Recent Progress in Neural-Network-based Quantum Monte Carlo

Weiluo Ren

ByteDance Research

renweiluo@bytedance.com

@International Workshop on Artificial Intelligence for Theoretical Sciences

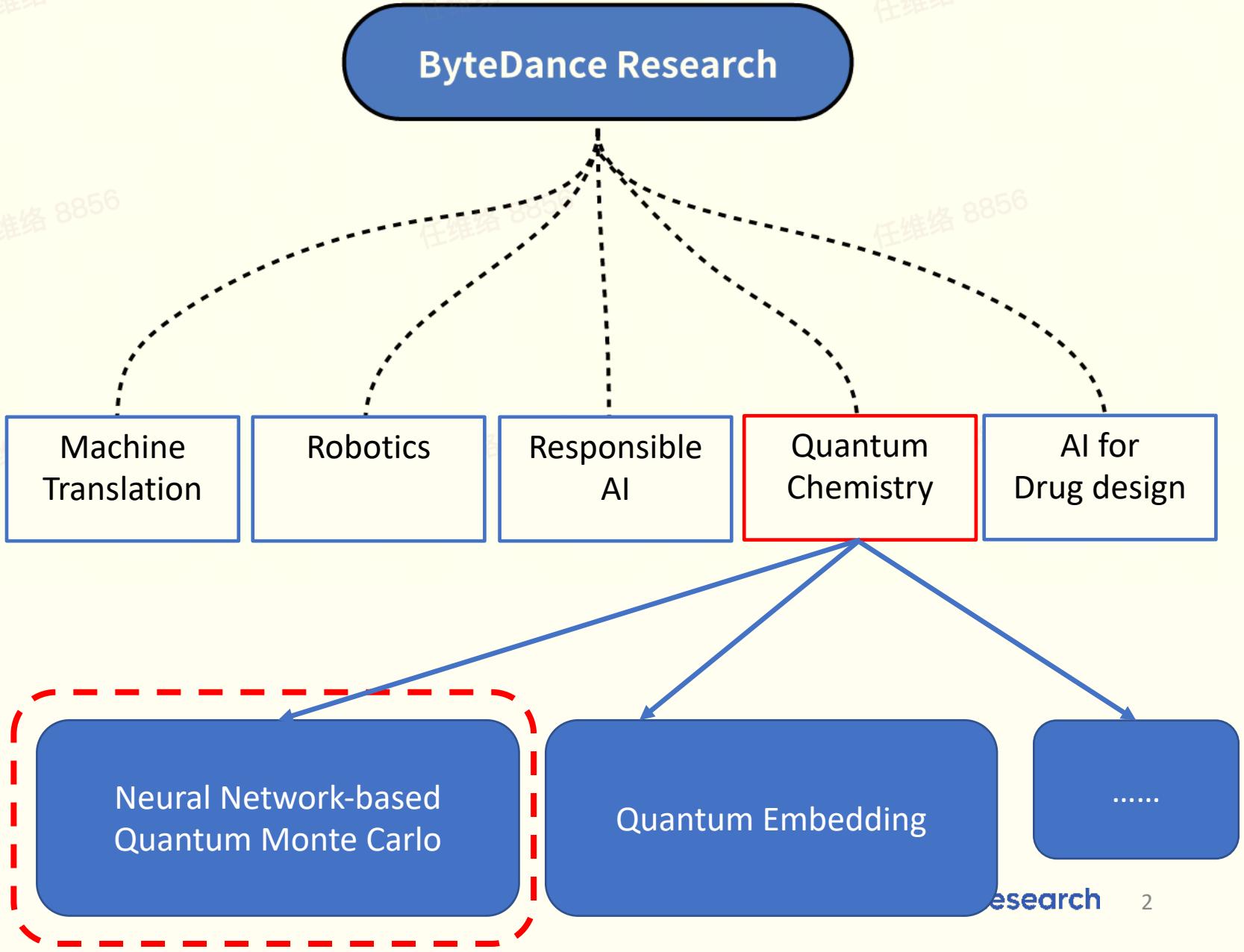
ByteDance 字节跳动

抖音 抖音火山版 飞书 今日头条

西瓜视频 番茄小说 住小帮 轻颜相机

剪映 火山引擎 懂车帝 幸福里

And TikTok !



Joint work with

Prof. Ji Chen

School of Physics

Prof. Liwei Wang

Prof. Di He

Department of Machine Intelligence

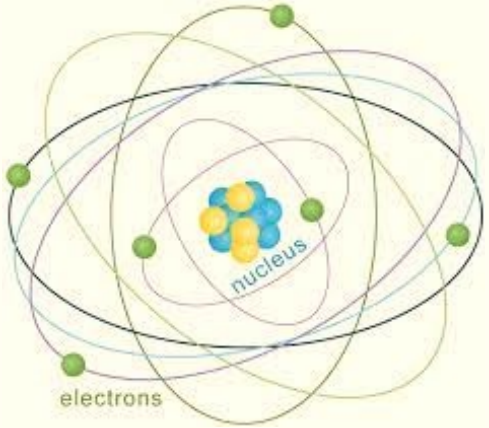
Outline

Introduction to NNQMC

Excited States with NNQMC

Novel Phase of Moire systems

Electronic Structure



Schrödinger Equation

$$\hat{H}\psi = E\psi$$

Electronic part

Nucleus-Electron part

$$\hat{H} = \boxed{-\frac{1}{2}\sum_i \nabla_i^2} + \boxed{\sum_{ij}\frac{1}{r_{ij}}} - \boxed{\sum_{ai}\frac{Z_a}{r_{ai}}}$$

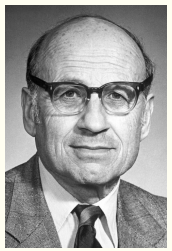
Kinetic energy of electrons

Repulsion of electrons

Attraction between nuclei and electrons

(Born–Oppenheimer Approximation)

Ab initio methods in Quantum Chemistry



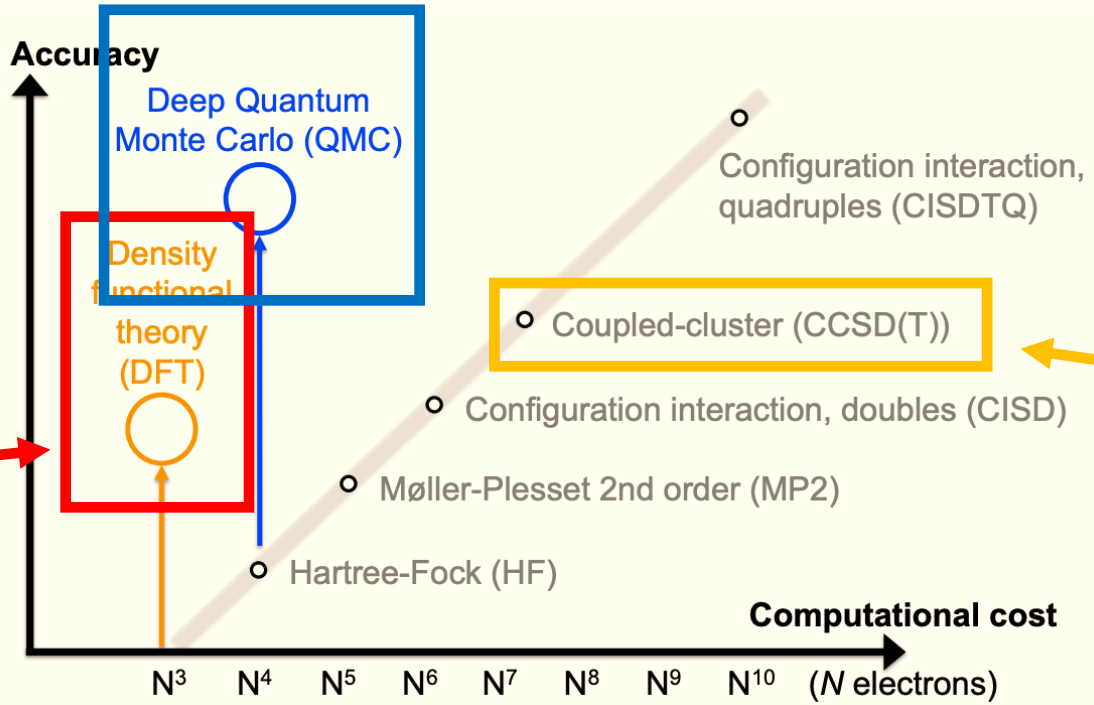
Density Functional Theory (DFT)

is the Go-To method

for Quantum Chemistry

NNQMC

New player on the field



H J Kulik et al 2022 Electron. Struct. 4 023004

We will focus on
Variational Monte Carlo today

CCSD(T)

