Scalable and accurate simulations of the Hubbard model with neural quantum states

Ao Chen
Division of Chemistry and Chemical Engineering
California Institute of Technology, Pasadena, USA
Center for Computational Quantum Physics
Flatiron Institute, New York, USA
chenao.phys@gmail.com

Christopher Roth
Center for Computational Quantum Physics
Flatiron Institute, New York, USA
croth@flatironinstitute.org

Zhou-Quan Wan
Center for Computational Quantum Physics
Flatiron Institute, New York, USA
zwan@flatironinstitute.org

Anirvan Sengupta
Center for Computational Quantum Physics
& Center for Computational Mathematics
Flatiron Institute, New York, USA
Department of Physics and Astronomy
Rutgers University, Piscataway, USA
asengupta@flatironinstitute.org

Antoine Georges
Collège de France, Paris, France
Center for Computational Quantum Physics
Flatiron Institute, New York, USA
CPHT, CNRS, École Polytechnique, Paris, France
DQMP, Université de Genève, Genève, Suisse
ageorges@flatironinstitute.org

Abstract

Neural quantum states (NQS) recently had a major impact on quantum many-body physics by suggesting novel wave-functions parametrized by neural networks. Nevertheless, NQS faces great challenges in fermionic systems, especially the Fermi-Hubbard model (FHM), a minimal model for unconventional superconductivity (SC). FHM poses three major numerical difficulties: large system sizes, correlated ground states, and competing low-energy states. In this work, we introduce the hidden fermion Pfaffian state (HFPS), which integrates ANN with Pfaffian that naturally encodes SC pairings. HFPS alleviates the three problems and provides state-of-the-art accuracy in FHM. Our results highlight the effectiveness of NQS and point toward promising future directions for resolving the longstanding challenge of unconventional SC in FHM.

1 Introduction

An outstanding challenge in many-body physics is to identify exotic phases of matter, for instance unconventional superconductivity (SC) [1, 2], from microscopic Hamiltonians. The Fermi-Hubbard

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Machine Learning and the Physical Sciences (ML4PS) Workshop at NeurIPS 2025.

model (FHM) is one of the most important models for this purpose, whose Hamiltonian is

$$\hat{\mathcal{H}} = -t \sum_{\langle ij \rangle, \sigma} \hat{c}^{\dagger}_{i,\sigma} \hat{c}_{j,\sigma} + U \sum_{i} \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow}, \tag{1}$$

where $\langle ij \rangle$ represents nearest neighbors, and t and U are respectively hopping and interacting strength. FHM is notoriously difficult to solve numerically, and the existence of unconventional SC in FHM remains controversial, despite numerous computational efforts [3, 4, 5, 6, 7, 8, 9].

An increasingly more popular approach is the variational Monte Carlo (VMC) [10]. It searches for the ground state of $\hat{\mathcal{H}}$ by utilizing a variational state $|\psi_\theta\rangle$ with parameters θ and minimizing the variational energy $E_\theta = \langle \psi_\theta | \hat{\mathcal{H}} | \psi_\theta \rangle \,/\, \langle \psi_\theta | \psi_\theta \rangle$. From a machine learning point of view, E_θ is a loss function to be minimized without a reference dataset, analogous to reinforcement learning. In VMC, one employs Monte Carlo samples to estimate E_θ, and uses gradient descent and its more accurate variants to minimize E_θ. Nevertheless, VMC still faces great challenges as listed below. These challenges call for variational states that are both scalable and expressive.

- **Large system sizes:** Stripe patterns in FHM induce strong finite-size effects. Accurate simulations therefore require sufficiently large system sizes to resolve these microscopic structures.
- **Correlated ground state:** Unconventional SC emerges in the strongly correlated regime characterized by large U and finite doping. In this regime, electron interactions are too strong to be treated as simple perturbative corrections to non-interacting fermions.
- **Competing states:** The SC ground state competes with numerous normal states of nearly degenerate energies. Convergence to competing states leads to solutions that fail to capture SC.

