Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Solving the Hubbard model with Neural Quantum States

Dingshun Lv

ywlds@163.com

ByteDance Seed https://orcid.org/0000-0003-0573-6490

Yuntian Gu

Peking University

Wenrui Li

Peking University

Heng Lin

Tsinghua University

Bo Zhan

Institute of Physics, Chinese Academy of Sciences

Ruichen Li

ByteDance

Yifei Huang

ByteDance Seed

Di He

Peking University

Yantao Wu

Institute of Modern Physics, Chinese Academy of Sciences

Tao Xiang

Institute of Physics, Chinese Academy of Sciences https://orcid.org/0000-0001-5998-7338

Mingpu Qin

Key Laboratory of Articial Structures and Quantum Control, School of Physics and Astronomy, Shanghai Jiao Tong University

Liwei Wang

Peking University

Article

Keywords: Neural Quantum States, Hubbard Model, Stripe order, High-Tc superconductivit

Posted Date: November 19th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7215963/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Solving the Hubbard model with Neural Quantum States

Yuntian Gu[1,2†], Wenrui Li[1,2†], Heng Lin[2,3†], Bo Zhan[2,4†], Ruichen Li[1,2], Yifei Huang[2], Di He[1*], Yantao Wu[4*], Tao Xiang[4*], Mingpu Qin[5*], Liwei Wang[1*], Dingshun Lv[2*]

[1*]State Key Laboratory of General Artificial Intelligence, School of Intelligence Science and Technology, Peking University, China.
[2*]ByteDance Seed, China.
[3]State Key Laboratory of Low-Dimensional Quantum Physics, Department of Physics, Tsinghua University, China.
[4*]Institute of Physics, Chinese Academy of Sciences, China.
[5*]Key Laboratory of Artificial Structures and Quantum Control (Ministry of Education), School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China.

*Corresponding author(s). E-mail(s): dihe@pku.edu.cn; yantaow@iphy.ac.cn; txiang@iphy.ac.cn; qinmingpu@sjtu.edu.cn; wanglw@pku.edu.cn; lvdingshun@bytedance.com;
†These authors contributed equally to this work.

Abstract

The rapid development of neural quantum states (NQS) has established it as a promising framework for studying quantum many-body systems. In this work, by leveraging the cutting-edge transformer-based architectures and developing highly efficient optimization algorithms, we achieve the state-of-the-art results for the doped two-dimensional (2D) Hubbard model, arguably the minimum model for high-Tc superconductivity. Interestingly, we find different attention heads in the NQS ansatz can directly encode correlations at different scales, making it capable of capturing long-range correlations and entanglements in strongly correlated systems. With these advances, we establish the half-filled stripe in the ground state of 2D Hubbard model with the next nearest neighboring hopping, consistent with experimental observations in cuprates. Our work establishes NQS as a powerful tool for solving challenging many-fermions systems.

Keywords: Neural Quantum States, Hubbard Model, Stripe order, High-Tc superconductivity.

1 Main

The Hubbard model [1–3] is the iconic model for studying the correlation effect in many-electron systems. Though its Hamiltonian is simple, the physics it can harbor is extremely rich, including quantum magnetism, Mott insulator, charge and spin density waves, and so on. Following the discovery of cuprate high-Tc superconductors, the two-dimensional (2D) Hubbard model has attracted increasing attention due to its potential connection to the microscopic mechanism of high-Tc superconductivity [4, 5]. The study of the 2D Hubbard model relies mainly on numerical many-body methods due to the lack of analytic solutions. In the past few decades, a variety of many-body methods were developed, including density matrix renormalization group (DMRG) [6] and its higher dimension generalization in the language of tensor network states [7, 8], such as projected entangled pair states (PEPS) [9], quantum embedding methods such as dynamical mean-field theory (DMFT)[10], density matrix embedding theory (DMET)[11], and different flavors of quantum Monte Carlo approaches [12–14] and so on. In the past, different methods usually gave inconsistent results due to the limitations of each method [15]. In recent years, large scale cross-validation has been performed among different methods, and consensus has begun to reach about the ground state of the doped 2D Hubbard model, with [16] and without the next nearest neighboring hopping [17].

On another front, researchers are also striving to develop more accurate and efficient new methods, aiming to provide more reliable results for strongly correlated systems like the Hubbard model to unveil the underlying exotic physics. In recent years, benefiting from the rapid development in the fields of neural networks and machine learning, Neural Quantum States (NQS) have emerged as a promising new approach for solving quantum many-body systems. Significant progress has been made in the development of advanced NQS methods in quantum systems since their inception by Carleo et al. [26], which have markedly improved the accuracy of NQS approaches, establishing them as a powerful framework for investigating systems in quantum chemistry [27–29] and interacting spin models [30–32], as shown in Figure 1(A). In light of these significant achievements, a crucial question arises: how effective is this new paradigm for the study of the more challenging strongly correlated fermionic lattice models?

In this direction, extensive effort has been made to investigate the utilization of NQS methods in the context of the Hubbard model, particularly in the most challenging 2D lightly doped cases. Early attempts to apply NQS approaches to the 2D Hubbard model[18] faced challenges in accurately representing fermionic nodal structures. Subsequent work by Luo et al.[20] introduced a neural network backflow ansatz that substantially improved the representation power of the wavefunction, demonstrating its effectiveness on systems as large as 12×8. More recently in 2022, Robledo et

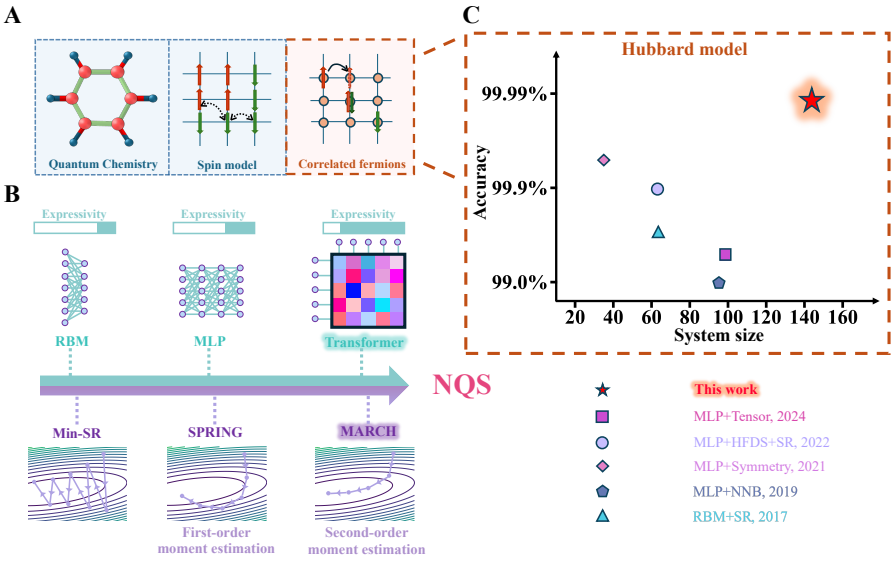


Fig. 1 Application, architecture, optimization, and performance of Neural Quantum States (NQS) on the Hubbard model. (A) Guided by the variational principle, NQS are powerful and versatile, applicable to a wide range of quantum systems, including ab initio quantum chemistry, spin models, and correlated fermionic models. (B) Various neural network architectures can be used to construct NQS, with increasing expressivity from Restricted Boltzmann Machines (RBM) [18], Multi-Layer Perceptrons (MLP) [19–22] to Transformers. These networks are optimized using methods inspired by quantum information, including Min-SR [23, 24], SPRING [25], and the advanced MARCH method proposed in this work. (C) A performance comparison of different NQS methods on the 2D Hubbard model, plotting simulation accuracy against system size. We benchmark the accuracy on systems where exact solutions are available (see Supplementary Section C for details). This work achieves higher accuracy on larger system sizes than previously reported in other notable NQS studies, demonstrating a significant advancement in the field.

al. [21] achieved highly accurate results for various geometries of the Hubbard model up to 8×8 by combining Hidden-Fermion Determinant States (HFDS) with Stochastic Reconfiguration (SR) [33–35]. Despite this notable progress, scaling NQS simulations to even larger, more physically relevant parameter regimes is hindered by network architectural limitations and optimization barriers[22, 36].