is Golden Standard method

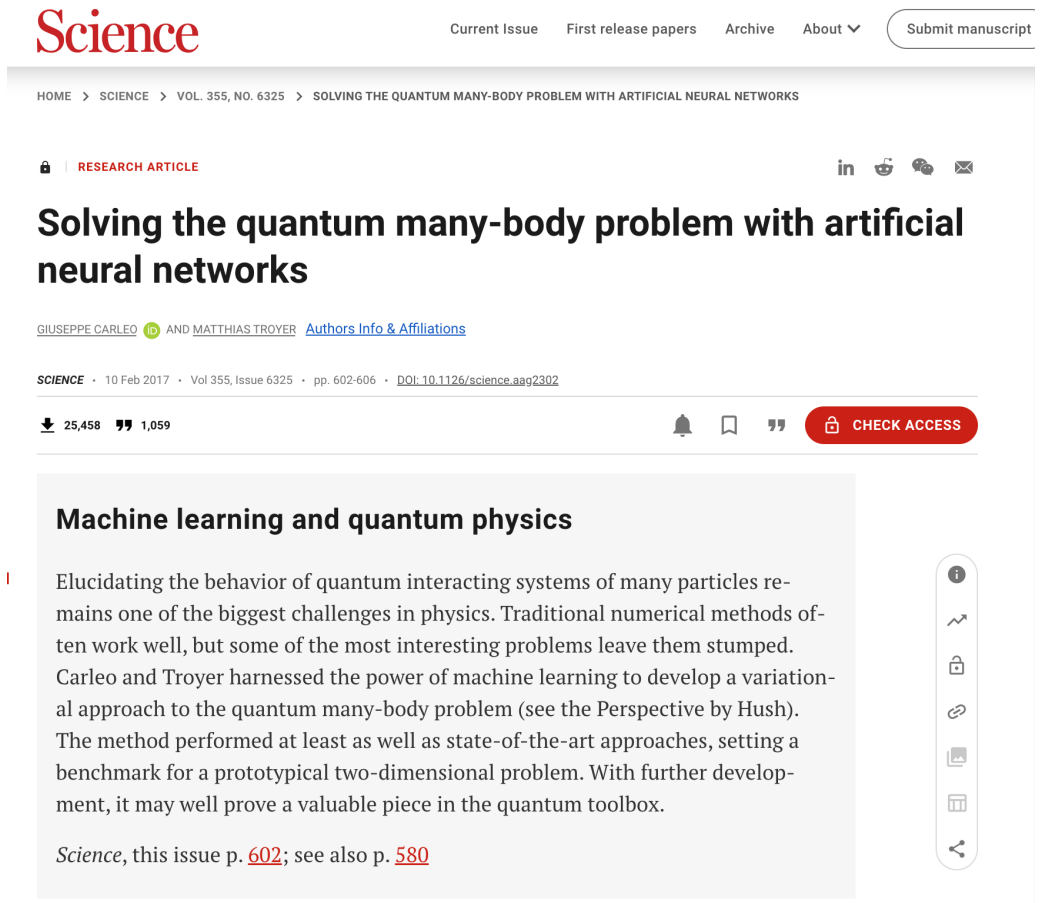
In Quantum Chemistry,

accurate but expensive.

Neural Quantum State (NQS)

Neural Network + QMC but for lattice models and/or fermionic systems with second quantization.

We will focus on the real-space QMC in this talk



Science

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 355, NO. 6325 > SOLVING THE QUANTUM MANY-BODY PROBLEM WITH ARTIFICIAL NEURAL NETWORKS

RESEARCH ARTICLE

Solving the quantum many-body problem with artificial neural networks

GIUSEPPE CARLEO AND MATTHIAS TROYER Authors Info & Affiliations

SCIENCE • 10 Feb 2017 • Vol 355, Issue 6325 • pp. 602-606 • DOI: 10.1126/science.aag2302

25,458 1,059 CHECK ACCESS

Machine learning and quantum physics

Elucidating the behavior of quantum interacting systems of many particles remains one of the biggest challenges in physics. Traditional numerical methods often work well, but some of the most interesting problems leave them stumped. Carleo and Troyer harnessed the power of machine learning to develop a variational approach to the quantum many-body problem (see the Perspective by Hush). The method performed at least as well as state-of-the-art approaches, setting a benchmark for a prototypical two-dimensional problem. With further development, it may well prove a valuable piece in the quantum toolbox.

Science, this issue p. 602; see also p. 580

Carleo, Giuseppe, and Matthias Troyer. Science 355.6325 (2017): 602-606.

Real-space Quantum Monte Carlo

Variational Monte Carlo

$$E_0 = min < \psi|\hat{H}|\psi >$$

ψ : quantum state

😔 Sub-optimal accuracy

😊 Have access to the wavefunction explicitly

😊 No sign-problem

Provide trial wavefunction

Diffusion Monte Carlo

$$\psi \rightarrow e^{-\hat{H}\Delta t}\psi$$

Schrödinger Equation

Variational Principle

Convert PDE-solving to optimization

Ground State Energy
(Smallest eigenvalue)

$$\hat{H}\psi = E\psi$$

Energy of ψ

$$E_0 = \min_{\psi} \boxed{<\psi|\hat{H}|\psi>}$$

ψ : quantum state

Variational Monte Carlo

$$< \psi|\hat{H}|\psi > = \int \psi(X)\hat{H}\psi(X)dX = E_{X\sim\psi^2}\frac{\hat{H}\psi(X)}{\psi(X)}$$

ψ: Real-space wavefunction
X: electron configuration in real-space $\begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}$ n: number of electrons

In practice, the energy is estimated by

1. Sampling a number of electron configurations $X_1, \ldots, X_N$ from distribution Ψ^2
2. Calculating the function values $\frac{\hat{H}\psi(X_i)}{\psi(X_i)}$ (called "local energy")
3. Taking average $\mathrm{E} \approx \frac{1}{N}\sum_i \frac{\hat{H}\psi(X_i)}{\psi(X_i)}$

