag{4}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

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 $S = S_1 \otimes S_2 \otimes \ldots \otimes S_N$ (Figure 1).

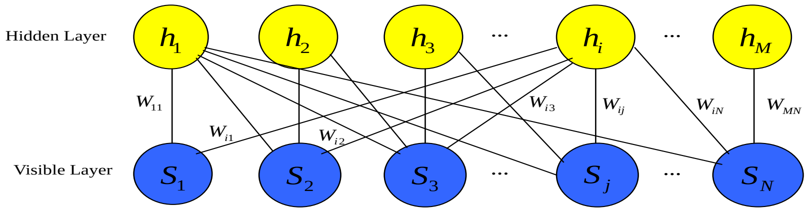

Figure 1. Artificial neural network encoding an NNQS. It is a restricted Boltzmann machine architecture that features a set of N visible artificial neurons (blue disks) and a set of M hidden neurons (yellow disks). For each value $\Lambda_{k_1k_2\ldots k_N}$ of the input observable S, the neural network computes the value of the $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N})$.

NNQWF can be reduced to

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

It is can be described by the following “quantum artificial neural network", see Figure 2 where $a = 0$ and

$$\textstyle\sum_{b_i}(x_1, x_2, \ldots, x_N) = b_i + \sum_{j=1}^{N} x_j, 2\cosh(z) = e^z + e^{-z}, \Pi(y_1, y_2, \ldots, y_M) = \Pi_{i=1}^{M} y_i,$$

and the final outcome $\Psi_{S,\Omega}(\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_N})$ is given by (5).

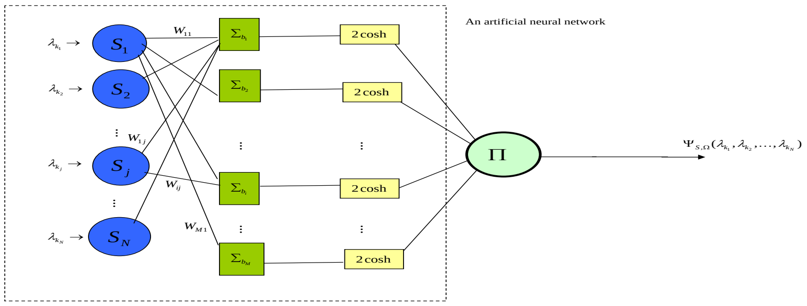

Figure 2. Quantum artificial neural network with parameter $\Omega = (0, b, W)$.

We call this network a quantum artificial neural network because that its inputs 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. *[20] Suppose that $|\Psi'_{S',\Omega'}\rangle$ and $|\Psi''_{S'',\Omega''}\rangle$ are two NNQSs with parameters*

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

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

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at *Entropy* **2019**, *21*, 82; doi:10.3390/e21010082

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

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

Proposition 2. *[20] Suppose that $|\Psi_{S,\Omega}\rangle$ is an NNQS and $U = U_1 \otimes U_2 \otimes \ldots \otimes U_N$ 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.*

As the end of this section, we discuss a special classes of NNQSs.

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

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

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

In this case, the NNQS (4) becomes

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1k_2\ldots k_N}\in\{1,-1\}^N} \Psi_{S,\Omega}(\lambda_{k_1},\lambda_{k_2},\ldots,\lambda_{k_N})|\psi_{k_1}\rangle\otimes|\psi_{k_2}\rangle\otimes\ldots\otimes|\psi_{k_N}\rangle. \tag{6}$$

This leads to the NNQS induced in [7] and discussed in [13]. 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 Eq.(4) and 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,

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

We call ϵ the *best relative error* between $E_H(S,\Omega)$ and E_{exact}. The neural network quantum state $|\Psi_{S,\Omega}\rangle$ corresponding to the minimum of ϵ is the best neural network representation of the ground state of H.

Generally, $E_H(S,\Omega)\ge E_{exact}$. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

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

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

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 $\epsilon\ge0$. Generally,

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

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 $E_{exact}\ge E'_{exact}+E''_{exact}$ that

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

□

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

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 NNQSs with parameters*

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

respectively. Let

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

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

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 $\epsilon'\epsilon'' \le \epsilon^0$ 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

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

Thus, we have

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

□

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

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

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

and $\epsilon = \epsilon'$ where

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at *Entropy* **2019**, *21*, 82; doi:10.3390/e21010082

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

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

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

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* $H = diag(1,2,3,4)$. *It is easily to see that the minimum eigenvalue of* H *is* 1, *the ground state is* $|00\rangle$.

Next we use spin-z NNQSs

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1k_2}\in\{1,-1\}^2} \Psi_{S,\Omega}(\lambda_{k_1},\lambda_{k_2})|\psi_{k_1}\rangle\otimes|\psi_{k_2}\rangle.$$

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

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

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

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

we define a function g by

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

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

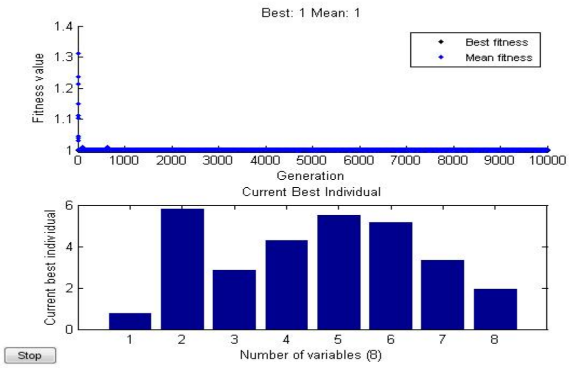

Figure 3. Numerically minimize g over $x_1, x_2, \ldots, x_8$ by optimization.

