

Solving frustrated quantum many-particle models with convolutional neural networks

Xiao Liang,^{1,2} Wen-Yuan Liu,^{1,2,3} Pei-Ze Lin,^{1,2} Guang-Can Guo,^{1,2} Yong-Sheng Zhang,^{1,2,*} and Lixin He^{1,2,†}

¹*Laboratory of Quantum Information, University of Science and Technology of China, Hefei, 230026, China*

²*Synergetic Innovation Center of Quantum Information and Quantum Physics,
University of Science and Technology of China, Hefei, 230026, China*

³*Department of Physics, The Chinese University of Hong-Kong, Shatin, New Territories, Hong Kong*

(Dated: April 19, 2019)

Recently, there has been significant progress in solving quantum many-particle problem via machine learning based on restricted Boltzmann machine (RBM). However, it is still highly challenging to solve frustrated models via machine learning, which has not been demonstrated so far. In this work, we design a brand new convolutional neural network (CNN) to solve such quantum many-particle problems. We demonstrate, for the first time, of solving the highly frustrated spin-1/2 J_1 - J_2 antiferromagnetic Heisenberg model on square lattices via CNN. The energy per site achieved by the CNN is even better than previous string-bond-state calculations. Our work therefore open up a new routine to solve challenging frustrated quantum many-particle problems using machine learning.

PACS numbers:

The successes of machine learning in image recognition,¹ playing GO,^{2,3} etc., stimulate interests in using machine learning techniques to solve physics problems, such as designing optical experiments,⁴ processing signals in finding gravity waves,^{5,6} quantum teleportation photon spot recognition,⁷ quantum phase distinguishing,^{8–10} inferring Hamiltonian solutions,^{11,12} classifying quantum states based on topological invariants,¹³ classifying spin configurations into phases,¹⁴ solving ground states of Bose-Hubbard model,¹⁵ and so on.

Recently machine learning has been applied to study quantum many-particle problems.^{16,19} This posed serious challenges, because it requires much higher precision than traditional machine learning problems. It has been demonstrated that by un-supervised learning, the restricted Boltzmann machine (RBM) can solve the ground states of transverse-field Ising model and anti-ferromagnetic Heisenberg model in rather high precision.¹⁶ The ground states of those models obey the Marshall-Peierls sign rule (MPSR)¹⁷ where the wave functions can be represented by positive numbers, suitable for RBM with real parameters.

The representation ability of RBM for some advanced states has been investigated, and often connects to another state-of-the-art method for many-particle problems, namely the tensor network state (TNS) method.¹⁸ The equivalence between RBM and TNS has been investigated in Ref.^{19–21} It has been suggested that RBM can represent the quantum states beyond area law.¹⁹ Furthermore, the extension of RBM, namely deep Boltzmann machine (DBM), can be transformed into tensor networks, and DBM can be a general representation of quantum many-particle states.²²

In the community of machine learning, it is believed that convolutional neural networks (CNN) have shown more efficiency comparing to the neural networks within dense connections.²³ Recent works try to bridge the state-of-the-art convolutional neural network to the ten-

sor network states.²⁴ It has been proposed that the information reuse introduced by overlapping convolution filters can enhance the state representation ability. Convolutional neural network has shown effectiveness in finding the critical points in phase transitions of Ising models,²⁵ which infers the network's valid state representation ability.

However, so far solving frustrated quantum many-particle models on large two-dimensional lattices is still challenging for neural networks. The ability to solve frustrated model is a milestone for neural networks, because the non-frustrated models can be efficiently solved to extreme high accuracy by quantum Monte Carlo (QMC) methods. In this work, we built a CNN to solve the ground states of two-dimensional frustrated quantum spin models, more specifically the antiferromagnetic spin-1/2 J_1 - J_2 Heisenberg model on square lattices. This model is highly frustrated, and considered as one of the most interesting and challenging spin models.^{26,27} The network is built using the elements from state-of-the-art neural networks: a convolution layer, a max-pooling layer and a transposed convolution layer. We name such CNN as the convolutional quantum neural state (CQNS). We solve the ground states at $J_2 = 0$ and $J_2/J_1 = 0.5$ and it is shown that CQNS can achieve the ground energies lower than those obtained by the previous string bond state (SBS) calculations.^{27,28} We further compare the spin correlation functions, focusing on the convolution filters's size effects. We show that if the size of the filter is too small, CQNS can not capture the correct long range spin correlations, which leads to poor energy precisions. However, for systems with the correlation length shorter than the system size, it is efficient to use relative smaller convolution filters. The CQNS is a general method which can be easily applied to study other quantum many-particle problems.