Variational

Monte Carlo

Optimize over a class of wavefunctions

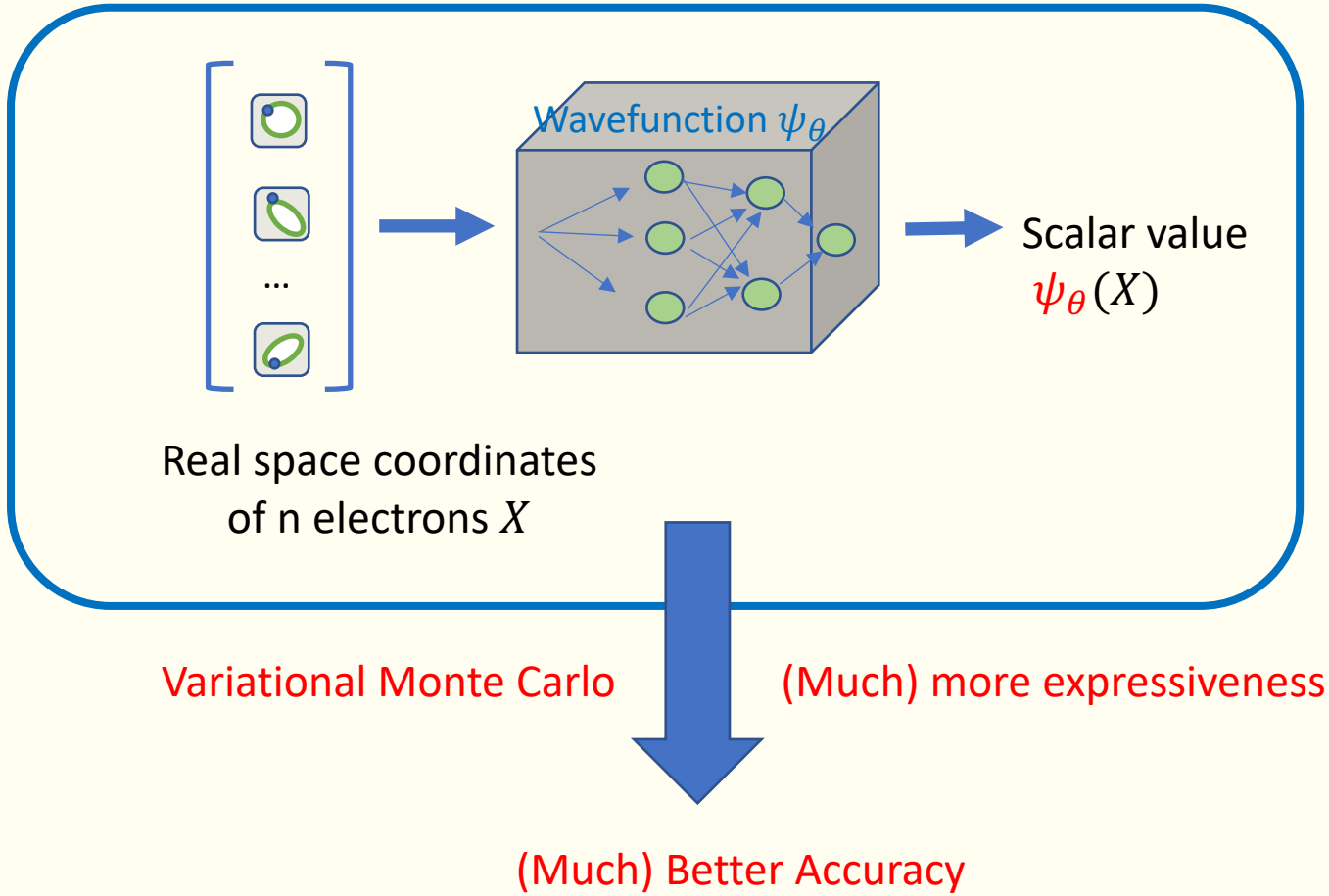
Use Monte Carlo to calculate energy

$$E_0 = \min_{\psi} < \psi|\hat{H}|\psi >$$

1. Parametrize a class of wavefunctions
2. Optimize the parameters to minimize energy

Neural Network as Wavefunction

$$E_0 = \min_{\psi} < \psi | \hat{H} | \psi >$$



Physical constraints for electronic wavefunctions

Permutation Anti-symmetry

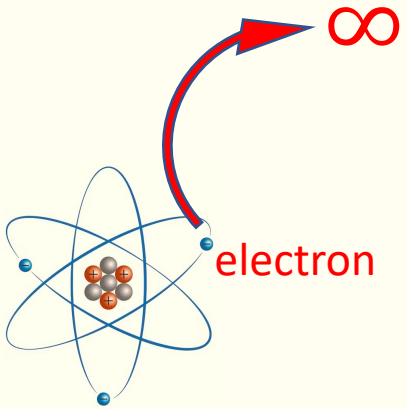
$$F\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_i \\ \dots \\ x_j \\ \dots \\ x_n \end{pmatrix} = -F\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_j \\ \dots \\ x_i \\ \dots \\ x_n \end{pmatrix}$$

Electron position vector: dim-3

Ignore spin for now

Boundary Condition

$$F\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_i \\ \dots \\ x_j \\ \dots \\ x_n \end{pmatrix} \longrightarrow 0$$



Explicit form of wavefunctions

Slater Determinant

$$\psi(x_1, \dots, x_n) = det \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_n) & \cdots & \phi_n(x_n) \end{bmatrix}$$

$\phi_1, \dots, \phi_n$ are one-electron orbitals

With Backflow

$$\psi\,(x_1, \dots, x_n)$$

$$= det \begin{bmatrix} \Phi_1\,(x_1; x_2, \dots, x_n) & \cdots & \Phi_n\,(x_1; x_2\,, \dots, x_n) \\ \vdots & \ddots & \vdots \\ \Phi_1\,(x_n; x_1, \dots, x_{n-1}) & \cdots & \Phi_n(x_n; x_1, \dots, x_{n-1}) \end{bmatrix}$$

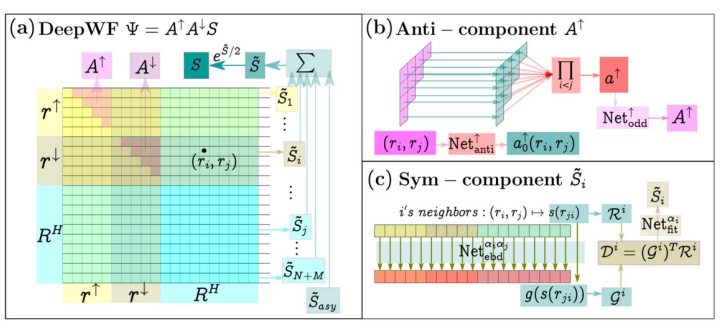
$\Phi_1, \dots, \Phi_n$ are “many-electron orbitals”,

Modeled with neural networks, carefully

designed to satisfy anti-symmetry

DeepWF

2018 Princeton University

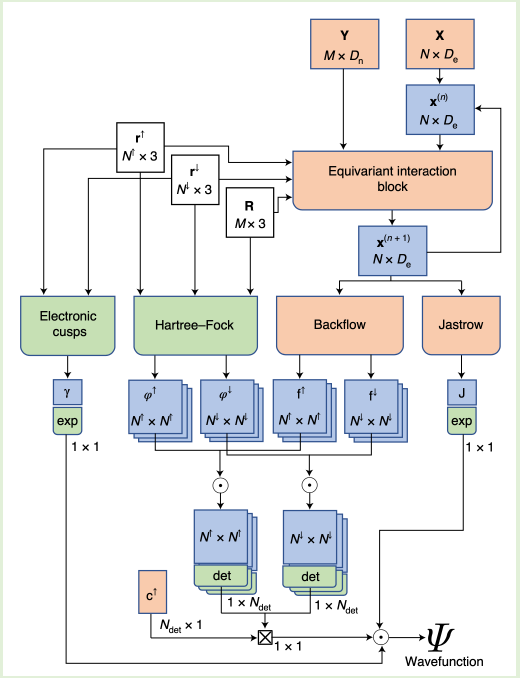


Han, J., Zhang, L., & Weinan, E. (2019). Solving many-electron Schrödinger equation using deep neural networks. *Journal of Computational Physics*, *399*, 108929.

PauliNet

2019 FU Berlin

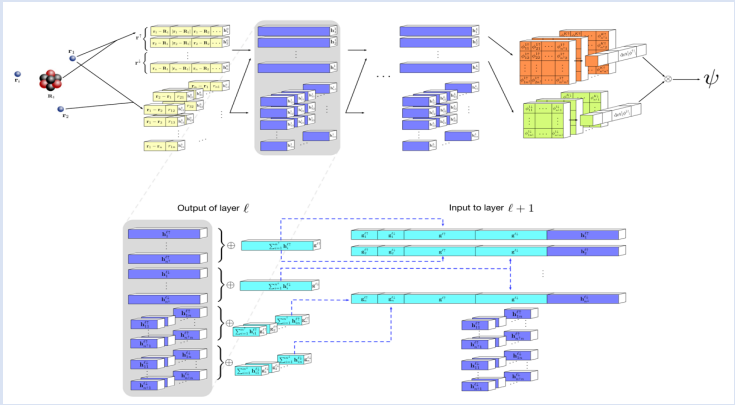
Authors are also in Microsoft now



Hermann, J., Schätzle, Z., & Noé, F. (2020). Deep-neural-network solution of the electronic Schrödinger equation. *Nature Chemistry*, *12*(10), 891-897.

FermiNet

2019 DeepMind & ICL



Pfau, D., Spencer, J. S., Matthews, A. G., & Foulkes, W. M. C. (2020). Ab initio solution of the many-electron Schrödinger equation with deep neural networks. *Physical Review Research*, *2*(3), 033429.

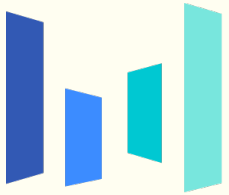
Train the neural network

- The wavefunction ψ is modelled by a (specially designed) neural network.
- Loss Function: Energy $\mathrm{E} = <\psi|\hat{H}|\psi> = E_{X\sim\psi^2}\frac{\hat{H}\psi(X)}{\psi(X)}$
- Data: The electronic configurations sampled from ψ^2.
- Procedure
  1. Sampling a number of electron configurations $X_1, \ldots, X_N$ from distribution ψ^2
  2. Calculating the function values $\frac{\hat{H}\psi(X_i)}{\psi(X_i)}$
  3. Taking average $\mathrm{E} \approx \frac{1}{N}\sum_i \frac{\hat{H}\psi(X_i)}{\psi(X_i)}$
  4. Gradient descent and update ψ, then loop over.

Related Works from ByteDance Research & PKU

State of the Art

(since 2021)



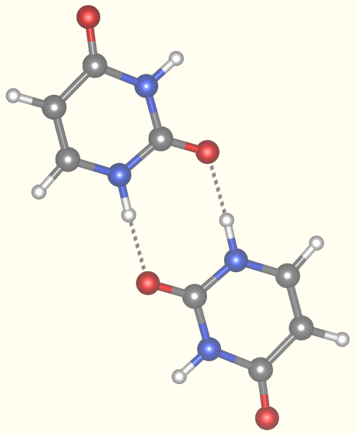
• Novel phases for Moire systems	**arXiv:2406.11134**
• Spin symmetry enforcement	**Nat. Comput. Sci. in press**
• Break through the computational bottleneck of Laplacian	**Nat. Mach. Intell. 6, 209 (2024)**
• Electric polarization in solids	**Phys. Rev. Lett. 132, 176401 (2024)**
• Better extrapolation scheme	**Mach. Learn.: Sci. Technol. 015016 (2024)**
• From VMC to DMC	**Nat. Commun. 14, 1860 (2023)**
• From molecules to solids	**Nat. Commun. 13, 7895 (2022)**
• Incorporate effective core potential	**Phys. Rev. Research, 4, 013021 (2022)**