Recently, a novel variational wave-function, namely neural quantum states (NQS), was introduced based on artificial neural networks (ANN) [11], which has become a mainstream method for spin systems [12, 13, 14, 15]. Nevertheless, the development of NQS for fermionic systems has progressed somewhat more slowly, especially on FHM, due to the three challenges listed above [16, 17, 18].

To construct an NQS that accurately captures possible SC in FHM, we combine ANNs with the Pfaffian wave-function that expresses SC pairings at the mean-field level, which we call hidden fermion Pfaffian states (HFPS) [19]. HFPS provides an architecture compatible with deep ANNs and large-scale VMC simulations. In our numerical experiments, HFPS outperforms existing variational methods and reaches state-of-the-art accuracy. Furthermore, suitable pairing fields can be utilized to initialize HFPS to avoid competing low-energy states.

2 Hidden fermion Pfaffian state

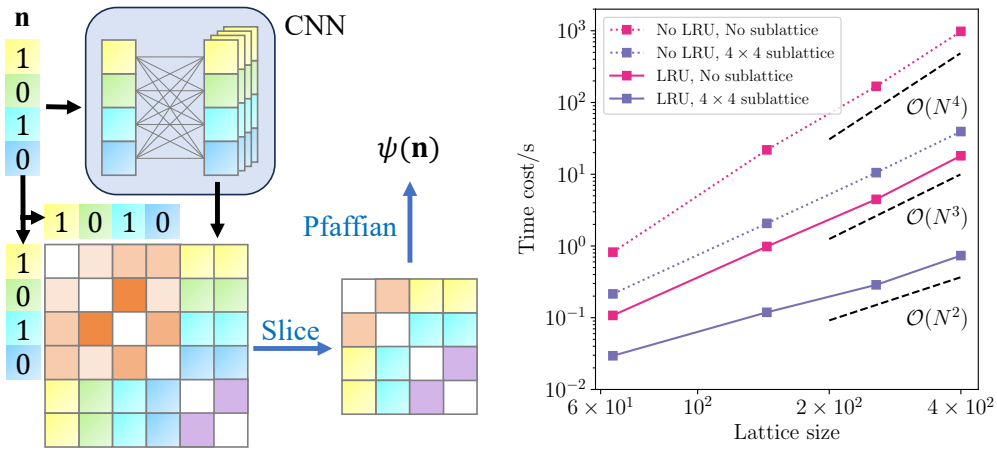
2.1 ANN-augmented Pfaffian

The Pfaffian wave-function is the solution of a general bilinear Hamiltonian. The Slater determinant state and the BCS state [20], for instance, are both subsets of the Pfaffian state. The Pfaffian wave-function with a conserved number of particles can be written as $\psi_{\mathrm{pf}}(\mathbf{n}) = \mathrm{pf}(\mathbf{n} \star \mathbf{F} \star \mathbf{n})$, where $\mathbf{F}$ is a skew-symmetric matrix parametrizing ψ_{pf}, and $\star$ denotes slicing the full matrix $\mathbf{F}$ according to the occupation number $\mathbf{n}$, along both rows ($\mathbf{n} \star \mathbf{F}$) and columns ($\mathbf{F} \star \mathbf{n}$). In VMC, ψ_{pf} has been a popular variational wave-function [21, 22].

In Ref.[18], it is shown that correlations between particles can be encoded by 'hidden' degrees of freedom on an enlarged Hilbert space. Here, we generalize this concept to Pfaffian wave-functions. We expand the Hilbert space with M orbitals and N visible fermions to include additional $\tilde{M}$ hidden orbitals and $\tilde{N}$ hidden fermions. To retain the wave-function $\psi(\mathbf{n})$ in the original Hilbert space, we assume that the hidden occupation number $\tilde{\mathbf{n}}$ depends on the visible occupation number $\mathbf{n}$, which is implicitly encoded by ANNs. The wave-function, similar to ψ_{pf}, is given by