To illustrate the structure of CQNS, we begin with an example of a one dimensional spin chain. The spin chain has finite length $L = 4$ with periodic boundary

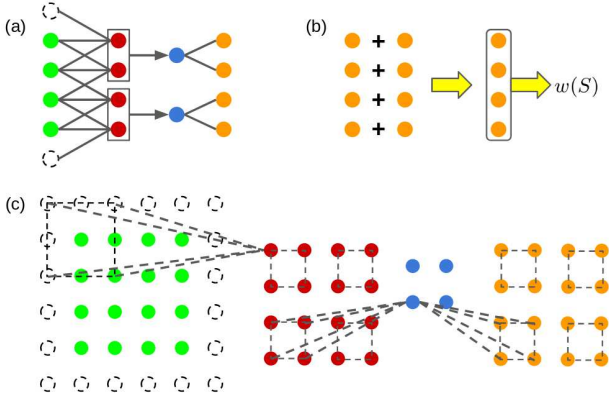


FIG. 1: (Color online) (a) The structure of a CQNS for a periodic spin chain with length $L = 4$, and the convolution filter length $K = 3$. Because of padding the convolution filter scans on the chain length of six. The green neurons are the input states whereas the red neurons are the outputs after convolution. Max-pooling is performed by two neurons with the stride number of two, where the outputs are the blue neurons. The orange neurons are the outputs after transposed convolution. (b) The final chain of neurons is the element-wise summation of the output chains from all convolution filters. (c) An example of a two dimensional CQNS on a 4×4 square lattice with periodic boundary condition and the size of the convolution filter is $K = 3$. The max-pooling and transposed convolution are also performed on the two-dimensional lattice.

conditions (PBC). The size of the convolution filter is taken to be $K = 3$. The network structure is shown in Fig. 1(a). The input spins are labeled as the green neurons, which are presented in the Pauli matrix σ_z base. To keep the output dimension unchanged after convolution, we pad additional spin sites on the head and the tail of the spin chain according to PBC. The number of sites padded on each side of the spin chain is $(K - 1)/2$. For $K = 3$, one additional spin site is padded on both the head and the tail of the spin chain. The reason to keep the output dimension unchanged after convolution is to use the advantage of max-pooling. In image recognition tasks, max-pooling reduces the output dimensions and introduces position invariance,²⁹ thus it is possible to use small number of convolution filters to grasp all possible spin configurations in the receptive field, especially when the filter size is large. The output neurons after max-pooling are denoted by the blue neurons. After max-pooling, the dimension is restored to the dimension of the spin lattice by a transposed convolution.

A quantum state of a spin lattice can be represented in the form of $|\Psi\rangle = \sum_{\mathbf{S}} w(\mathbf{S})|\mathbf{S}\rangle$, where $|\mathbf{S}\rangle = |s_1, s_2, \dots, s_N\rangle$ is a spin configuration, and $w(\mathbf{S})$ is its coefficient. For a given spin configuration $|\mathbf{S}\rangle$, CQNS outputs the wave function coefficient $w(\mathbf{S})$, which is the product of all the final output neurons. Figure 1(a) depicts the network structure of using one convolution filter, which generates one chain of output neurons. When N convolution filters are used, there will be N chains of

output neurons. To correlate different filters, we sum all the output neurons on the same site generated from all filters, as shown in Fig. 1(b) denotes. The wave-function coefficient $w(\mathbf{S})$ is obtained by taking the products of all the neurons in the final chain.

One can easily extend the above one-dimensional neural network to higher dimensions. Figure 1(c) depicts a two-dimensional CQNS with a $K \times K$ convolution filter, with $K = 3$. After convolution layer, a max-pooling is performed on adjacent four neurons followed by a transposed convolution restores the dimensions. The wave-function coefficient $w(\mathbf{S})$ is the product of all the neurons on the final output plane.