Make NNQMC more practical

- NNQMC + Effective Core Potential
  - Less electrons in calculations.
  - *Li, Xiang, et al. Physical Review Research 4.1 (2022): 013021.*
- NN + Diffusion Monte Carlo
  - Less optimization steps
  - *Ren, Weiluo, et al. Nature Communications 14.1 (2023): 1860.*
- Forward Laplacian
  - Address computation bottleneck: Laplacian Operator
  - *Li, Ruichen, et al. Nature Machine Intelligence 6.2 (2024): 209-219.*

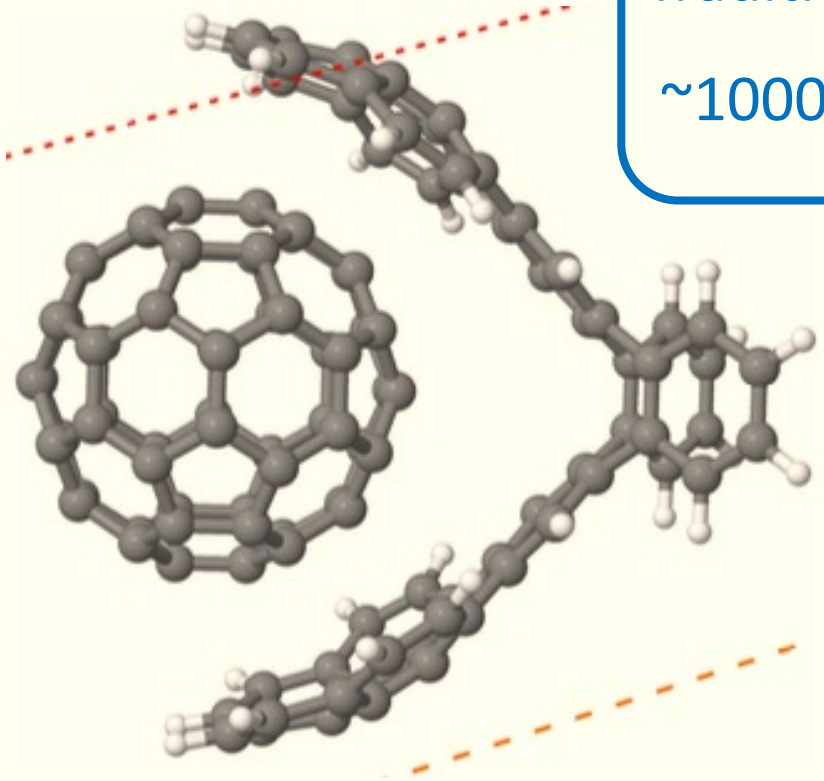
We can now handle:

100~150 electrons



Traditional QMC can handle

~1000 electrons

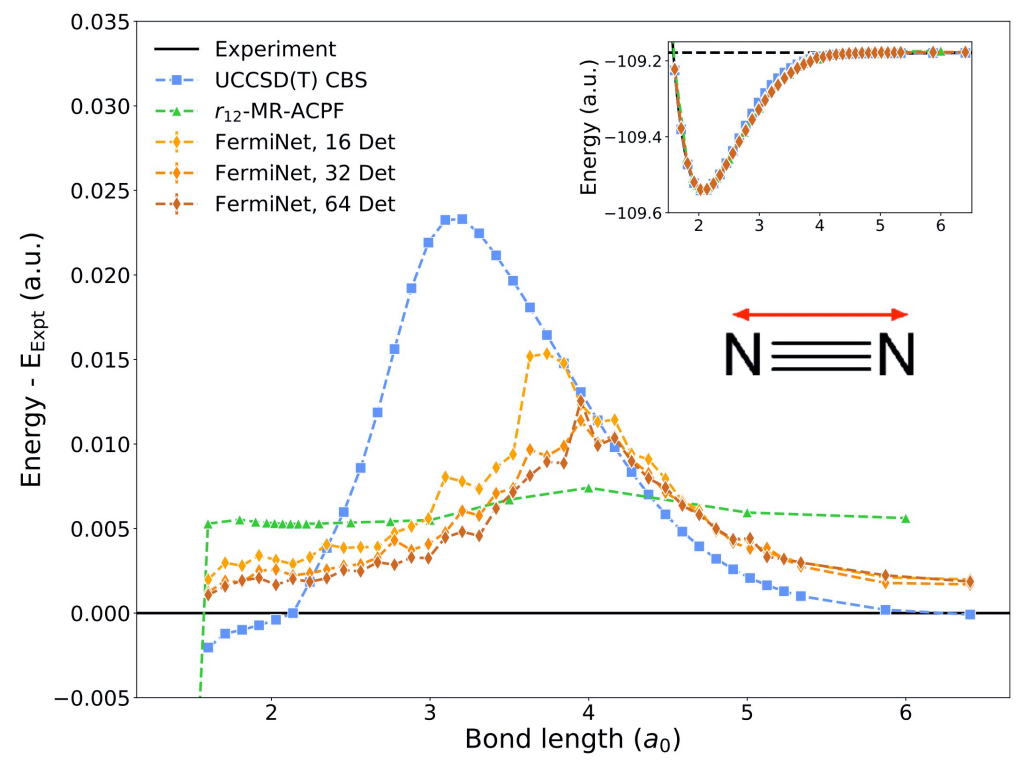


Al-Hamdani, Yasmine S., et al. "Interactions between large molecules pose a puzzle for reference quantum mechanical methods." Nature Communications 12.1 (2021): 3927.

N_2 energy curve

FermiNet from Google DeepMind (2020)

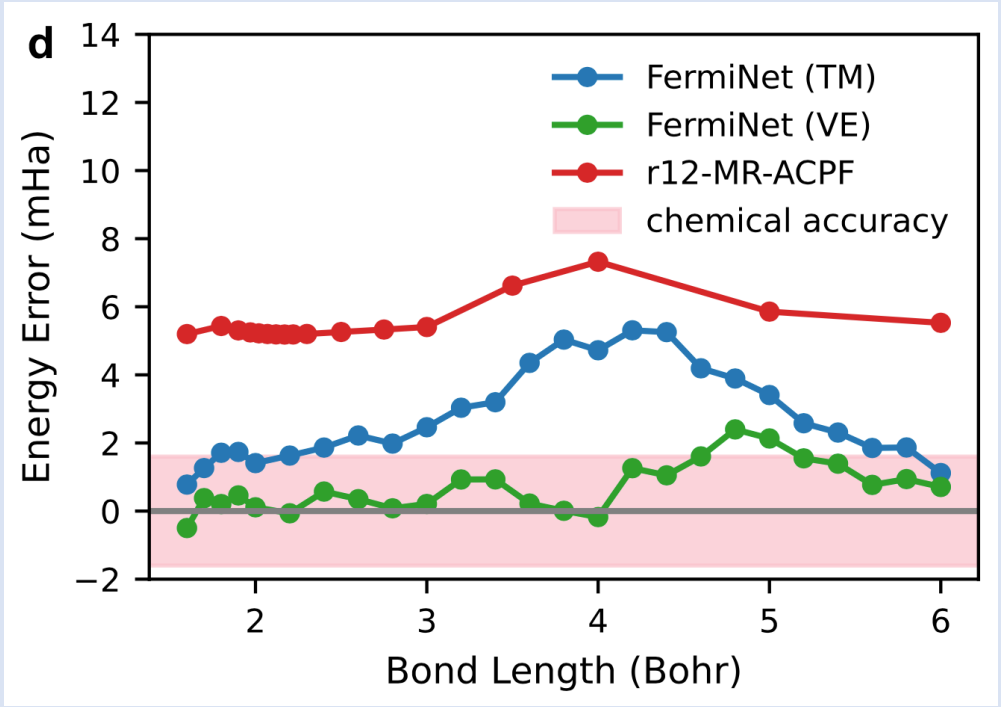
Beat Golden Standard Method



Pfau, David, et al. Physical Review Research 2.3 (2020): 033429.

Our work, better extrapolation scheme (2023)

Chemical Accuracy (almost)



Fu, Weizhong, et al. Machine Learning: Science and Technology 5.1 (2024): 015016.