$$\psi(\mathbf{n}) = \mathrm{pf}\left[(\mathbf{n}, \tilde{\mathbf{n}}) \star \begin{pmatrix} \mathbf{F}^{vv} & \mathbf{F}^{vh} \\ -(\mathbf{F}^{vh})^T & \mathbf{F}^{hh} \end{pmatrix} \star (\mathbf{n}, \tilde{\mathbf{n}}) \right] = \mathrm{pf} \begin{pmatrix} \mathbf{n} \star \mathbf{F}^{vv} \star \mathbf{n} & \mathbf{n} \star \tilde{\mathbf{F}}^{vh}(\mathbf{n}) \\ -\tilde{\mathbf{F}}^{vh}(\mathbf{n})^T \star \mathbf{n} & \tilde{\mathbf{F}}^{hh} \end{pmatrix}, \tag{2}$$

where $\tilde{\mathbf{F}}^{vh}(\mathbf{n}) = \mathbf{F}^{vh} \star \tilde{\mathbf{n}}$ is an $M \times \tilde{N}$ matrix parametrized by ANNs, and $\tilde{\mathbf{F}}^{hh} = \tilde{\mathbf{n}} \star \mathbf{F}^{hh} \star \tilde{\mathbf{n}}$ is an $\tilde{N} \times \tilde{N}$ matrix. The variational parameters of the full wave-function include directly parametrized $\mathbf{F}^{vv}$ and $\tilde{\mathbf{F}}^{hh}$, and ANN parameters θ. An illustration of HFPS architecture is presented in Fig. 1(a).



(a) HFPS architecture (b) Time cost

Figure 1: Illustration of HFPS architecture in Eq. (2) and its time scaling. (a) A spinless fermion example with four orbitals, $M = 4$, two visible fermions, $N = 2$, and two hidden fermions, $\tilde{N} = 2$. (b) Time cost of HFPS forward pass with and without the acceleration given by low-rank updates (LRU) and sublattices. The test is performed on an A6000 GPU with a batch size of 1000.

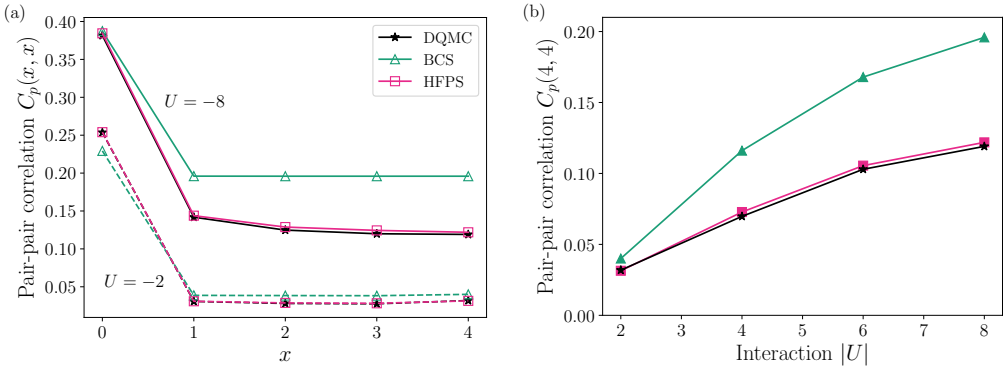


Figure 2: Pair-pair correlation $C_p(\mathbf{r})$ in the 8×8 attractive FHM.