In this work, we make significant progress in expanding the scalability and accuracy of the NQS approach for solving the Hubbard model with two key innovations as shown in Figure 1(B). First, a novel NQS architecture is proposed, which leverages the self-attention mechanism [37], a technique proven to be very effective in scaling up [38]. The built-in long-range structure makes the NQS ansatz different from tensor network states, which consist of only local tensors. This ensures that the NQS can precisely capture long-range correlations and entanglements in strongly correlated systems and can accurately handle larger systems, which is infeasible for other neural

network architectures like Restricted Boltzmann Machine (RBM) [18] and Multi-Layer Perceptron (MLP) [19–22]. Second, we propose an enhanced optimization algorithm, termed Moment-Adaptive ReConfiguration Heuristic (MARCH), designed to achieve substantially faster and more stable convergence than previous optimization algorithms such as SR [33–35], Min-SR [23, 24] and SPRING [25], especially for large and frustrated systems.

With these architectural and algorithmic advancements, we achieve state-of-the-art variational ground state energy of the 2D Hubbard model at unprecedented sizes up to 16×16, for both open and periodic boundary conditions, as shown in Figure 1(C). As a result, we confirm the filled stripe ground state in the pure Hubbard model and identify a half-filled stripe ground state with the next nearest neighboring hopping $t' = -0.2$, consistent with experimental observations in cuprate [39]. Moreover, in the $t' = -0.2$ case, we find the stripe is horizontal, i.e., stripe arrangements prefer the longer side of the geometry. This is quite different from previous calculations on cylinders [16], indicating the effectiveness of our method in eliminating the boundary effects. We also have a careful analysis of the structure of our NQS ansatz and find that different attention heads can directly encode correlations at different scales after optimization, making it capable to capture long-range correlations and entanglements in the Hubbard model. Our developments in architectures and optimization strategies establish NQS as a powerful and transformative paradigm for solving the Hubbard model. This framework extends readily to a broader range of strongly correlated systems, offering a new lens to explore new exotic quantum many-body states resulting from strong correlations.

2 Results

2.1 A transformer-based NQS framework

The Hamiltonian of the Hubbard model is:

$$\hat{H} = -t \sum_{\langle ij \rangle, \sigma} \hat{c}^\dagger_{i\sigma} \hat{c}_{j\sigma} - t' \sum_{\langle\langle ij \rangle\rangle, \sigma} \hat{c}^\dagger_{i\sigma} \hat{c}_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}. \tag{1}$$

Here $\hat{c}^\dagger_{i\sigma}$ ($\hat{c}_{j\sigma}$) is the creation (annihilation) operator for an electron with spin σ at site $i(j)$, $\hat{n}_{i\sigma} = \hat{c}^\dagger_{i\sigma} \hat{c}_{i\sigma}$ is the particle number operator. The terms $\langle ij \rangle$ and $\langle\langle ij \rangle\rangle$ denote pairs of nearest neighboring and next nearest neighboring sites, respectively. We study the 2D Hubbard model defined on square lattices with size $L_x \times L_y$. Both open boundary conditions (OBC) and periodic boundary conditions (PBC) are adopted in this work. We set the nearest neighboring hopping amplitude $t = 1$ as the unit of energy. The on-site Coulomb repulsion is fixed at $U = 8$, a value relevant to cuprates [40, 41]. We consider both the pure Hubbard case ($t' = 0$) and the case with the next nearest neighboring hopping ($t' = -0.2$). In this work, we focus on the lightly doped regime with hole concentration of $\delta = 1/8$, which is the most intriguing region in both the 2D Hubbard model and in cuprates.

Figure 2 presents an overview of our NQS algorithm. As shown in Figure 2(A), we employ a transformer-based NQS, parameterized by θ, to represent the amplitude

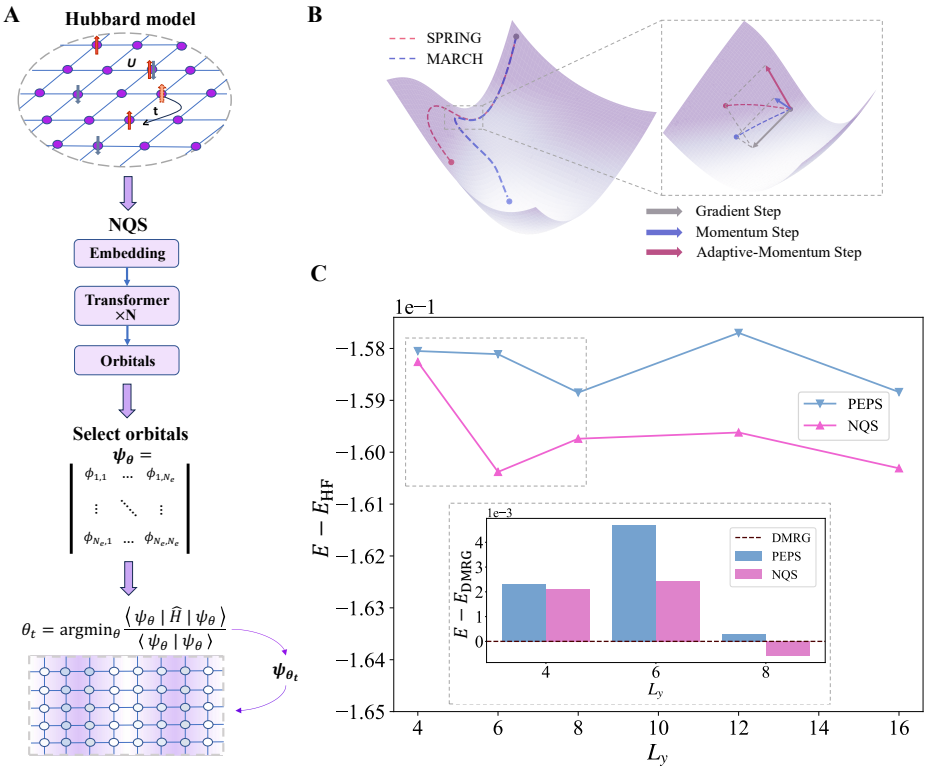


Fig. 2 NQS architecture, optimization scheme, and benchmark performance on the Hubbard model. (**A**) An input lattice configuration is processed through a transformer-based neural network to generate backflow orbitals. These orbitals are then used to construct the NQS wavefunction ψ_θ. The variational parameters θ are optimized to find the ground state. (**B**) Optimization trajectories for MARCH and SPRING. MARCH follows a more direct path by dynamically adapting its step size for each parameter. The inset shows how SPRING's reliance on momentum might lead to overshooting, whereas MARCH's adaptive-momentum step prevents oscillations and accelerating convergence to the minimum. (**C**) Comparison of ground-state energies with PEPS (bond dimension $D \geq 20$) [42] for the pure Hubbard model on $16 \times L_y$ system with OBC. Our NQS results consistently achieve lower variational energies than PEPS across different lattice geometries. The inset shows we also outperform reference DMRG energies [42] (with 32000 SU(2) multiplets) when $L_y = 8$.