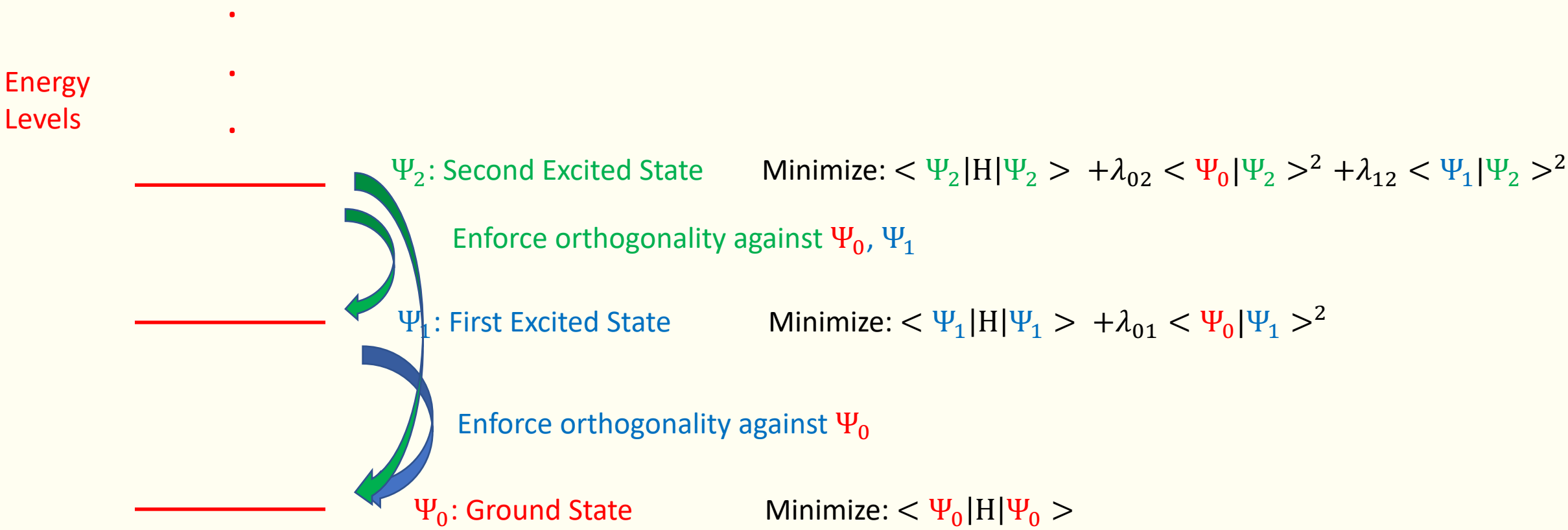
Outline

Introduction to NNQMC

Excited States with NNQMC

Novel Phase of Moire systems

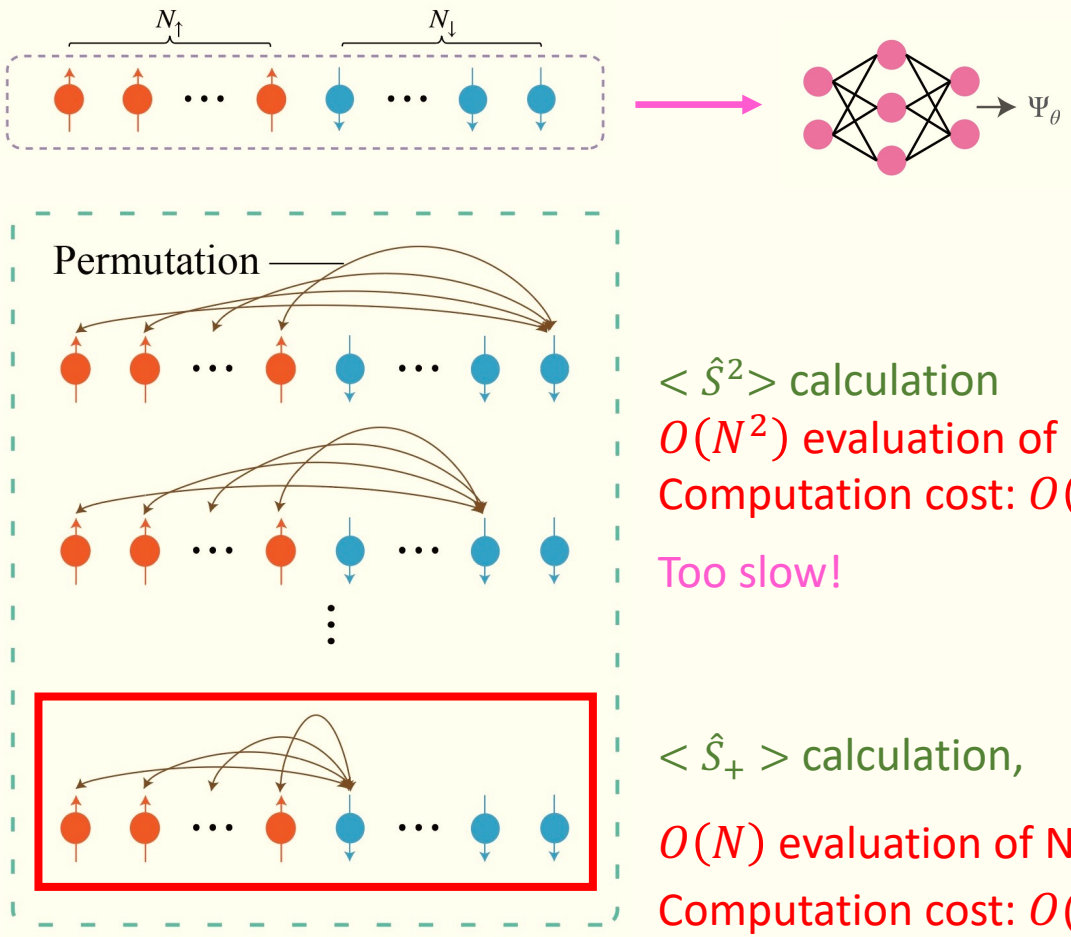
Penalty to enforce orthogonality



Enforce Spin-Symmetry in NNQMC

Li, Zhe, et, al. arXiv:2406.01222, Nature Computational Science in press

- Spin-pure state
  - Eigen-state of $\hat{S}^2$ operator
- Add "deviation of $<\hat{S}^2>$ to target value" to loss
  - Similar to the overlap for excited states
- $<\hat{S}^2>$ calculation is expensive
- Consider $\hat{S}_+$ operator instead
  - One order cheaper
  - Still able to simulate eigenstate of $\hat{S}^2$



$<\hat{S}^2>$ calculation

$O(N^2)$ evaluation of NN

Computation cost: $O(N^5)$

Too slow!

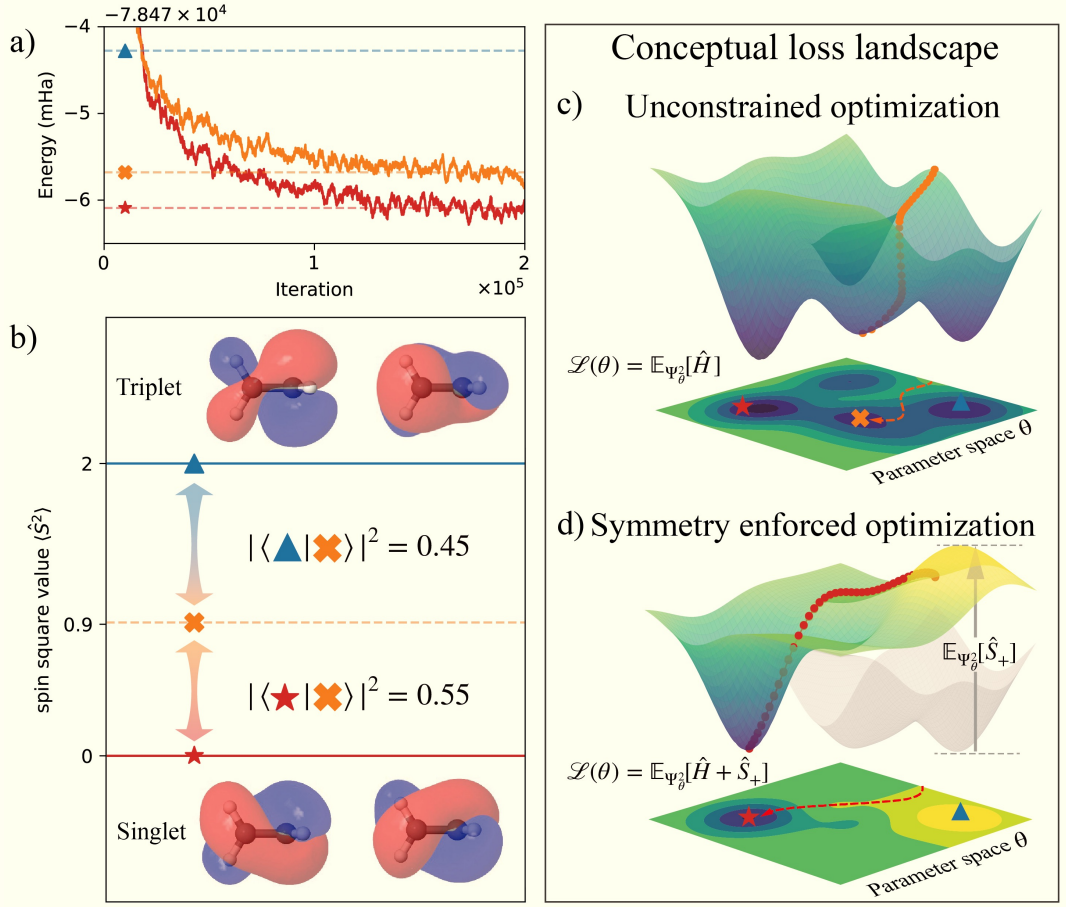
$<\hat{S}_+>$ calculation,

$O(N)$ evaluation of NN

Computation cost: $O(N^4)$

dubbed as "$\hat{S}_+$ penalty"