To reveal how the CQNS works, we use the $L=4$ spin chain as an example. Considering a convolution filter with length $K = 3$ and one additional spin padding in both the head and the tail of the spin chain, the convolution operation can be written in the matrix form:

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = b + \begin{bmatrix} w_1 & w_2 & w_3 & 0 & 0 & 0 \\ 0 & w_1 & w_2 & w_3 & 0 & 0 \\ 0 & 0 & w_1 & w_2 & w_3 & 0 \\ 0 & 0 & 0 & w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} s_4 \\ s_1 \\ \vdots \\ s_1 \end{bmatrix}, \quad (1)$$

where $\mathbf{S} = [s_4, s_1, s_2, s_3, s_4, s_1]^T$ is the input spin configuration and b is the bias of the convolution filter. The output neurons after convolution are the input neurons of the max-pooling, i.e.,

$$\begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} \max(p_1, p_2) \\ \max(p_3, p_4) \end{bmatrix}. \quad (2)$$

We then perform transposed convolution, the operation of transposed convolution,

$$\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} d_1 & 0 \\ d_2 & 0 \\ 0 & d_1 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}. \quad (3)$$

to restore the original size of the lattice.

Equation 1- 3 describe the transformations between the input spins and the output neurons when using one convolution filter. Using M convolution filters lead to M output chains, as shown in Fig. 1(b). The final output chain is the direct summation of the output chains from all convolution filters, i.e.,

$$\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \sum_{m=1}^M \begin{bmatrix} d_1^{(m)} g_1^{(m)} \\ d_2^{(m)} g_1^{(m)} \\ d_1^{(m)} g_2^{(m)} \\ d_2^{(m)} g_2^{(m)} \end{bmatrix}. \quad (4)$$

The final wave function coefficient is the product of all neurons in the final output chain, i.e., $w(\mathbf{S}) = h_1 h_2 h_3 h_4$. For the system with N sites, the final expression for $w(\mathbf{S})$ is,

$$w(\mathbf{S}) = \sum_{n_1, n_2, \dots, n_N} g(n_1, n_2, \dots, n_N) s_1^{n_1} s_2^{n_2} \dots s_N^{n_N}, \quad (5)$$

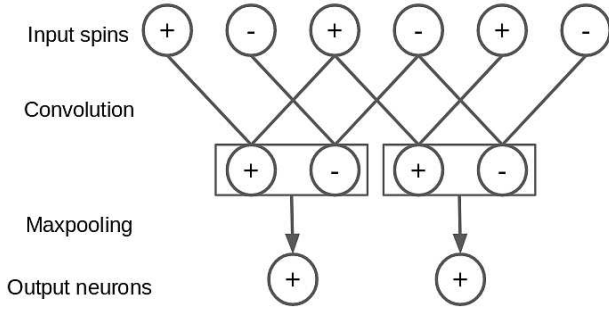


FIG. 2: Diagram of CQNS with only one convolution filter. The kernel of the filter are $[0.5, 0, 0.5]$. When the input spin configuration is a Néel state, the output neurons after convolution are also positive and negative distributed alternatively. After the max-pooling, all the output neurons are positive, which leads to a non-zero wave-function coefficient. For all other input spin configurations, the coefficients are zero.

where $n_1 + n_2 + \dots + n_N \leq N$, and $g(n_1, n_2, \dots, n_N)$ is given by CQNS.

It is easy to show that CQNS can exactly represent the classical Néel state, which has alternating spin up and spin down in the lattice, and is the ground state of a classical antiferromagnetic Heisenberg model, using single convolution filter. The filter kernel has the form $[0.5, 0, 0.5]$, and biases are taken to be zero. Figure 2 depicts the intermediate neurons after convolution and max-pooling when the input state is a Néel state. The output neurons after max-pooling are all positive, thus the wave-function coefficient $w(\mathbf{S}) > 0$. It is easy to verify that at least one of the output neurons after max-pooling is zero if the input state is not the Néel state, and therefore $w(\mathbf{S}) = 0$. For a quantum antiferromagnetic Heisenberg model more filters are needed to represent the ground state. As an example, for the spin chain of $L = 4$, it is easy to verify that the ground state can be exactly represented by the CQNS with four filters, which are $[-0.5, 0.5, 0.5]$, $[0.5, 0.5, -0.5]$, $[0.5, -0.5, 0.5]$ and $[0.5, 0, 0.5]$ and biases for the filters are -1.5, -1.5, -0.5 and 0 respectively.