$\psi_\theta(\mathbf{n})$ associated with each configuration $|\mathbf{n}\rangle$. Our transformer-based NQS assigns a unique high-dimensional embedding vector to each site i. These embeddings are subsequently processed through a series of multi-head attention blocks, the core component of modern large language models [43]. The attention block is defined as:

$$\text{Attn}(Q, K, V) = \text{softmax}(QK^\top / \sqrt{d_H})V, \tag{2}$$

where $Q, K, V \in \mathbb{R}^{N \times d_H}$ are feature matrices produced by neural networks. Here, N represents the number of sites in the Hubbard model, d_H is the hidden dimension of the attention block. The i-th row of these matrices corresponds to the i-th site in

the Hubbard model, arranged in dictionary order. The normalized attention scores S_{attn}, given by $\text{softmax}(QK^\top/\sqrt{d_H})$, allow the block to aggregate information from all sites by assigning weights. Thus, the attention block can capture complex, long-range correlations, which is crucial for modeling strongly correlated systems like the Hubbard model. The multi-head attention block employs distinct feature matrices to capture diverse correlation patterns in the system. After self-attention layers, a linear layer is employed to produce backflow orbitals $M(\mathbf{n})$. The amplitude is then derived as the determinant of orbitals occupied by electrons.

In our work, we aim to find the ground state by minimizing the total energy of the wavefunction. In previous NQS studies [26], SR-based optimizers have been used to minimize the energy. These optimizers effectively simulate the imaginary time evolution, enhancing the convergence rate over the standard first-order optimizer by exploiting the quantum geometric tensor [44]. One particular method is the SPRING algorithm [25], which leverages the history of gradients to accelerate convergence and smooth out the optimization path. Building on this, we introduce the MARCH algorithm. MARCH enhances SPRING by also incorporating an estimate of the second moment of the gradients, drawing a direct parallel to how the popular Adam optimizer [45] improves upon Momentum [46]. This addition of second-moment information is crucial for navigating the complex and high-dimensional energy landscapes that typically exist in quantum systems. It provides an estimate of the gradient's variance, allowing MARCH to adapt the learning rate for each parameter individually. This adaptive capability offers a significant advantage, particularly when encountering challenging regions like saddle points or flat plateaus where the gradient is small. While a momentum-based method like SPRING might overshoot, MARCH can use the second-moment information to maintain progress and "march" confidently out of these areas. As illustrated in Figure 2(B), this results in a more direct and efficient path to the energy minimum, leading to faster and more robust convergence.

These combined architectural and algorithmic advancements enable us to accurately tackle 2D Hubbard model with unprecedented sizes beyond the reach of previously NQS methods. We first perform benchmark calculations for systems under PBC for the pure Hubbard model ($t' = 0$) at half-filling, where numerical exact ground state energies can be obtained by Auxiliary Field Quantum Monte Carlo (AFQMC) [47]. The NQS energies are consistent with the AFQMC results within twice the statistical errors (see Supplementary Table 1).

We then benchmark the new NQS method on the doped pure Hubbard model under OBC with varying sizes L_x and L_y. Direct comparisons of ground-state energies with state-of-the-art DMRG and PEPS [42] results are presented in Figure 2(C) (with Hartree-Fock energies as reference). In narrow systems such as $L_y \leq 6$, the DMRG energy is the lowest. However, when reaching the 2D limit by increasing the width of the system, where DMRG fails to capture the area law of entanglement entropy, the NQS results surpass DMRG. Noticeably, in the entire range of L_y we studied, NQS achieves lower variational energies than those reported by state-of-the-art PEPS with bond dimension $D \geq 20$ [42].

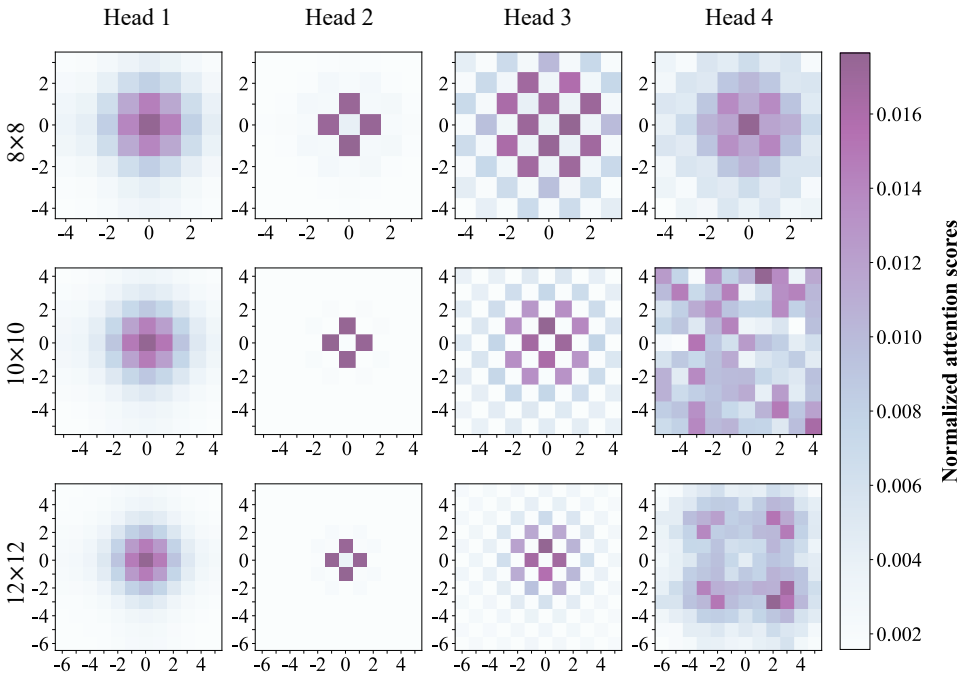


Fig. 3 Visualization of the attention mechanism. Shown is how a lattice site attends to another target site in the 2D pure Hubbard model for lattice sizes of 8×8, 10×10, and 12×12 at half filling. The four attention heads, which are equivalent and have been rearranged by pattern for clarity, each correspond to a specific physical aspect: short-range correlations, nearest-neighbor hopping, antiferromagnetic correlations, and complex and delocalized correlations in the ground state.

2.2 Attention mechanism reveals physical correlations

To gain deeper insight into how the transformer architecture captures the intricate physics of the Hubbard model, we examine the attention patterns learned by the network [48]. In our transformer ansatz, the heads after optimization are found to specialize in capturing the correlations in different scales present in the ground-state wave function. This emergent specialization is analogous to observations in fields like natural language processing, where different attention heads in a transformer learn distinct, interpretable patterns such as syntactical dependencies [37]. In Figure 3, we visualize the normalized attention scores S_{attn} for the half-filling systems, which quantify the learned pairwise importance between sites, from the first layer for representative systems. Similar results for the doped cases or deeper layers can be found in Supplementary Figure 1-10. The results are striking: the first head clearly learns to prioritize local interactions, with S_{attn} that decay with the distance between sites, capturing short-range correlations. The second head specializes in nearest-neighbor hopping, with its attention focusing almost exclusively on the four adjacent sites, which is a

fundamental process in the Hubbard model. The third head successfully identifies the emergent antiferromagnetic order, assigning high attention scores in a checkerboard pattern characteristic of this magnetic correlation. The final head appears to capture more complex, delocalized phase information, exhibiting a sophisticated pattern that is crucial for accurately representing the full complexity of the wave function. This specialization demonstrates the remarkable ability of the transformer's attention mechanism to autonomously discover and disentangle the fundamental physical correlations governing the system.