Spin symmetry helps ground state calculation



Ethylene C_2H_4

★ singlet state: symmetry enforced

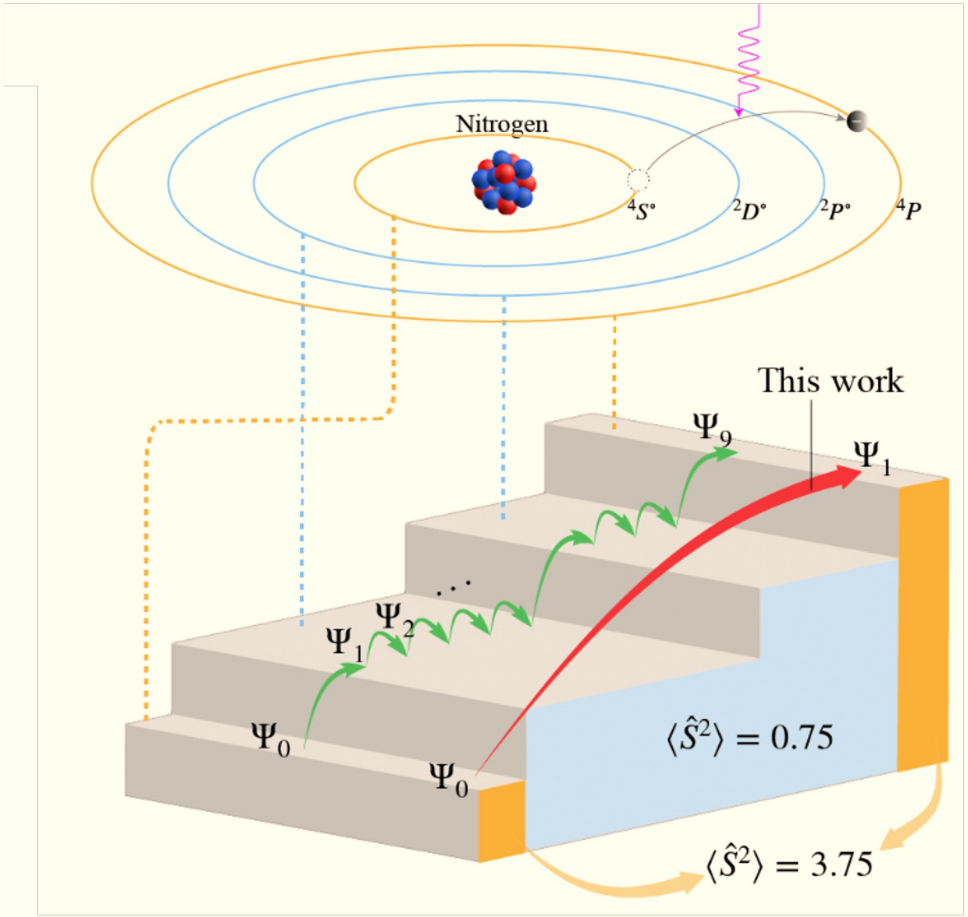
✖ singlet state: free of constraint (contaminated)