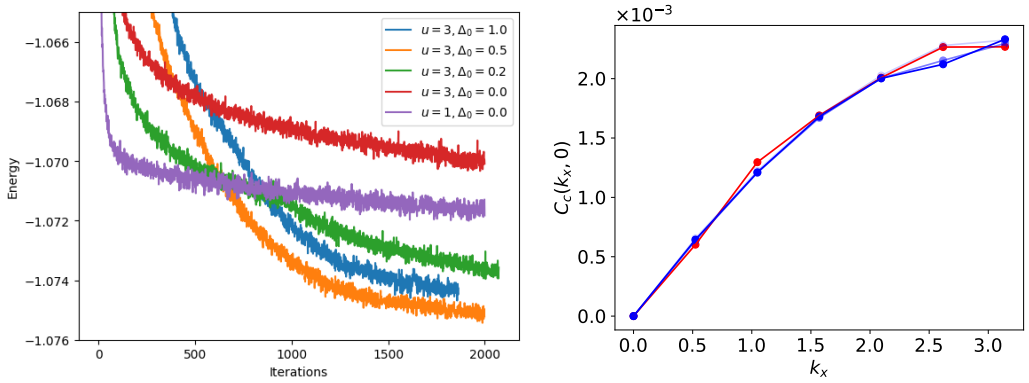
2.2 Time scaling

The overall complexity of translation symmetrized fermionic NQS is $\mathcal{O}(N^4)$, in which $\mathcal{O}(N^3)$ comes from Pfaffian (or determinant in other works) and $\mathcal{O}(N)$ comes from translation symmetry. To solve the challenge of large system sizes in FHM, we utilize two useful techniques in HFPS, namely low-rank updates (LRU) [23] and sublattices [22], to lower the cost to $\mathcal{O}(N^2)$ given suitable initialization. The details of these techniques are discussed in Ref. [19] and implemented by `lrux` [24] and `quantax` [25]. In Fig. 1(b), we show the growth of time cost over increasing system sizes. The forward pass complexity is reduced from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^2)$ as expected. For the largest 20×20 lattice, the forward pass is accelerated by a factor of 10^3. Therefore, we expect HFPS to be an ideal NQS architecture for simulating FHM.

3 Numerical experiments

3.1 Superconductivity with attractive interactions

When the on-site interaction of FHM is attractive, the electrons form Cooper pairs and exhibit conventional SC, as described by the BCS mean-field theory [20]. To quantify SC in a system



(a) Training curves

(b) Converged charge structure factors

Figure 3: Competing states distinguished by different initializations.

with a conserved particle number, we define the pair-pair correlation $C_p(\mathbf{r}) = \langle \hat{\Delta}^\dagger(\mathbf{r})\hat{\Delta}(\mathbf{0})\rangle$ with $\hat{\Delta}(\mathbf{r}) = \hat{c}_{\mathbf{r},\uparrow}\hat{c}_{\mathbf{r},\downarrow}$ to observe the off-diagonal long-range order that reflects SC. In Fig. 2, we show $C_p(\mathbf{r})$ given by several methods in the 8×8 attractive FHM at 1/8 doping (particle density $n = 7/8$) and periodic boundary condition (PBC). The BCS theory is a qualitatively correct description and provides good accuracy under weak interaction. However, BCS theory tends to overestimate long-range correlations under strong interaction because of the mean-field approximation. Consequently, it is not a quantitatively accurate description for strongly interacting electrons. HFPS, on the other hand, is trained to encode correlations through an ANN and achieves great accuracy in predicting $C_p(\mathbf{r})$, within the error bar of the exact result given by determinant quantum Monte Carlo (DQMC) across all interaction strengths. Therefore, HFPS provides a reliable approach in studying the behavior of systems with SC and strong correlations.

3.2 Initialization with pairing field

In order to test our method in the regime with strong correlations and possible unconventional SC, we study FHM with hole doping and strong repulsive interactions. In this regime, FHM develops many competing states with nearly degenerate energies, where proper initialization is key to avoiding convergence to competing states. For FHM with $U \geq 0$, the optimal Pfaffian state with the lowest energy is equivalent to a Slater determinant state without pairing. Nevertheless, we find that we can sometimes improve our final energies by initializing the pure Pfaffian part $\mathbf{F}^{vv}$ of Eq. (2) with an additional d-wave pairing field, i.e.