2.3 Filled stripe order in pure Hubbard model

The ground state of the lightly doped pure ($t' = 0$) 2D Hubbard model was established as the filled stripe before [16]. Recently Liu et al. [42] observed a crossover of the stripe direction with the increase of system widths for systems under OBC, indicating the existence of a large finite size effect in systems with open boundaries. We performed systematic calculations of the pure Hubbard model for different sizes under both OBC and PBC. For systems with OBC, we find stripe states that match well with the results obtained from PEPS but with lower energies (see Figure 2(C)). In particular, our results also demonstrate a crossover between horizontal and vertical stripes as a function of system width (from $L_y = 8$ to $L_y = 12$, see Supplementary Figure 14-15). But when we switch to PBC, the crossover vanishes (see Supplementary Figure 17-18), indicating the crossover is a finite size and open boundary effect. The ability to handle systems under PBC with NQS can mitigate the finite size effect, allowing to obtain more reliable results in the thermodynamic limit. By studying systems under PBC, we observe a robust, filled stripe phase with wavelength of $\lambda = 8$ as shown in Figure 4(A), a result that is consistent with previous findings [16]. Simulations on smaller 8×8 and 12×12 lattices (see Supplementary Figure 19-20) show clear finite-size effects due to their incommensurability with the wavelength $\lambda = 8$.

2.4 Half-filled stripe order in t'-Hubbard model

It was found that the doped pure Hubbard model doesn't host superconductivity [49]. To pursue superconducting ground states, the simplest modification is to include the next nearest neighboring hopping (t') term. On one hand, a finite t' can account for the electron and hole doping asymmetry in the phase diagram of cuprates. On the other hand, electronic structure calculations [40, 50] show a finite $t' \approx -0.2$ term indeed exists in the one band effective Hamiltonian of cuprates. A recent study by Xu et al. [17] found that the phase diagram of the Hubbard model with $t' = -0.2$ agrees well with the cuprates, where fractional filling stripe coexists with the d-wave superconductivity in the hole doping region. The t' term frustrates the anti-ferromagnetic order in the parent compound and makes the doped stripe state fragile. Importantly, large finite size and boundary effect [51] was found for the wavelength of the stripe state in the hole doping region, which fluctuates in the vicinity of $\lambda = 4$ [17].

We perform systematic calculations for systems under PBC with $t' = -0.2$. With PBC, we can get rid of the boundary effect. Our calculations on large sizes show the ground state for the $t' = -0.2$ Hubbard model is clearly the half-filled stripe, i.e., with

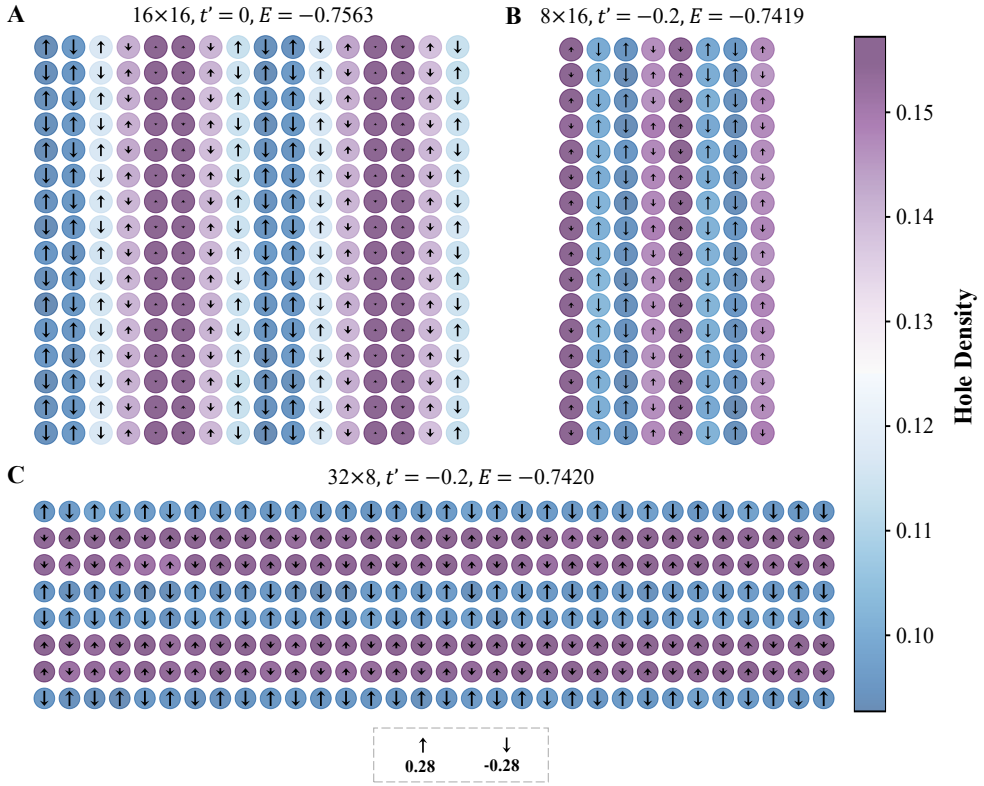


Fig. 4 Hole and spin density distributions in the ground state for different lattice sizes with $U = 8$ and hole doping $\delta = 1/8$ under PBC in the Hubbard model. The magnitudes of the spin density are represented by the sizes of the arrows while the direction is denoted by the direction of the arrows. Hole density is depicted using a color scale. Stripes can be clearly seen in all three systems. (A) Results for 16×16 pure Hubbard model where the stripe state with $\lambda = 8$ is observed. (B) Results for 16×12 system with $t' = -0.2$ where the stripe state with $\lambda = 4$ is observed. (C) Results for 32×8 system with $t' = -0.2$ where the stripe state with $\lambda = 4$ is observed. Notice that the stripes in the ground state for $t' = -0.2$ in (B) and (C) are horizontal, even though both vertical and horizontal stripes are allowed in the systems. More results can be found in the supplementary materials.

wavelength $\lambda = 4$, as shown in Figure 4(B). These results indicate that the previous found fluctuating wavelengths for stripe may result from the finite size and boundary conditions. The $\lambda = 4$ stripe agrees well with the experimental observation of cuprates [39], providing further evidence to support the relevance of the t' Hubbard model to cuprates.