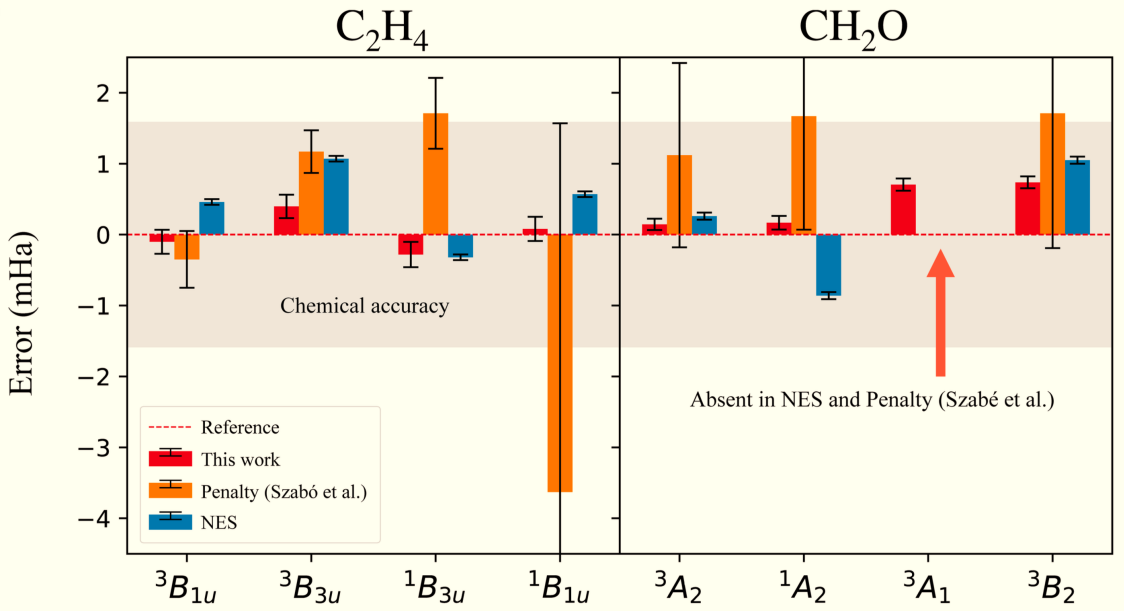
▲ triplet state

Li, Zhe, et, al. arXiv:2406.01222, Nature Computational Science in press

Spin symmetry helps excited state calculation

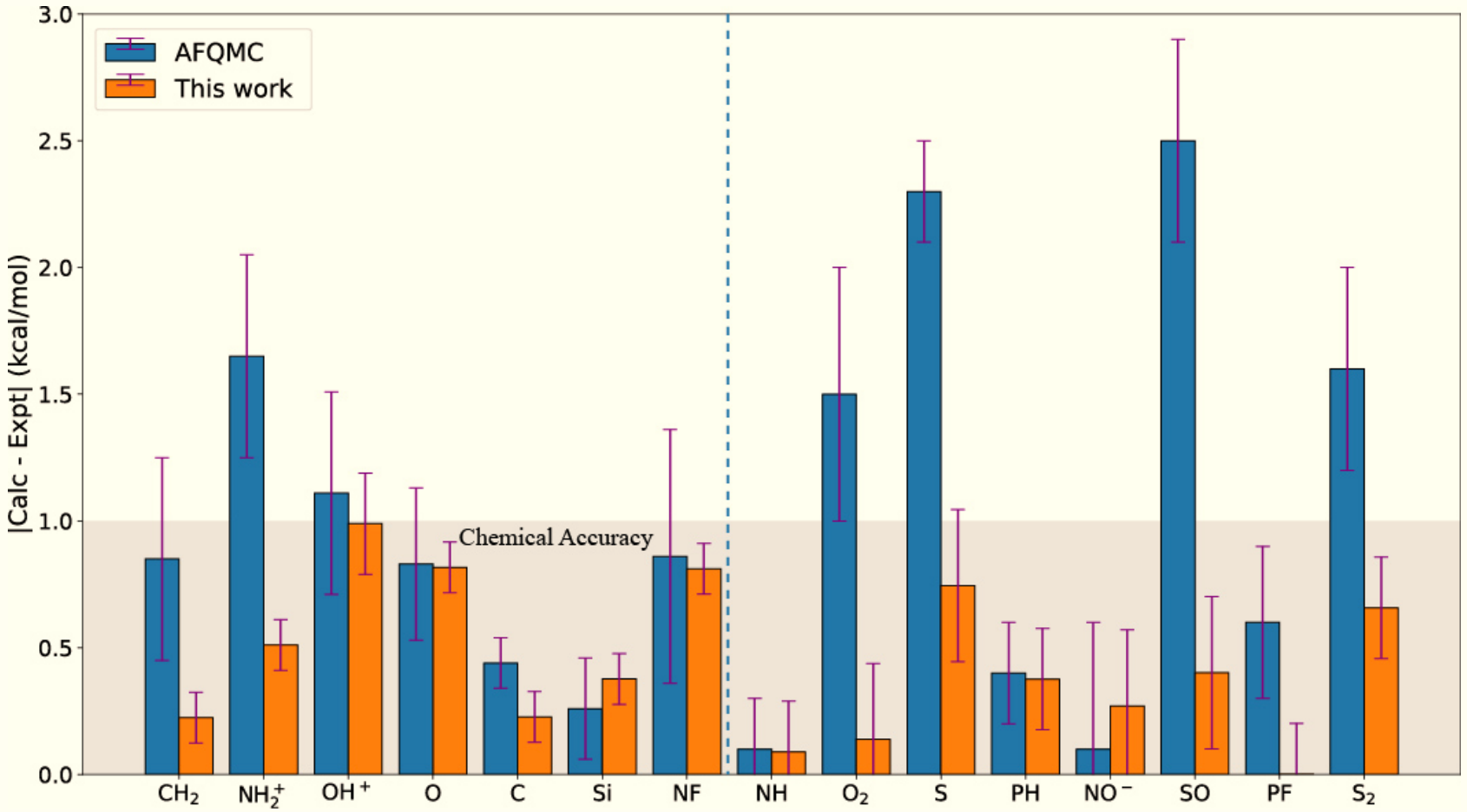
High excited states





Li, Zhe, et, al. arXiv:2406.01222, Nature Computational Science in press

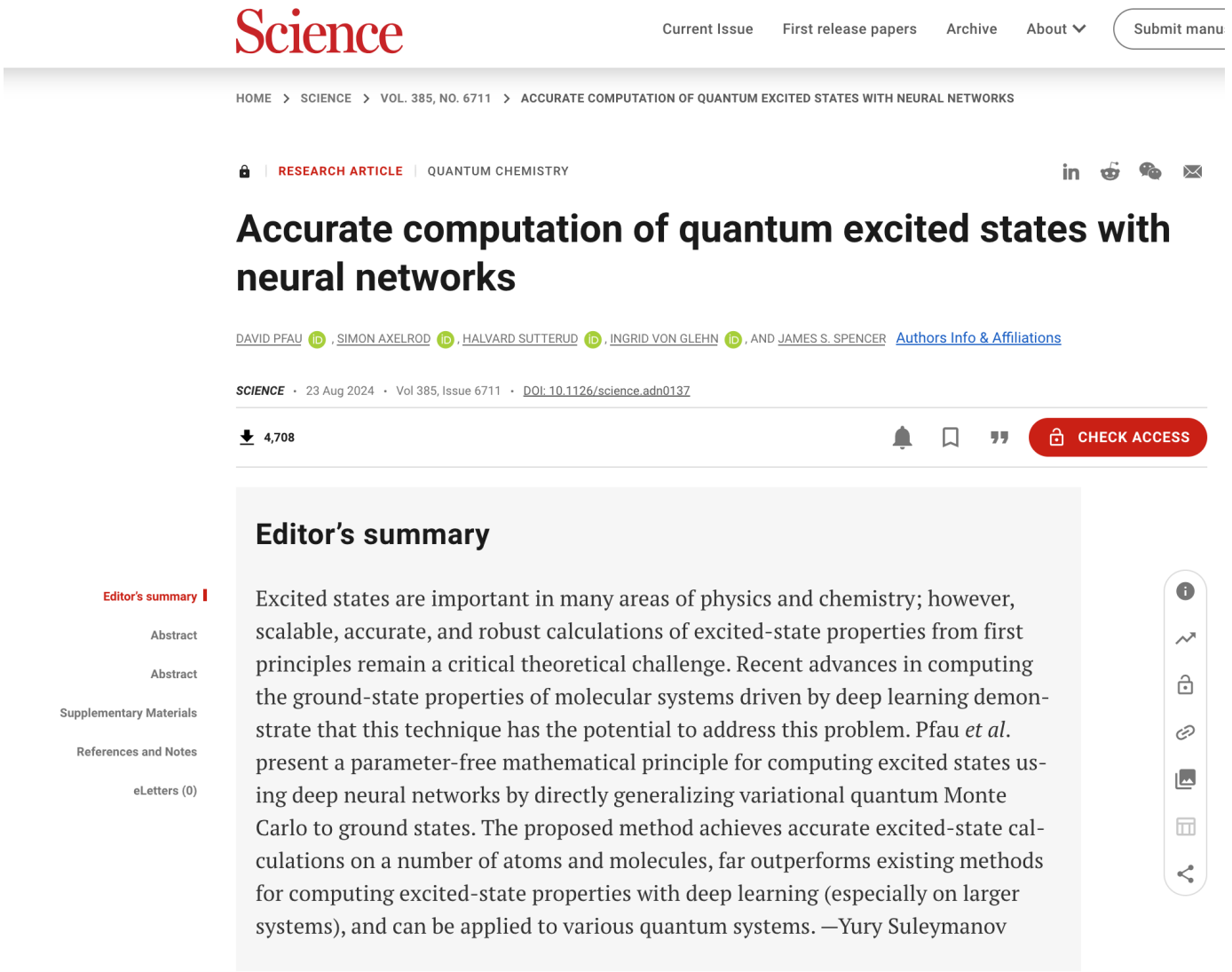
Singlet-triplet gap of biradicals



Li, Zhe, et, al. arXiv:2406.01222, Nature Computational Science in press

Natural Excited State

Pfau, David, et al. *Science* 385.6711 (2024): eadn0137.



Slater Determinant

$$\psi(x_1, \dots, x_n) = det \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_n) & \cdots & \phi_n(x_n) \end{bmatrix}$$

NNVMC (backflow)

$$\Psi(X) = \Psi(x_1, \dots, x_n)$$
$$= det \begin{bmatrix} \Phi_1(x_1; x_2, \dots, x_n) & \cdots & \Phi_n(x_1; x_2, \dots, x_n) \\ \vdots & \ddots & \vdots \\ \Phi_1(x_n; x_1, \dots, x_{n-1}) & \cdots & \Phi_n(x_n; x_1, \dots, x_{n-1}) \end{bmatrix}$$

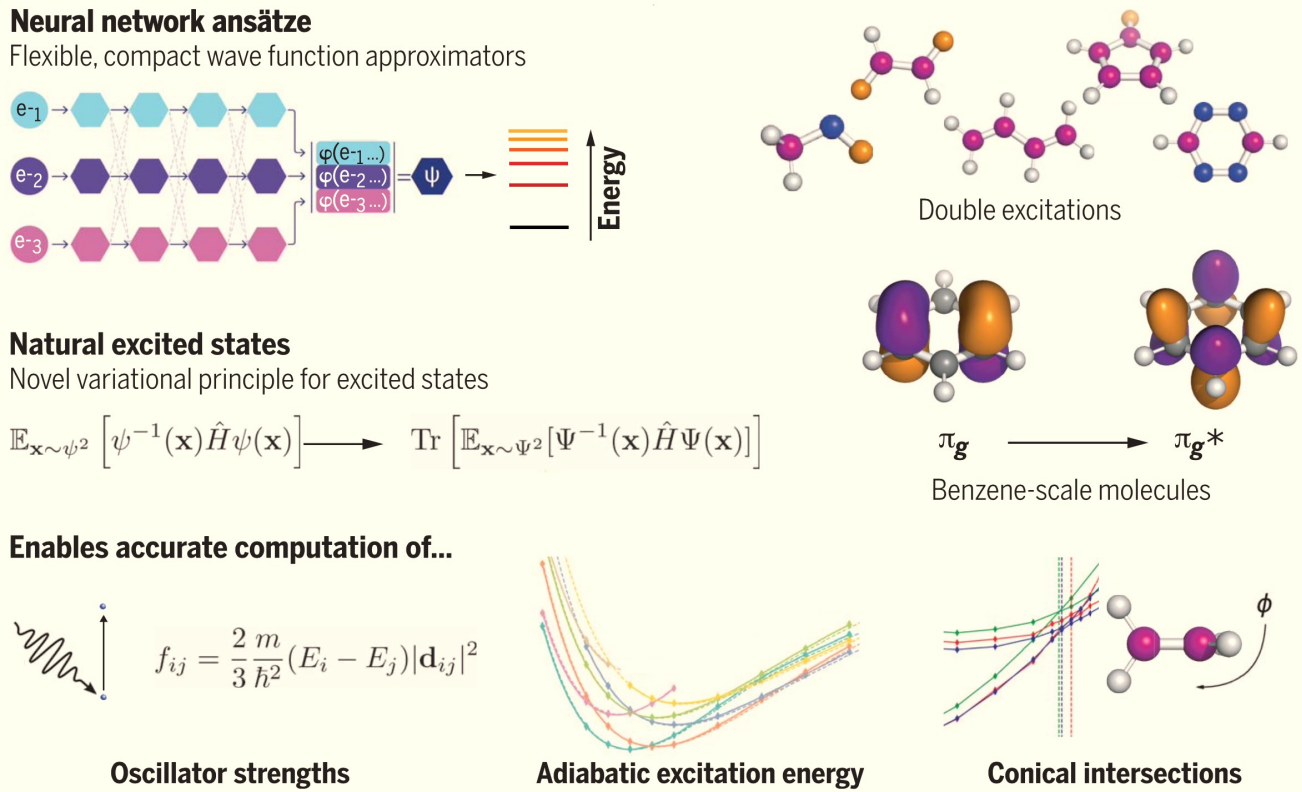
Natural Excited State

$$\boldsymbol{\Psi}(\boldsymbol{X^1}, \dots, \boldsymbol{X^k}) = det \begin{bmatrix} \Psi_1(\boldsymbol{X^1}) & \cdots & \Psi_k(\boldsymbol{X^1}) \\ \vdots & \ddots & \vdots \\ \Psi_1(\boldsymbol{X^k}) & \cdots & \Psi_k(\boldsymbol{X^k}) \end{bmatrix}$$

- Ψ_1,…, Ψ_k are guaranteed to be non-collapsed.
- We minimize the sum of their energies.
- Optimize for the best k-dim subspace

Pfau, David, et al. *Science* 385.6711 (2024): eadn0137.