We benchmark the CQNS on the two dimensional spin-1/2 J_1 - J_2 Heisenberg model,

$$H = \sum_{\langle i, j \rangle} J_1 \mathbf{s}_i \cdot \mathbf{s}_j + J_2 \sum_{\langle\langle i, j \rangle\rangle} \mathbf{s}_i \cdot \mathbf{s}_j, \quad (6)$$

where $\langle i, j \rangle$ denotes the nearest neighbour sites and $\langle\langle i, j \rangle\rangle$ denotes the next nearest neighbour sites. We assume both $J_1, J_2 > 0$, and take $J_1 = 1$ in all calculations in this work. J_1 - J_2 model is a one of the most interesting and challenging quantum spin models due to the strong frustrated interactions, and the ground state near $J_2/J_1 = 0.5$ is still under intensive debate.³⁰⁻³³ In this work, we only focus on the ground state energies of the model.

The total energy of the system can be calculated as:

$$E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{1}{\sum_{\mathbf{S}} w_{\mathbf{S}}^2} \sum_{\mathbf{S}} w_{\mathbf{S}}^2 \sum_{\mathbf{S}'} \frac{w_{\mathbf{S}'}}{w_{\mathbf{S}}} H_{\mathbf{S}'\mathbf{S}}, \quad (7)$$

where $H_{\mathbf{S}'\mathbf{S}} = \langle \mathbf{S} | H | \mathbf{S}' \rangle$. Our goal is to minimize the total energy, which can be achieved by the stochastic gradient method,³⁴ where the energy E and the gradients G are calculated via Monte Carlo sampling over spin configurations,

$$\begin{aligned} E &= \langle E_{\mathbf{S}} \rangle, \\ G &= \langle O_{\mathbf{S}} E_{\mathbf{S}} \rangle - \langle E_{\mathbf{S}} \rangle \langle O_{\mathbf{S}} \rangle, \end{aligned} \quad (8)$$

where $E_{\mathbf{S}} = \sum_{\mathbf{S}'} \frac{w(\mathbf{S}')}{w(\mathbf{S})} H_{\mathbf{S}'\mathbf{S}}$ and $O_{\mathbf{S}} = \frac{1}{w(\mathbf{S})} \frac{\partial w(\mathbf{S})}{\partial a_i}$. Here a_i are all the parameters in the CQNS.

However, the neural networks are highly nonlinear, therefore the system may be easily trapped in some local minima via the stochastic gradient descending method. To overcome this difficult, we optimize CQNS via a recently developed replica-exchange molecular dynamics method, which has been successfully applied to optimize the SBS,²⁸ one type of TNS. We map the quantum many-particle problem to a classical mechanical problem, in which we treat the parameters of CQNS as the generalized coordinates of the system. We optimize the energy of the virtual system using a replica exchange molecular dynamics method. By exchanging the system configurations among higher and lower temperatures, it can explore large phase space and therefore effectively avoid being stuck in the local minima.³⁵⁻³⁷ Details of the method are presented in Ref.²⁸

In our simulations, we use $M = 56$ temperatures. Initially the temperatures distribute exponentially between the highest ($1/\beta_0 = 10^{-3}$) and lowest ($1/\beta_0 = 10^{-5}$) temperatures. The MD step length is set to $\Delta t = 0.01$. For each temperature we start with random weights and biases. During the simulations, we adjust the temperatures after configuration exchange for 10 times, whereas there are 210 MD steps between the exchanges of two configurations. For each MD step, we sample 5000 spin configurations. The energies used for replica exchange are averaged over 200 MD steps.