By using Matlab, we find

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

We obtain

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

Meanwhile, the corresponding NNQS is

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

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

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

Example 2. *Suppose that*

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

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

Next we use spin-z NNQSs

$$|\Psi_{S,\Omega}\rangle = \sum_{\Lambda_{k_1k_2\ldots k_N}\in\{1,-1\}^N} \Psi_{S,\Omega}(\lambda_{k_1},\lambda_{k_2},\ldots,\lambda_{k_N})|\psi_{k_1}\rangle\otimes|\psi_{k_2}\rangle\otimes\ldots\otimes|\psi_{k_N}\rangle,$$

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

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

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

$$a_1 = x_1 + x_2\mathrm{i}, a_2 = x_3 + x_4\mathrm{i}, b_1 = x_5 + x_6\mathrm{i}, b_2 = x_7 + x_8\mathrm{i},$$

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

using Matlab(see Figure 4), we find

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

where

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

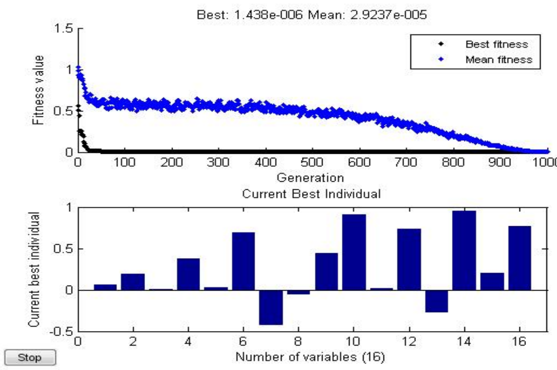

Figure 4. Numerically minimize ϵ by optimization.

Meanwhile, the corresponding NNQS is

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

then normalized NNQS is

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

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

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2018 doi:10.20944/preprints201812.0218.v1

Peer-reviewed version available at Entropy 2019, 21, 82; doi:10.3390/e21010082

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

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

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

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

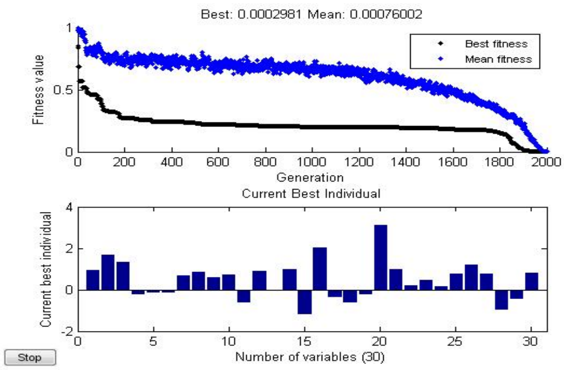

Figure 5. Numerically minimize ϵ by optimization.

Meanwhile, the corresponding NNQS is

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

then normalized NNQS is

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

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, local unitary transformations of Hamiltonians. Besides, our method have been illustrated with two examples.

Acknowledgments: 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) and Shaanxi Province Innovation Ability Support Program (2018KJXX-054).

Author Contributions: The work of this paper was accomplished by Ying Yang, Chengyang Zhang and Huaixin Cao. Moreover, all authors have read the paper carefully and approved the research contents that were written in the final manuscript. We thank Dr. Wenfei Cao of Shaanxi Normal University for his help in writing algorithms.

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

References

1. Kolmogorov, A.N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Amer. Math. Soc. Transl* **1963**, *28*, 55.
2. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Networks* **1989**, *2*, 359.
3. Cybenko, G. Approximation by superposition of a sigmoidal function. Math. *Control signal* **1989**, *2*, 303.
4. Funahashi, K. On the approximate realization of continuous mappings by neural networks, *Neural Networks* **1989**, *2*, 183.
5. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* **1991**, *4*, 251.
6. Le Roux, N.; Bengio, Y. Representational power of restricted boltzmann machines and deep belief networks. *Neural Comput.* **2008**, *20*, 1631.
7. Carleo, G.; Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **2017**, *355*, 602.
8. Ackley, D.H.; Hinton, G.E.; Sejnowski, T.J. A learning algorithm for Boltzmann machines. *Cogn. Sci.* **1985**, *9*, 147.
9. Cai, Z. Approximating quantum many-body wave functions using artificial neural networks. arXiv:1704.05148.
10. Saito, H. Solving the bose-hubbard model with machine learning. *J. Phys. Soc. Jpn.* **2017**, *86*, 093001.
11. Gao, X.; Duan, L.M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* *8*, **2017**, 662.
12. Huang, Y.; Moore, J.E. Neural network representation of tensor network and chiral states. arXiv:1701.06246.
13. Deng, D.L.; Li, X.P.; Das Sarma, S. Machine learning topological states. *Phys. Rev. B* **2017**, *96*, 195145.
14. Deng, D.L.; Li, X.; Sarma, S.D. Exact machine learning topological states. arXiv:1609.09060.
15. 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.
16. Levin, M.; Wen, X.G. Detecting topological order in a ground state wave function. *Phys. Rev. Lett.* **2006**, *96*, 110405.
17. Nomura, Y.; Darmawan, A.; Yamaji, Y.; Imada, M. Restricted-Boltzmann-machine learning for solving strongly correlated quantum systems. *Phys. Rev. B* **2006**, *96*, 205152.
18. Clark, S.R. Unifying neural-network quantum states and correlator product states via tensor networks. arXiv:1710.03545.
19. Kaubruegger, R.; Pastori, L.; Budich, J.C. Chiral topological phases from artificial neural networks. arXiv:1710.04713
20. Yang, Y.; Cao, H.X.; Zhang, Z.J. Neural network representations of ground states. *Sci. China-Phys. Mech. Astron.* Submitted.