$$\hat{\mathcal{H}}' = \hat{\mathcal{H}} + \Delta_0 \sum_{\langle ij \rangle} \mathrm{sign}(i,j)(\hat{c}_{i,\uparrow}\hat{c}_{j,\downarrow} - \hat{c}_{i,\downarrow}\hat{c}_{j,\uparrow} + h.c.), \tag{3}$$

where $\hat{\mathcal{H}}$ is the original Hubbard Hamiltonian in Eq. (1) with a weaker interaction, Δ_0 is the strength of the pairing field, and $\mathrm{sign}(i,j)$ denotes the sign for d-wave pairing. A finite Δ_0 provides pairings in the initialized wave-function, which is crucial for avoiding convergence to competing normal states. After training $\mathbf{F}^{vv}$, Δ_0 is set back to 0 to allow NQS to converge to the true ground state.

In Fig. 3(a), we show the training curves of HFPS at $1/6$ doping and $U = 4$ on 12×12 FHM, given different initializations. Although states initialized with finite Δ_0 have higher energies in the beginning, they converge to states with lower energies and stronger pairings. To observe the structure of these states, we measure the charge structure factor $C_c(\mathbf{k}) = \sum_{\mathbf{r}} C_c(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}$, where $C_c(\mathbf{r}) = \langle \hat{n}_{\mathbf{r}} \cdot \hat{n}_{\mathbf{0}}\rangle - \langle \hat{n}_{\mathbf{r}}\rangle \cdot \langle \hat{n}_{\mathbf{0}}\rangle$. The charge density wave (CDW) of the normal state generates a peak of $C_c(\mathbf{k})$ at $\mathbf{k} = (\pi/3, 0)$, as shown by the red curve in Fig. 3(b). On the other hand, the states initialized by a finite Δ_0 have weakened $C_c(\mathbf{k})$ at $\mathbf{k} = (\pi/3, 0)$, revealing suppressed CDW. These results indicate that suitable initialization of the Pfaffian is the key to resolving competing states.

We benchmark our approach on the 16×4 FHM at $1/8$ filling and $U = 8$ with PBC. In this system, HFPS and the recent Transformer backflow [26] greatly outperform earlier variational results. As presented in Table 1, HFPS outperforms transformer backflow after imposing pairing initialization

Method	Initialization	Final energy E/N
Transformer backflow	Stripe state	-0.76298
HFPS	Stripe state	-0.76256
HFPS	Stripe state with pairing	$\mathbf{-0.76413}$

Table 1: Variational energy in the 16×4 FHM with $U = 8$ and $1/8$ doping.

with $\Delta_0 = 0.2$. Although this system is believed to have no long-range pair-pair correlation, hence no SC, it still benefits from the short-range correlations imposed during initialization.

4 Discussion

In this paper, we discuss three major challenges – large system sizes, correlated ground states, and competing low-energy states – that NQS faces in simulating FHM. To overcome these problems, we propose HFPS that combines NQS with Pfaffians. HFPS has a favorable $\mathcal{O}(N^3)$ scaling as discussed in 2.2, alleviating the problem of large system sizes in FHM. In 3.1, we further show that HFPS can efficiently capture the strong correlation of the SC ground state through its deep ANN. Finally, HFPS successfully distinguishes the ground state and low-lying competing states by comparing converged variational energies of different initial biases, as presented in 3.2.

The development of HFPS advances fermionic NQS from proof-of-principle studies on small systems to large-scale simulations at the frontier of correlated physics. We expect HFPS to provide new insights into challenging fermionic problems, particularly the possible unconventional SC in FHM.

References

[1] J. G. Bednorz and K. A. Müller. Possible hightc superconductivity in the ba-la-cu-o system. *Zeitschrift für Physik B Condensed Matter*, 64(2):189–193, Jun 1986. ISSN 1431-584X. doi: 10.1007/BF01303701. URL https://doi.org/10.1007/BF01303701.