We study J_1 - J_2 model on a 10×10 two-dimensional square lattice, with periodic boundary conditions. Figure 3 depicts the numerical results for $J_2 = 0$, and $J_2 = 0.5$, with different number of filters and filter sizes. The filter sizes are taken to be $K \times K$.

The energies per site for $J_2 = 0$ calculated with different number of filters and filter sizes are shown in Fig. 3(a). We also show the total energies calculated from other methods as comparisons. The dotted line denotes the ground energy per site obtained by the SBS, $E = -0.6581$,²⁷ and the dash-dotted line denotes the result of the entangled plaquette states $E = -0.6699$,³⁸ where as the dashed line is the result of quantum Monte carlo $E = -0.6715$.³⁸ The solid dots are the results obtained by CQNS. The energy for $K = 3$ and $M = 128$ filters is -0.6461. The energy can be improved by increasing

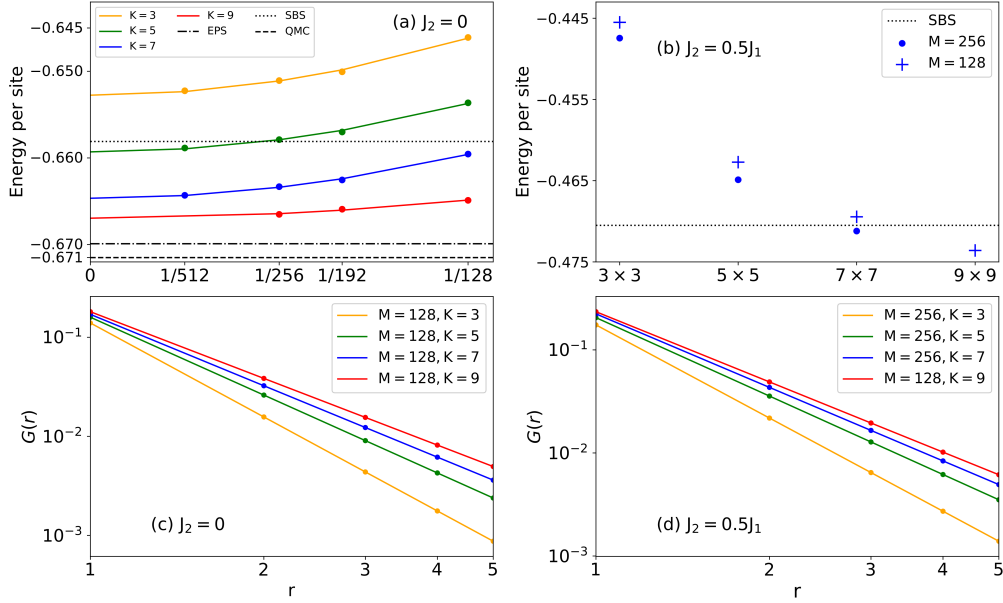


FIG. 3: (Color online) (a) The ground state energies calculated by CQNS at $J_2 = 0$ as functions of the number of filters M and size of filters K . The reference ground energies obtained by the string-bond-state (SBS), entangled-plaquette-state (EPS) and QMC are also shown in the dotted line (-0.6581), the dash-dotted line (-0.6699) and the dashed line (-0.6715), respectively. (b) The ground state energies calculated by CQNS at $J_2 = 0.5J_1$ as functions of M and K . The ground state energy obtained by SBS is shown as the dotted line (-0.4705). The fitted spin correlation function G_r are shown in (c) for $J_2=0$, and (d) for $J_2=0.5J_1$ for different number of filters M and size of filters K .

the number of filters. When using 512 filters, the energy decreases to -0.6522. However for the small filters, further increasing the number of filters does not improve the energy much. On the other hand, increasing filter size can dramatically improve the energy. For example, for filter number $M=128$, the energy is -0.6536 for $K = 5$, -0.6596 for $K = 7$ and -0.6649 for $K = 9$. As shown in the figure, when $K = 9$, the energy per site converges faster than using smaller filters, the energies for $M=128$, 192 and 256 filters are -0.6649, -0.6659 and -0.6665, respectively. Based on the numerical results, CQNS with filter size of $K=7, 9$ can exceed SBS. From the numerical results, it is revealed that larger filter number and larger filter size lead to more accurate ground states, however increasing filter size is more efficient than increasing filter number.