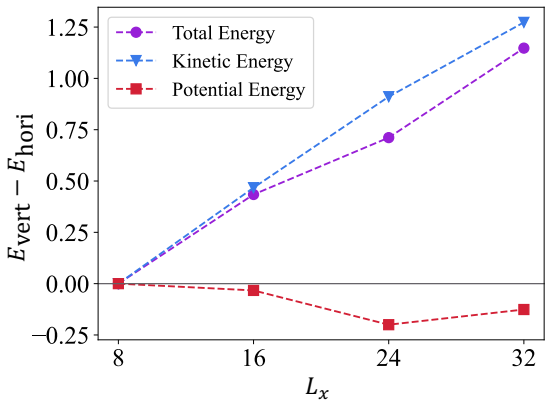


Fig. 5 Energy difference between vertical stripe and horizontal stripe. Kinetic energy, potential energy, and total energy results are all shown. We systematically study systems $L_x \times 8$ with L_x ranging from 8 to 32, with $t' = -0.2$ and $\delta = 1/8$. We can find that the horizontal stripe has lower energy for all the systems (notice that for 8×8 system, vertical and horizontal stripes are the same). Moreover, the results show that the gain of energy in the horizontal stripe is mainly from the kinetic energy.

2.5 Horizontal stripes are energetically favoured over vertical stripes

Previous numerical studies on stripe order in the Hubbard and related models have been conducted on cylindrical geometries with DMRG [52–56]. However, such boundary conditions and their related finite-size effect can inherently favor the formation of vertical stripe orientations. To alleviate the possible boundary effect, we study systems with PBC in both directions. We systematically study systems with size $L_x \times 8$ where L_x is ranged from 8 to 32, with $t' = -0.2$ and $\delta = 1/8$ (see Equation 1). We consistently observe that a horizontal stripe configuration is energetically favored in the ground state while the system can accommodate half-filled stripe in both horizontal and vertical directions (see Figure 4(B) and (C), Figure 5 and Supplementary Figure 21-26).

These results suggest the existence of a tendency toward fewer long stripes rather than multiple fragmented ones in the systems, different from previous studies on cylinders [17]. From Figure 5, we can also find that the gain of energy for the horizontal stripe state is mainly from the kinetic energy, while the potential energy in the horizontal stripe is slightly higher. A simple explanation is that the formation of a long "river" of holes in the horizontal stripe is beneficial to the kinetic energy. A comprehensive understanding of this observation needs further investigation.

For rectangular systems with a width not divisible by 8, a horizontal stripe with wavelength $\lambda = 4$ is impossible, so "frustration" appears. One such example is the 16×10 system, for which we find that the trend for horizontal stripe wins and the ground state has a wavelength of $\lambda = 5$ while the vertical stripe state with $\lambda = 4$ has a higher energy (see Supplementary Figure 28-29). In another example for the 16×12

system, the trend for wavelength 4 wins in the competition and the ground state is a vertical stripe (see Supplementary Figure 30).

For the pure Hubbard model, it is also possible that the true ground state is a horizontal stripe. But the resolution of this issue requires the study of systems that can accommodate stripes with a wavelength 8 in both vertical and horizontal directions. The smallest size for such systems under PBC is 32×16, which is beyond our present capacity.

3 Discussion

In this work, we propose a new NQS framework for studying quantum many-body systems by leveraging the cutting-edge machine learning architectures and developing highly accurate optimization algorithms. State-of-the-art results for the 2D doped Hubbard model with unprecedented sizes are achieved with this new framework. NQS enables the study of 2D Hubbard model under periodic boundary conditions, which is crucial to get rid of the boundary effect and makes the extrapolation to the thermodynamic limit reliable. By studying 2D Hubbard model with large systems under periodic boundary conditions, we confirm the filled stripe phase in the ground state of the doped pure ($t' = 0$) Hubbard model. When the next nearest neighboring hopping term ($t' = -0.2$) is included, the ground state is found to be half-filled stripes, consistent with experimental observation in cuprates. Noticeably, the direction of the stripe for $t' = -0.2$ is determined to be horizontal when the width and the length are not equal, different from previous studies [17]. These results indicate that previous calculations on systems with open boundary suffer from the boundary effect and underscore the effectiveness of the new NQS method in gaining more reliable results.

We also have a careful analysis of the attention patterns learned in the ground state of 2D Hubbard model. Interestingly, we find that after the energy optimizations, different heads within our NQS specialize in encoding the correlations at different scales of the ground-state wave function. This special structure enables our NQS ansatz to effectively capture long-range correlations and entanglements, which is crucial for the accurate calculation of strongly correlated quantum many-body systems.

There is much room for the future development of NQS. On one hand, the development of more efficient optimization approaches could enable the study of even large system sizes, which is crucial for more reliable finite size scaling in the doped Hubbard model. On the other hand, more sophisticated NQS ansatz for the calculation of dynamic and finite temperature properties [57] are needed to further unveil the physics of the Hubbard model and establish the microscopic theory for cuprate. We anticipate that NQS framework will continue benefiting from the development in the broad machine learning community. Nevertheless, our results demonstrate that NQS is already a powerful framework for solving challenging many-fermions systems.

4 Methods

4.1 Variational Monte Carlo

VMC is a quantum Monte Carlo method based on the variational principle of quantum mechanics. Upon choosing a wavefunction ansatz ψ_θ, minimizing the energy of this ansatz with respect to the parameters θ makes ψ_θ an approximation to the ground state. In this work, we adopt a neural network as the wavefunction ansatz. The neural network takes occupation number $\mathbf{n}$ as input and outputs the amplitude of the wavefunction, $\langle \mathbf{n}|\psi_\theta\rangle = \psi_\theta(\mathbf{n})$. The energy of the wavefunction can be written as the expectation value of the local energy, i.e.

$$\bar{E}_\theta = \frac{\langle\psi_\theta|\hat{H}|\psi_\theta\rangle}{\langle\psi_\theta|\psi_\theta\rangle} = \mathbb{E}_{\mathbf{n}\sim\psi_\theta^2/\|\psi_\theta\|^2}[E^\theta_{\mathrm{loc}}(\mathbf{n})], \tag{3}$$

where

$$E^\theta_{\mathrm{loc}}(\mathbf{n}) = \sum_{\mathbf{n}'} \frac{\psi_\theta(\mathbf{n}')}{\psi_\theta(\mathbf{n})} \langle \mathbf{n}'| \hat{H} |\mathbf{n}\rangle \tag{4}$$

is the local energy. Exact calculation of the energy using Equation 3 is infeasible in an exponentially large Hilbert space. However, we can approximate this summation using the Monte Carlo method. A set of occupation numbers $\{\mathbf{n}_i\}_{i=1}^B$ is sampled from the probability distribution defined by the squared amplitude of the wavefunction using Markov chain Monte Carlo (MCMC) techniques. The local energy of occupation numbers, $E^\theta_{\mathrm{loc}}(\mathbf{n})$, can be efficiently calculated from the Hubbard model Hamiltonian, with the number of summation terms in Equation 4 scaling linearly with the size of the system. Therefore, parameters of the neural network can be optimized to approachthe ground state of the Hubbard model:

$$\theta^* = \arg\min_\theta \bar{E}_\theta. \tag{5}$$

4.2 Wavefunction ansatz

As illustrated in Figure 2(A), we employ a neural quantum state (NQS) ansatz to represent the many-body quantum state $|\psi\rangle$ of a fermionic system with N lattice sites. Each individual site corresponds to a 4 dimensional physical Hilbert space spanned by $|0\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, and $|\uparrow\downarrow\rangle$. We encode the four possible tokens into a d dimensional vector via an embedding matrix $E \in \mathbb{R}^{4\times d}$. To account for spatial arrangement, crucial due to the permutation equivariance of self-attention, these embeddings are augmented with learnable positional encodings. To this end, we add the matrix $P \in \mathbb{R}^{N\times d}$ to the input embedding. The embedded input can be compactly written into a matrix $X^{(0)} = [E(n_1) + P_1, ..., E(n_N) + P_N]^\top \in \mathbb{R}^{N\times d}$. Then, L Transformer blocks follow, each of which transforms the input via the following equation:

$$Y^{(l)} = X^{(l-1)} + \mathrm{Attn}^{(l)}(X^{(l-1)}), \tag{6}$$

$$X^{(l)} = Y^{(l)} + \mathrm{FFN}^{(l)}(Y^{(l)}), \tag{7}$$

where $l \in [L]$, $\mathrm{Attn}^{(l)}$ and $\mathrm{FFN}^{(l)}$ denote the multi-head self-attention layer and the feed-forward network for the l-th Transformer block, respectively:

$$\mathrm{Attn}^{(l)}(X) = \sum_{h=1}^{H} \mathrm{softmax}\left(XW_Q^{(l,h)}(XW_K^{(l,h)})^\top/\sqrt{d_H}\right) XW_V^{(l,h)}(W_O^{(l,h)})^\top, \tag{8}$$

$$\mathrm{FFN}^{(l)}(X) = \sigma(XW_F^{(l)}). \tag{9}$$

The variational parameters $W_Q^{(l,h)}, W_K^{(l,h)}, W_V^{(l,h)}, W_O^{(l,h)} \in \mathbb{R}^{d\times d_H}$ are the query, key, value, and output matrices of the head h, respectively. $d_H = d/H$ is the hidden dimension of each head. $W_F^{(l)} \in \mathbb{R}^{d\times d}$ is the weight matrix in the $\mathrm{FFN}^{(l)}$. The activation σ is chosen as SiLU [58].

The final output $X^{(L)}$ is transformered by a linear layer to produce K sets of backflow orbitals $M^k \in \mathbb{R}^{2N\times N_e}$, where $k \in [K]$. The Slater matrix Φ^k is formed by selecting rows of M^k corresponding to the occupied position in the input configuration $\mathbf{n}$. The overall NQS wavefunction is a linear combination of these determinants:

$$\langle \mathbf{n}|\psi\rangle = \sum_{k=1}^{K} \det[\Phi^k]. \tag{10}$$

4.3 Optimization

Stochastic reconfiguration (SR) is one of the most popular optimization methods for NQS. In this work, however, we propose an advanced SR algorithms for superior performance, named Moment-Adaptive ReConfiguration Heuristic (MARCH), most directly inspired by the SPRING algorithm [25]. MARCH combines ideas from the extremely successful Adam optimizer in traditional machine learning [45] and the recently introduced efficient Min-SR algorithm [23, 24].

The SR parameter update is typically defined as

$$\mathrm{d}\theta = \arg\min_{\mathrm{d}\theta'} \mathbb{E}_{\mathbf{n}}\left[|O(\mathbf{n})\mathrm{d}\theta' - \epsilon(\mathbf{n})|^2\right] \approx \arg\min_{\mathrm{d}\theta'} \frac{1}{B}\sum_{i=1}^{B} |O(\mathbf{n}_i)\mathrm{d}\theta' - \epsilon(\mathbf{n}_i)|^2, \tag{11}$$

where $O(\mathbf{n}) = \frac{\partial}{\partial\theta}\log|\psi_\theta(\mathbf{n})| - \mathbb{E}_{\mathbf{n}'}[\frac{\partial}{\partial\theta}\log|\psi_\theta(\mathbf{n}')|]$, $\epsilon(\mathbf{n}) = -\delta\tau(E_{\mathrm{loc}}(\mathbf{n}) - \mathbb{E}_{\mathbf{n}'}[E_{\mathrm{loc}}(\mathbf{n}')])$, B denotes batch size in MCMC sampling.

To express the problem in a standard linear algebra form, we construct a matrix $\tilde{O}$ and a vector $\tilde{\epsilon}$ by stacking the contributions from each of the B samples in the batch:

$$\tilde{O} = \begin{pmatrix} O(\mathbf{n}_1) \\ O(\mathbf{n}_2) \\ \vdots \\ O(\mathbf{n}_B) \end{pmatrix} \in \mathbb{R}^{B\times N_{\mathrm{params}}} \quad \text{and} \quad \tilde{\epsilon} = \begin{pmatrix} \epsilon(\mathbf{n}_1) \\ \epsilon(\mathbf{n}_2) \\ \vdots \\ \epsilon(\mathbf{n}_B) \end{pmatrix} \in \mathbb{R}^{B}. \tag{12}$$

Since $\tilde{O}$ might be ill-conditioned, the conventional method is to add some form of regularization. For example, a Tikhonov regularization is usually added to yield the regularized problem as follows:

$$\mathrm{d}\theta = \arg\min_{\mathrm{d}\theta'} \frac{1}{\lambda}||\tilde{O}\mathrm{d}\theta' - \tilde{\epsilon}||^2 + ||\mathrm{d}\theta'||^2. \tag{13}$$

MARCH differs from previous SR algorithms by adaptively translating and scaling $\mathrm{d}\theta$ individually for each parameter based on the estimation of the first and second moments.

The overall objective function can be written as:

$$\mathrm{d}\theta_k = \arg\min_{\mathrm{d}\theta'} \frac{1}{\lambda}||\tilde{O}\mathrm{d}\theta' - \tilde{\epsilon}||^2 + ||\mathrm{diag}(v_{k-1})^{1/4}(\mathrm{d}\theta' - \phi_{k-1})||^2, \tag{14}$$

where k denotes the current step, $\mathrm{diag}(v_{k-1})$ is the diagonal matrix whose main diagonal is composed of v_{k-1}. The first-order moment estimator, $\phi_k = \mu\mathrm{d}\theta_k$, acts as a momentum term. Intuitively, ϕ carries the momentum, which accelerates updates in directions of persistent gradients and dampens oscillations. The second moment estimator, $v_k = \beta v_{k-1} + (\mathrm{d}\theta_k - \mathrm{d}\theta_{k-1})^2$, tracks the volatility of the gradient. A large v_k indicates that the gradient is changing erratically, so the update of the parameter is suppressed to prevent divergence. Conversely, a small v_k implies a stable gradient, allowing the parameter to be updated more aggressively. The behavior of these estimators is governed by the hyperparameters μ, β, and λ, which are specified in Supplementary Table 7.

Therefore, the explicit update formula can be derived. We first define $\pi = \mathrm{diag}(v_{k-1})^{1/4}(\mathrm{d}\theta' - \phi_{k-1}), U = \tilde{O}\mathrm{diag}(v_{k-1})^{-1/4}, \zeta = \tilde{\epsilon} - \tilde{O}\phi_{k-1}$. The equation 14 can be recast as:

$$\pi = \arg\min_{\pi'} \frac{1}{\lambda}||U\pi' - \zeta||^2 + ||\pi'||^2. \tag{15}$$

We can reduce the computational budget from $\mathcal{O}(N_{\text{params}}^3)$ to $\mathcal{O}(B^2 N_{\text{params}})$ by Woodbury matrix identity:

$$\begin{aligned}
\pi &= (U^T U + \lambda I)^{-1} U^T \zeta && (16)\\
&= \frac{1}{\lambda}\left[I - U^T(\lambda I + UU^T)^{-1}U\right] U^T\zeta && (17)\\
&= \frac{1}{\lambda}\left[U^T - U^T(\lambda I + UU^T)^{-1}(UU^T + \lambda I - \lambda I)\right]\zeta && (18)\\
&= U^T(\lambda I + UU^T)^{-1}\zeta. && (19)
\end{aligned}$$