[2] M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Q. Wang, and C. W. Chu. Superconductivity at 93 k in a new mixed-phase y-ba-cu-o compound system at ambient pressure. *Phys. Rev. Lett.*, 58:908–910, Mar 1987. doi: 10.1103/PhysRevLett.58.908. URL https://link.aps.org/doi/10.1103/PhysRevLett.58.908.

[3] Bo-Xiao Zheng, Chia-Min Chung, Philippe Corboz, Georg Ehlers, Ming-Pu Qin, Reinhard M. Noack, Hao Shi, Steven R. White, Shiwei Zhang, and Garnet Kin-Lic Chan. Stripe order in the underdoped region of the two-dimensional hubbard model. *Science*, 358(6367):1155–1160, 2017. doi: 10.1126/science.aam7127. URL https://www.science.org/doi/abs/10.1126/science.aam7127.

[4] Andrew S. Darmawan, Yusuke Nomura, Youhei Yamaji, and Masatoshi Imada. Stripe and superconducting order competing in the hubbard model on a square lattice studied by a combined variational monte carlo and tensor network method. *Phys. Rev. B*, 98:205132, Nov 2018. doi: 10.1103/PhysRevB.98.205132. URL https://link.aps.org/doi/10.1103/PhysRevB.98.205132.

[5] Mingpu Qin, Chia-Min Chung, Hao Shi, Ettore Vitali, Claudius Hubig, Ulrich Schollwöck, Steven R. White, and Shiwei Zhang. Absence of superconductivity in the pure two-dimensional hubbard model. *Phys. Rev. X*, 10:031016, Jul 2020. doi: 10.1103/PhysRevX.10.031016. URL https://link.aps.org/doi/10.1103/PhysRevX.10.031016.

[6] Hao Xu, Hao Shi, Ettore Vitali, Mingpu Qin, and Shiwei Zhang. Stripes and spin-density waves in the doped two-dimensional hubbard model: Ground state phase diagram. *Phys. Rev. Res.*, 4:013239, Mar 2022. doi: 10.1103/PhysRevResearch.4.013239. URL https://link.aps.org/doi/10.1103/PhysRevResearch.4.013239.

[7] Alexander Wietek. Fragmented cooper pair condensation in striped superconductors. *Phys. Rev. Lett.*, 129:177001, Oct 2022. doi: 10.1103/PhysRevLett.129.177001. URL https://link.aps.org/doi/10.1103/PhysRevLett.129.177001.

[8] Sandro Sorella. Systematically improvable mean-field variational ansatz for strongly correlated systems: Application to the hubbard model. *Physical Review B*, 107(11):115133, 2023.

[9] Hao Xu, Chia-Min Chung, Mingpu Qin, Ulrich Schollwöck, Steven R. White, and Shiwei Zhang. Coexistence of superconductivity with partially filled stripes in the hubbard model. *Science*, 384 (6696):eadh7691, 2024. doi: 10.1126/science.adh7691. URL https://www.science.org/doi/abs/10.1126/science.adh7691.

[10] Federico Becca and Sandro Sorella. *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017. doi: 10.1017/9781316417041.

[11] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017. ISSN 0036-8075. doi: 10.1126/science.aag2302. URL https://science.sciencemag.org/content/355/6325/602.

[12] Yusuke Nomura and Masatoshi Imada. Dirac-type nodal spin liquid revealed by refined quantum many-body solver using neural-network wave function, correlation ratio, and level spectroscopy. *Phys. Rev. X*, 11:031034, Aug 2021. doi: 10.1103/PhysRevX.11.031034. URL https://link.aps.org/doi/10.1103/PhysRevX.11.031034.

[13] Christopher Roth, Attila Szabó, and Allan H. MacDonald. High-accuracy variational monte carlo for frustrated magnets with deep neural networks. *Phys. Rev. B*, 108:054410, Aug 2023. doi: 10.1103/PhysRevB.108.054410. URL https://link.aps.org/doi/10.1103/PhysRevB.108.054410.