At $J_2 = 0.5$, the model is highly frustrated, where QMC suffers from sign problem, and therefore is much more difficult to calculate. In Fig. 3(b), we show the calculated energies with different filter size K and number of filters M . The ground state energy per site calculated by SBS is -0.4705, which is shown in dotted line as a comparison. According to the figure, the energy has limited improvement from $M=128$ to 256 filters. However increasing filter size can significantly improve the energy. For $M=128$, the energy is -0.4455 for $K=3$, -0.4627 for $K=5$ and -0.4736 for $K=9$. For $M=256$, the energy per site is -0.4475 for $K = 3$, -0.4649 for $K = 5$ and -0.4715 for $K = 7$. CQNS can exceed SBS with filter size $K \geq 7$.

We further calculate the spin correlation functions, and the results are shown in Fig. 3(c) and (d) for $J_2=0$, and

$J_2=0.5$ respectively. The error bars are smaller than the width of the solid lines and therefore not shown. The spin correlation functions $C(r)$ are calculated as,

$$C(r) = \frac{1}{2L^2} \sum_{i,j} (\langle \mathbf{s}_{i+r,j} \mathbf{s}_{i,j} \rangle + \langle \mathbf{s}_{i,j+r} \mathbf{s}_{i,j} \rangle). \quad (9)$$

The calculated results show that the correlation functions obey power law decay. We fit $C(r)$ as $C(r) = C_0 + \alpha r^{-\beta}$. We plot the correlation functions $G(r) = \alpha r^{-\beta}$, i.e., the constants have been subtracted. Because of the periodic boundary condition, the maximal distance is $r = L/2$. For $J_2 = 0$, the exponents β from $K = 3$ to $K = 9$ is 3.15, 2.61, 2.39 and 2.24. For $J_2/J_1 = 0.5$, β from $K = 3$ to $K = 9$ is 3.00, 2.53, 2.37 and 2.26. From these results, one may see that using small filters greatly underestimate the decay length of the spin correlation functions, and therefore can not represent the ground state very well. Increase filter size may describe more accurate the correlation functions and therefore the better ground state wave-functions and energies. To faithfully represent a quantum state, the size of the convolution filters should cover the correlation length. Otherwise the neural network can not capture the long quantum correlations.

The CQNS is expected to work well for quantum systems with short correlation lengths, e.g., gapped systems. On the other hand, to simulate very large systems, which is much larger than their correlation lengths ξ , the CQNS is also advantageous. One may use a suitable filter size K , which is larger than ξ , but still much less than the

system sizes, therefore the number of tunable parameters in CQNS can be greatly reduced. We note that in this work, only a single convolution layers is used. The long range correlation may be captured via small filters by deep learning, namely, more than one convolution layers are stacked. Extending CQNS with multiple convolution layers is our future interest.

In this work, we design a convolutional quantum neural network, namely CQNS, built from the basic elements of the state-of-the-art convolutional neural networks to study quantum many-particle problems. We use CQNS to study the spin-1/2 J_1 - J_2 Heisenberg model which is strongly frustrated and challenging to solve. We have obtained fairly accurate ground state energies that are even better than previous SBS calculations.^{27,28} The CQNS is a general and flexible method that can be easily applied to different Hamiltonian, with various boundary conditions (open, periodic and cylindrical etc.) and in different dimensions. It therefore provide a new powerful tool to study the long-standing quantum many-particle problems.

The CQNS used in this paper is as simple as three

layers, where the first convolution layer plays a critical role, and the size of the filters must be large enough to capture the long range correlations. We note that there could be different convolutional network structures, and the present CQNS may not necessary be the optimal one. Especially deep neural network may further enhance the state representation ability, as it is possible to capture the long range correlation by multiple convolution layers with smaller filters in each layer. It recently has been proposed that deep CNN may efficiently represent the many-particle states that even beyond area law.²⁴ Design a practical deep CQNS is an interesting and promising route to after.