Finally, we can write down the parameters update formula as:

$$\mathrm{d}\theta = \mathrm{diag}(v_{k-1})^{-1/2}\tilde{O}^T(\tilde{O}\mathrm{diag}(v_{k-1})^{-1/2}\tilde{O}^T + \lambda I)^{-1}(\tilde{\epsilon} - \tilde{O}\phi_{k-1}) + \phi_{k-1}. \tag{20}$$

Acknowledgments. We thank ByteDance Seed for inspiration and encouragement, and Hang Li for his guidance and support. We thank Xiangjian Qian for providing

DMRG benchmark results. We are grateful to Yuxi Zhu for helping with graphics. D.H. is supported by National Science Foundation of China (NSFC62376007). Y.W. is supported by a start-up grant from IOP-CAS. T.X. is supported by National Natural Science Foundation of China (Grant No. 12488201). M.Q. acknowledges the support from the National Key Research and Development Program of MOST of China (2022YFA1405400), the National Natural Science Foundation of China (Grant No. 12274290) and the Innovation Program for Quantum Science and Technology (2021ZD0301902). L.W. is supported by National Science and Technology Major Project (2022ZD0114902) and National Science Foundation of China (NSFC92470123, NSFC62276005).

Author contributions. L.W. and D.L. conceived this project. Y.G. designed the algorithm, wrote the code and performed the calculation with helps from W.L., H.L., R.L and Y.H, all of whom contributed to preliminary explorations. Y.G. and W.L. visualized the data. B.Z., Y.W., M.Q. and T.X. provided valuable insight in the physics of the Hubbard model. Y.W. wrote the PEPS code for the benchmark with t' hopping. D.H. provided insights into neural network model design. All authors contributed to the analysis of the data, the discussion of the results and the writing of the manuscript.

Competing interests. There are no competing interests to declare.

Data and materials availability. All code and data will be made openly available upon publication.

References

[1] Hubbard, J.: Electron Correlations in Narrow Energy Bands. Proc. R. Soc. Lond. A **276**(1365), 238–257 (1963) https://doi.org/10.1098/rspa.1963.0204

[2] Qin, M., Schäfer, T., Andergassen, S., Corboz, P., Gull, E.: The hubbard model: A computational perspective. Annual Review of Condensed Matter Physics **13**(1), 275–302 (2022)

[3] Arovas, D.P., Berg, E., Kivelson, S.A., Raghu, S.: The hubbard model. Annual review of condensed matter physics **13**(1), 239–274 (2022)

[4] Anderson, P.W.: The resonating valence bond state in la_2cuo_4 and superconductivity. Science **235**(4793), 1196–1198 (1987) https://doi.org/10.1126/science.235.4793.1196 https://www.science.org/doi/pdf/10.1126/science.235.4793.1196

[5] Zhang, F.C., Rice, T.M.: Effective hamiltonian for the superconducting cu oxides. Phys. Rev. B **37**, 3759–3761 (1988) https://doi.org/10.1103/PhysRevB.37.3759

[6] White, S.R.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett. **69**, 2863–2866 (1992) https://doi.org/10.1103/PhysRevLett.69.2863

[7] Cirac, J.I., Pérez-García, D., Schuch, N., Verstraete, F.: Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. Rev. Mod. Phys. **93**, 045003 (2021) https://doi.org/10.1103/RevModPhys.93.045003

[8] Xiang, T.: Density Matrix and Tensor Network Renormalization. Cambridge University Press, Cambridge (2023)

[9] Verstraete, F., Cirac, J.I.: Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions (2004). https://arxiv.org/abs/cond-mat/0407066

[10] Georges, A., Kotliar, G., Krauth, W., Rozenberg, M.J.: Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. Rev. Mod. Phys. **68**, 13–125 (1996) https://doi.org/10.1103/RevModPhys.68.13

[11] Knizia, G., Chan, G.K.-L.: Density matrix embedding: A simple alternative to dynamical mean-field theory. Phys. Rev. Lett. **109**, 186404 (2012) https://doi.org/10.1103/PhysRevLett.109.186404

[12] Gubernatis, J., Kawashima, N., Werner, P.: Quantum Monte Carlo Methods: Algorithms for Lattice Models. Cambridge University Press, Cambridge (2016)

[13] Becca, F., Sorella, S.: Quantum Monte Carlo Approaches for Correlated Systems. Cambridge University Press, Cambridge (2017)

[14] Zhang, S. *Auxiliary-Field Quantum Monte Carlo for Correlated Electron Systems*, Vol. 3 of *Emergent Phenomena in Correlated Matter: Modeling and Simulation*, Ed. E. Pavarini, E. Koch, and U. Schollwöck (Verlag des Forschungszentrum Jülich, 2013)

[15] LeBlanc, J.P.F., Antipov, A.E., Becca, F., Bulik, I.W., Chan, G.K.-L., Chung, C.-M., Deng, Y., Ferrero, M., Henderson, T.M., Jiménez-Hoyos, C.A., Kozik, E., Liu, X.-W., Millis, A.J., Prokof'ev, N.V., Qin, M., Scuseria, G.E., Shi, H., Svistunov, B.V., Tocchio, L.F., Tupitsyn, I.S., White, S.R., Zhang, S., Zheng, B.-X., Zhu, Z., Gull, E.: Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms. Phys. Rev. X **5**, 041041 (2015) https://doi.org/10.1103/PhysRevX.5.041041

[16] Zheng, B.-X., Chung, C.-M., Corboz, P., Ehlers, G., Qin, M.-P., Noack, R.M., Shi, H., White, S.R., Zhang, S., Chan, G.K.-L.: Stripe order in the underdoped region of the two-dimensional Hubbard model. Science **358**(6367), 1155–1160 (2017) https://doi.org/10.1126/science.aam7127

[17] Xu, H., Chung, C.-M., Qin, M., Schollwöck, U., White, S.R., Zhang, S.: Coexistence of superconductivity with partially filled stripes in the hubbard model. Science **384**(6696), 7691 (2024)

[18] Nomura, Y., Darmawan, A.S., Yamaji, Y., Imada, M.: Restricted boltzmann machine learning for solving strongly correlated quantum systems. Physical Review B **96**(20), 205152 (2017)

[19] Inui, K., Kato, Y., Motome, Y.: Determinant-free fermionic wave function using feed-forward neural networks. Physical Review Research **3**(4), 043126 (2021)

[20] Luo, D., Clark, B.K.: Backflow transformations via neural networks for quantum many-body wave functions. Physical review letters **122**(22), 226401 (2019)

[21] Robledo Moreno, J., Carleo, G., Georges, A., Stokes, J.: Fermionic wave functions from neural-network constrained hidden states. Proceedings of the National Academy of Sciences **119**(32), 2122059119 (2022)

[22] Zhou, Y.-T., Zhou, Z.-W., Liang, X.: Solving fermi-hubbard-type models by tensor representations of backflow corrections. Physical Review B **109**(24), 245107 (2024)

[23] Chen, A., Heyl, M.: Empowering deep neural quantum states through efficient optimization. Nature Physics **20**(9), 1476–1481 (2024)

[24] Rende, R., Viteritti, L.L., Bardone, L., Becca, F., Goldt, S.: A simple linear algebra identity to optimize large-scale neural network quantum states. Communications Physics **7**(1), 260 (2024)