[14] Ao Chen and Markus Heyl. Empowering deep neural quantum states through efficient optimization. *Nature Physics*, 20(9):1476–1481, Sep 2024. ISSN 1745-2481. doi: 10.1038/s41567-024-02566-1. URL https://doi.org/10.1038/s41567-024-02566-1.

[15] Luciano Loris Viteritti, Riccardo Rende, Alberto Parola, Sebastian Goldt, and Federico Becca. Transformer wave function for two dimensional frustrated magnets: Emergence of a spin-liquid phase in the shastry-sutherland model. *Phys. Rev. B*, 111:134411, Apr 2025. doi: 10.1103/PhysRevB.111.134411. URL https://link.aps.org/doi/10.1103/PhysRevB.111.134411.

[16] Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada. Restricted boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B*, 96:205152, Nov 2017. doi: 10.1103/PhysRevB.96.205152. URL https://link.aps.org/doi/10.1103/PhysRevB.96.205152.

[17] Di Luo and Bryan K. Clark. Backflow transformations via neural networks for quantum many-body wave functions. *Phys. Rev. Lett.*, 122:226401, Jun 2019. doi: 10.1103/PhysRevLett.122.226401. URL https://link.aps.org/doi/10.1103/PhysRevLett.122.226401.

[18] Javier Robledo Moreno, Giuseppe Carleo, Antoine Georges, and James Stokes. Fermionic wave functions from neural-network constrained hidden states. *Proceedings of the National Academy of Sciences*, 119(32):e2122059119, 2022. doi: 10.1073/pnas.2122059119. URL https://www.pnas.org/doi/abs/10.1073/pnas.2122059119.

[19] Ao Chen, Zhou-Quan Wan, Anirvan Sengupta, Antoine Georges, and Christopher Roth. Neural network-augmented pfaffian wave-functions for scalable simulations of interacting fermions, 2025. URL https://arxiv.org/abs/2507.10705.

[20] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Microscopic theory of superconductivity. *Phys. Rev.*, 106:162–164, Apr 1957. doi: 10.1103/PhysRev.106.162. URL https://link.aps.org/doi/10.1103/PhysRev.106.162.

[21] Daisuke Tahara and Masatoshi Imada. Variational monte carlo method combined with quantum-number projection and multi-variable optimization. *Journal of the Physical Society of Japan*, 77 (11):114701, 2008. doi: 10.1143/JPSJ.77.114701. URL https://doi.org/10.1143/JPSJ.77.114701.

[22] Takahiro Misawa, Satoshi Morita, Kazuyoshi Yoshimi, Mitsuaki Kawamura, Yuichi Motoyama, Kota Ido, Takahiro Ohgoe, Masatoshi Imada, and Takeo Kato. mvmc—open-source software for many-variable variational monte carlo method. *Computer Physics Communications*, 235: 447–462, 2019. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2018.08.014. URL https://www.sciencedirect.com/science/article/pii/S0010465518303102.

[23] RuQing G. Xu, Tsuyoshi Okubo, Synge Todo, and Masatoshi Imada. Optimized implementation for calculation and fast-update of pfaffians installed to the open-source fermionic variational solver mvmc. *Computer Physics Communications*, 277:108375, 2022. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2022.108375. URL https://www.sciencedirect.com/science/article/pii/S0010465522000947.

[24] Ao Chen and Christopher Roth. Fast low-rank update of matrix determinants and pfaffians in JAX, 2025. URL https://github.com/ChenAo-Phys/lrux.

[25] Ao Chen and Christopher Roth. Quantax: Flexible neural quantum states based on QuSpin, JAX, and Equinox, 2025. URL https://github.com/ChenAo-Phys/quantax.

[26] Yuntian Gu, Wenrui Li, Heng Lin, Bo Zhan, Ruichen Li, Yifei Huang, Di He, Yantao Wu, Tao Xiang, Mingpu Qin, Liwei Wang, and Dingshun Lv. Solving the hubbard model with neural quantum states, 2025. URL https://arxiv.org/abs/2507.02644.