The authors appreciate insightful discussions with Shao-Jun Dong and Chao Wang. The construction and computation of neural networks in this paper is performed by the mainstream deep learning framework PyTorch. The numerical calculations have been done on the USTC HPC facilities. This work was supported by the National Natural Science Foundation of China (Grant No. 11674306, 61590932), National key R&D program (No. 2016YFA0301300, 2016YFA0317300).

-
- * Electronic address: yshzhang@ustc.edu.cn
† Electronic address: helx@ustc.edu.cn
- ¹ A. Krizhevsky, L. Sutskever and G. E. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*, Advances in Neural Information Processing Systems **25**, (2012).
 - ² D. Silver, *et. al.* *Mastering the game of Go with deep neural networks and tree search*, Nature **529**, 484-489 (2016).
 - ³ D. Silver, *et. al.* *Mastering the game of Go without human knowledge*, Nature **550**, 354-359 (2017).
 - ⁴ A. A. Melnikov, H. O. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger and H. J. Briegel, *Active learning machine learns to create new quantum experiments*, Proc. Natl. Acad. Sci. **115**, 1221 (2018).
 - ⁵ R. Biswas, *et. al.* *Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data*, Phys. Rev. D **88**, 062003 (2013).
 - ⁶ H. Gabbard, M. Williams, F. Hayes and C. Messenger, *Matching matched filtering with deep networks in gravitational-wave astronomy*, Phys. Rev. Lett. **120**, 141103 (2018).
 - ⁷ M. Krenn, J. Handsteiner, M. Fink, R. Fickler, R. Ursin, M. Malik and A. Zeilinger, *Twisted light transmission over 143 km*, Proc. Natl. Acad. Sci. **113**, 13648 (2016).
 - ⁸ L. Wang, *Discovering phase transitions with unsupervised learning*, Phys. Rev. B **94**, 195105 (2016).
 - ⁹ C. Wang and H. Zhai, *Unsupervised Learning of Frustrated Classical Spin Models 1: Principle Component Analysis*, Phys. Rev. B **96**, 144432 (2017).
 - ¹⁰ K. Chng, J. Carrasquilla, R. G. Melko and E. Khatami, *Machine Learning Phases of Strongly Correlated Fermions*, Phys. Rev. X **7**, 031038 (2017).
 - ¹¹ K. Mills, M. Spanner and I. Tamblyn, *Deep learning and the Schrödinger equation*, Phys. Rev. A **96**, 042113 (2017).
 - ¹² K. Mills and I. Tamblyn, *Deep neural networks for direct, featureless learning through observation: the case of 2d spin models*, Phys. Rev. E **97**, 032119 (2018).
 - ¹³ P. Zhang, H. Shen and H. Zhai, *Machine Learning Topological Invariants with Neural Networks*, Phys. Rev. Lett. **120**, 066401 (2018).
 - ¹⁴ A. Tanaka and A. Tomiya, *Detection of phase transition via convolutional neural network*, J. Phys. Soc. Jap. **86**, 063001 (2017).
 - ¹⁵ H. Saito, *Solving the Bose-Hubbard model with machine learning*, J. Phys. Soc. Jap. **86**, 093001 (2017).
 - ¹⁶ G. Carleo and M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science **355**, 602 (2017).
 - ¹⁷ A. Voigt, J. Richter and N. B. Ivanov, *Marshall-Peierls sign rule for excited states of the frustrated J1-J2 Heisenberg antiferromagnet*, Physica A **245**, 269-275(1997).
 - ¹⁸ J. Biamonte and V. Bergholm, *Tensor Networks in a Nutshell*, arXiv: 1708.00006v1.
 - ¹⁹ D. L. Deng, X. Li and S. DasSarma, *Quantum Entanglement in Neural Network States*, Phys. Rev. X **7**, 021021 (2017).
 - ²⁰ I. Glasser, N. Pancotti, M. August, I. D. Rodriguez and J. I. Cirac, *Neural-Network Quantum States, String-Bond States, and Chiral Topological States*, Phys. Rev. X **8**, 011006 (2018).
 - ²¹ J. Chen, S. Cheng, H. Xie, L. Wang and T. Xiang, *Equivalence of restricted Boltzmann machines and tensor network states*, Phys. Rev. B **97**, 085104 (2018).
 - ²² X. Gao and L. M. Duan, *Efficient representation of quantum many-body states with deep neural networks*, Nat. Commun. **8**, 662 (2017).
 - ²³ Y. LeCun, L. Botton, Y. Bengio and P. Haffner, *Gradient-Based Learning Applied to Document Recognition*, Proceedings of the IEEE, **86**, 2278-2324 (1998).
 - ²⁴ Y. Levine, O. Sharir, N. Cohen and A. Shashua, *Bridg-*