[25] Goldshlager, G., Abrahamsen, N., Lin, L.: A kaczmarz-inspired approach to accelerate the optimization of neural network wavefunctions. Journal of Computational Physics **516**, 113351 (2024)

[26] Carleo, G., Troyer, M.: Solving the quantum many-body problem with artificial neural networks. Science **355**(6325), 602–606 (2017)

[27] Pfau, D., Spencer, J.S., Matthews, A.G., Foulkes, W.M.C.: Ab initio solution of the many-electron schrödinger equation with deep neural networks. Physical review research **2**(3), 033429 (2020)

[28] Ren, W., Fu, W., Wu, X., Chen, J.: Towards the ground state of molecules via diffusion monte carlo on neural networks. Nature Communications **14**(1), 1860 (2023)

[29] Li, R., Ye, H., Jiang, D., Wen, X., Wang, C., Li, Z., Li, X., He, D., Chen, J., Ren, W., *et al.*: A computational framework for neural network-based variational monte carlo with forward laplacian. Nature Machine Intelligence **6**(2), 209–219 (2024)

[30] Nomura, Y., Imada, M.: Dirac-type nodal spin liquid revealed by refined quantum many-body solver using neural-network wave function, correlation ratio, and level

spectroscopy. Physical Review X **11**(3), 031034 (2021)

[31] Roth, C., Szabó, A., MacDonald, A.H.: High-accuracy variational monte carlo for frustrated magnets with deep neural networks. Physical Review B **108**(5), 054410 (2023)

[32] Yu, Q.-H., Lin, Z.-J.: Solving quantum many-particle models with graph attention network. Chinese Physics Letters **41**(3), 030202 (2024)

[33] Sorella, S.: Generalized lanczos algorithm for variational quantum monte carlo. Physical Review B **64**(2), 024512 (2001)

[34] Nightingale, M., Melik-Alaverdian, V.: Optimization of ground-and excited-state wave functions and van der waals clusters. Physical review letters **87**(4), 043401 (2001)

[35] Sorella, S., Casula, M., Rocca, D.: Weak binding between two aromatic rings: Feeling the van der waals attraction by quantum monte carlo methods. The Journal of chemical physics **127**(1) (2007)

[36] Karthik, V., Medhi, A.: Convolutional restricted boltzmann machine correlated variational wave function for the hubbard model on a square lattice. Physical Review B **110**(12), 125125 (2024)

[37] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

[38] Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., Amodei, D.: Scaling laws for neural language models. arXiv preprint arXiv:2001.08361 (2020)

[39] Tranquada, J., Sternlieb, B., Axe, J., Nakamura, Y., Uchida, S.-i.: Evidence for stripe correlations of spins and holes in copper oxide superconductors. nature **375**(6532), 561–563 (1995)

[40] Andersen, O., Liechtenstein, A., Jepsen, O., Paulsen, F.: Lda energy bands, low-energy hamiltonians, t', t'', $t_\perp(k)$, and $J_\perp$. Journal of Physics and Chemistry of Solids **56**(12), 1573–1591 (1995)

[41] Hirayama, M., Yamaji, Y., Misawa, T., Imada, M.: Ab initio effective hamiltonians for cuprate superconductors. Physical Review B **98**(13), 134501 (2018)

[42] Liu, W.-Y., Zhai, H., Peng, R., Gu, Z.-C., Chan, G.K.-L.: Accurate simulation of the hubbard model with finite fermionic projected entangled pair states. Phys. Rev. Lett. **134**, 256502 (2025) https://doi.org/10.1103/r4q9-4yvj

[43] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P.,

Neelakantan, A., Shyam, P., Sastry, G., Askell, A., *et al.*: Language models are few-shot learners. Advances in neural information processing systems **33**, 1877–1901 (2020)

[44] Cheng, R.: Quantum geometric tensor (fubini-study metric) in simple quantum system: A pedagogical introduction. arXiv preprint arXiv:1012.1337 (2010)

[45] Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization (2017). https://arxiv.org/abs/1412.6980

[46] Sutskever, I., Martens, J., Dahl, G., Hinton, G.: On the importance of initialization and momentum in deep learning. In: International Conference on Machine Learning, pp. 1139–1147 (2013). PMLR

[47] Qin, M., Shi, H., Zhang, S.: Benchmark study of the two-dimensional hubbard model with auxiliary-field quantum monte carlo method. Physical Review B **94**(8), 085103 (2016)

[48] Viteritti, L.L., Rende, R., Parola, A., Goldt, S., Becca, F.: Transformer wave function for two dimensional frustrated magnets: Emergence of a spin-liquid phase in the shastry-sutherland model. Physical Review B **111**(13), 134411 (2025)

[49] Qin, M., Chung, C.-M., Shi, H., Vitali, E., Hubig, C., Schollwöck, U., White, S.R., Zhang, S., Many-Electron Problem), S.C.: Absence of superconductivity in the pure two-dimensional hubbard model. Physical Review X **10**(3), 031016 (2020)

[50] Hirayama, M., Yamaji, Y., Misawa, T., Imada, M.: Ab initio effective hamiltonians for cuprate superconductors. Phys. Rev. B **98**, 134501 (2018) https://doi.org/10.1103/PhysRevB.98.134501

[51] Chung, C.-M., Qin, M., Zhang, S., Schollwöck, U., White, S.R.: Plaquette versus ordinary d-wave pairing in the $t^{'}$-hubbard model on a width-4 cylinder. Phys. Rev. B **102**, 041106 (2020) https://doi.org/10.1103/PhysRevB.102.041106

[52] White, S.R., Scalapino, D.J.: Density matrix renormalization group study of the striped phase in the 2d $t-J$ model. Phys. Rev. Lett. **80**, 1272–1275 (1998) https://doi.org/10.1103/PhysRevLett.80.1272

[53] Jiang, Y.-F., Zaanen, J., Devereaux, T.P., Jiang, H.-C.: Ground state phase diagram of the doped hubbard model on the four-leg cylinder. Phys. Rev. Res. **2**, 033073 (2020) https://doi.org/10.1103/PhysRevResearch.2.033073

[54] Huang, E.W., Mendl, C.B., Jiang, H.-C., Moritz, B., Devereaux, T.P.: Stripe order from the perspective of the Hubbard model. npj Quantum Materials **3**(1), 22 (2018) https://doi.org/10.1038/s41535-018-0097-0 arXiv:1709.02398 [cond-mat.str-el]

[55] Jiang, H.-C., Devereaux, T.P.: Superconductivity in the doped Hubbard model and its interplay with next-nearest hopping t'. Science **365**(6460), 1424–1428 (2019) https://doi.org/10.1126/science.aal5304

[56] Lu, X., Chen, F., Zhu, W., Sheng, D.N., Gong, S.-S.: Emergent superconductivity and competing charge orders in hole-doped square-lattice $t-j$ model. Phys. Rev. Lett. **132**, 066002 (2024) https://doi.org/10.1103/PhysRevLett.132.066002

[57] Šimkovic, F., Rossi, R., Georges, A., Ferrero, M.: Origin and fate of the pseudogap in the doped hubbard model. Science **385**(6715), 9194 (2024) https://doi.org/10.1126/science.ade9194 https://www.science.org/doi/pdf/10.1126/science.ade9194

[58] Elfwing, S., Uchibe, E., Doya, K.: Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural networks **107**, 3–11 (2018)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- npSI.pdf