- ing Many-Body Quantum Physics and Deep Learning via Tensor Networks, arXiv: 1803.09780v1.
- ²⁵ J. Carrasquilla and R. G. Melko, *Machine learning phases of matter*, Nat. Phys. **13**, 431-434 (2017).
 - ²⁶ L. Wang, Z.-C. Gu, F. Verstraete and X.-G. Wen, *Tensor-product state approach to spin-1/2 square J_1 - J_2 antiferromagnetic Heisenberg model: Evidence for deconfined quantum criticality*, Phys. Rev. B **94**, 075143 (2016).
 - ²⁷ A. Sfondrini, J. Cerrillo, N. Schuch and J. I. Cirac, *Simulating two- and three-dimensional frustrated quantum systems with string-bond states*, Phys. Rev. B **81**, 214426 (2010).
 - ²⁸ W. Liu, C. Wang, Y. Li, Y. Lao, Y. Han, G. C. Guo, Y. H. Zhao and L. He, *Replica exchange molecular dynamics optimization of tensor network states for quantum many-body systems*, J. Phys. Condens. Matter **27**, 085601 (2015).
 - ²⁹ D. Scherer, A. Müller and S. Behnke, *Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition*, 20th International Conference on Artificial Neural Networks (ICANN), Thessaloniki, Greece, September 2010.
 - ³⁰ L. Wang, D. Poilblanc, Z.-C. Gu, X.-G. Wen, and F. Verstraete, *Constructing a Gapless Spin-Liquid State for the Spin-1/2 J_1J_2 Heisenberg Model on a Square Lattice*, Phys. Rev. Lett. **111**, 037202 (2013).
 - ³¹ W. Liu, S. Dong, C. Wang, Y. Han, H. A.n, G. C. Can and L. He, *Gapless spin liquid ground state of spin-1/2 J_1 - J_2 Heisenberg model on square lattices*, arXiv: 1806.07031v1 (2018).
 - ³² W. J. Hu, F. Becca, A. Parola and S. Sorella, *Direct evidence for a gapless Z_2 spin liquid by frustrating Néel antiferromagnetism*, Phys. Rev. B **88**, 060402 (2013).
 - ³³ S. S. Gong, W. Zhu, D. N. Sheng, O. I. Motrunich and Matthew P. A. Fisher, *Plaquette Ordered Phase and Quantum Phase Diagram in the Spin-1/2 J_1 - J_2 Square Heisenberg Model*, Phys. Rev. Lett. **113**, 027201 (2014).
 - ³⁴ H. Robbins and S. Monro, *A Stochastic Approximation Method*, The Annals of Mathematical Statistics, **22** 400-407 (1951).
 - ³⁵ K. Cao, Z.-W. Zhou, G.-C. Guo and L. He, *Efficient numerical method to calculate the three-tangle of mixed states*, Phys. Rev. A **81**, 034302 (2010).
 - ³⁶ K. Cao, G.-C. Guo, D. Vanderbilt and L. He, *First-Principles Modeling of Multiferroic RMn_2O_5* , Phys. Rev. Lett. **103**, 257201 (2009).
 - ³⁷ E. Marinari, G. Parisi and J. J. Ruiz-Lorenzo, *Phase structure of the three-dimensional Edwards-Anderson spin glass*, Phys. Rev. B **58**, 14852 (1998).
 - ³⁸ F. Mezzacapo, N. Schuch, M. Boninsegni and J. I. Cirac, *Ground-state properties of quantum many-body systems: entangled-plaquette states and variational Monte Carlo*, New. J. Phys. **11**, 083026 (2009).