Natural Excited State



Natural excited states. Combining neural networks with a mathematical insight enables accurate calculations of challenging excited states of molecules.

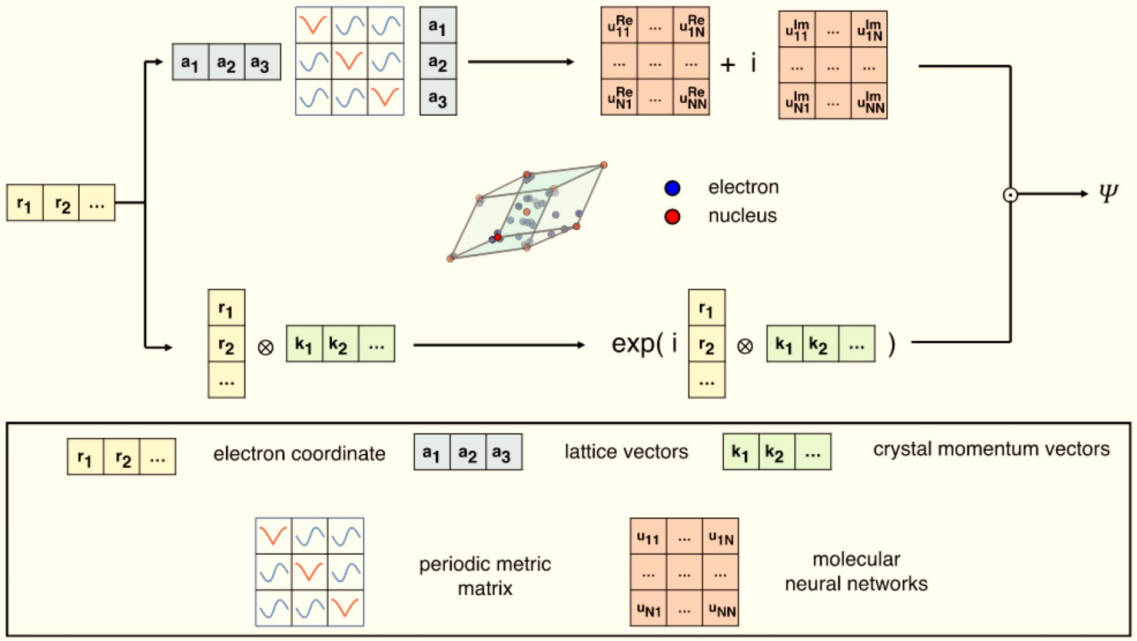
Outline

Introduction to NNQMC

Excited States with NNQMC

Novel Phase of Moire systems

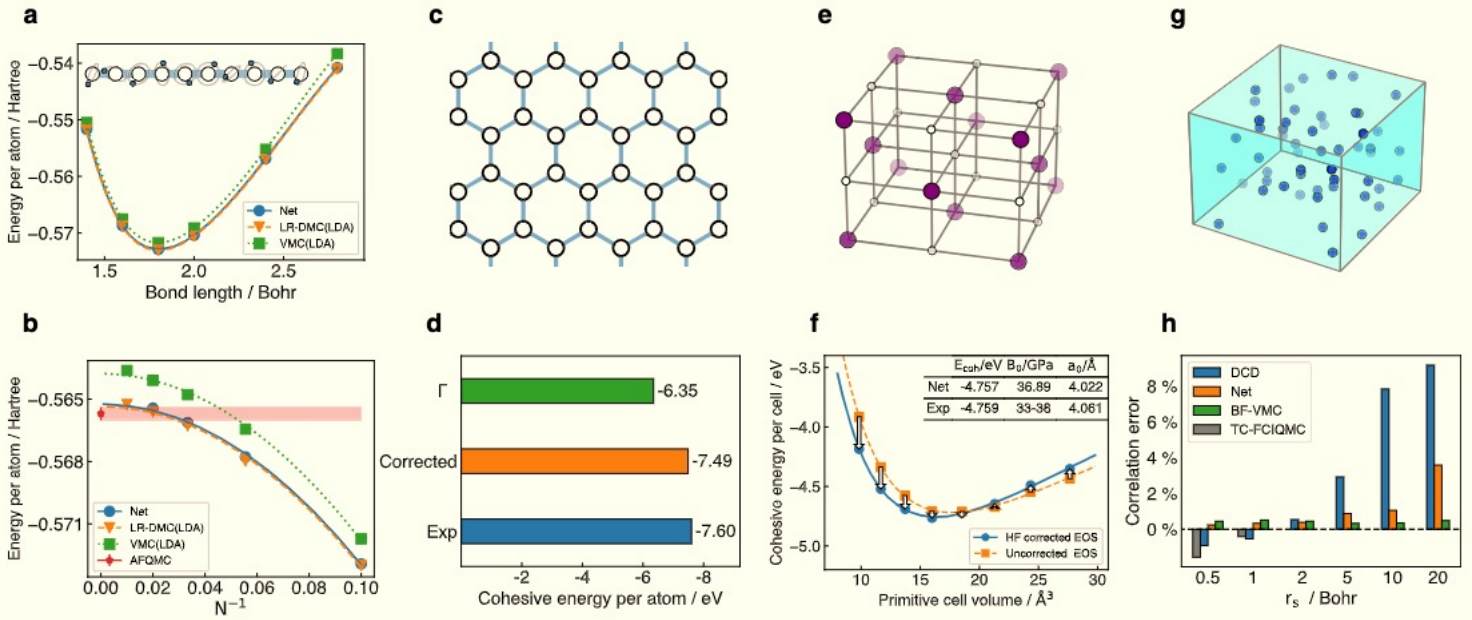
Modeling Real Solids - DeepSolid



- Modify the input the feature with periodic functions
- Carefully deal with the smoothness near the boundary.

Li, X., Li, Z., & Chen, J. (2022). *Nature Communications*, *13*(1), 7895.

DeepSolid cont’d



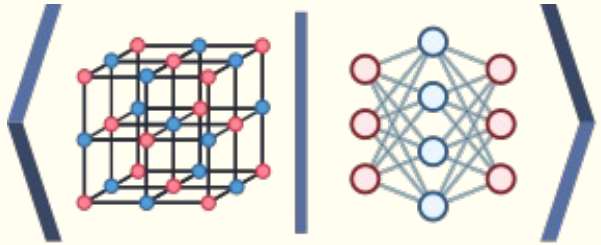
Li, X., Li, Z., & Chen, J. (2022). *Nature Communications*, *13*(1), 7895.

More on NNQMC for solids

Checkout our recent review

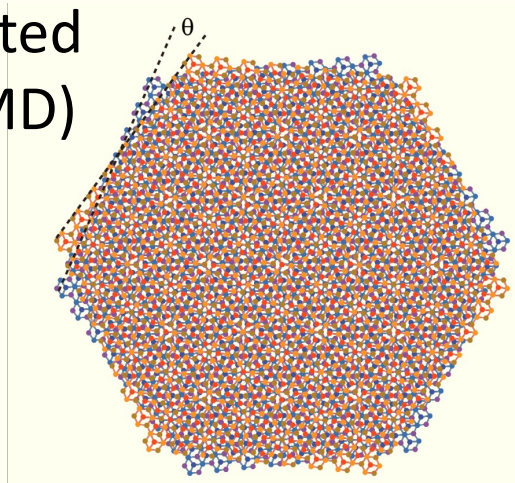
Qian, Y., Li, X., Li, Z., Ren, W., & Chen, J. (2024).
Deep learning quantum Monte Carlo for solids.

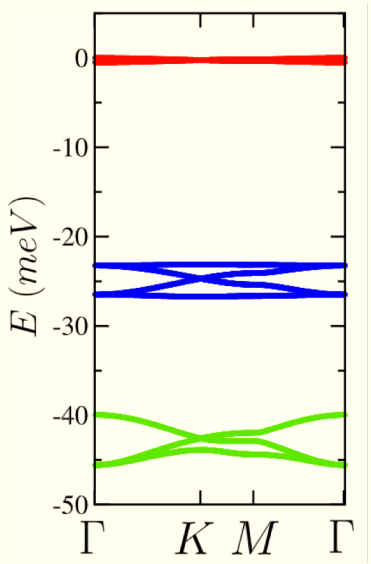
arXiv:2407.00707.



Moire Material

Twisted graphene or twisted metal dichalcogenide (TMD)





Flat band appears which signals strong correlation

Spontaneous symmetry breaking,
Mott insulator,
Wigner crystal,
Unconventional superconductivity,
Anomalous Quantum hall……

$$H = \sum_i -\frac{1}{2m^*}\Delta_i + V(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}\frac{1}{\epsilon\,|\,\mathbf{r}_i - \mathbf{r}_j\,|}$$

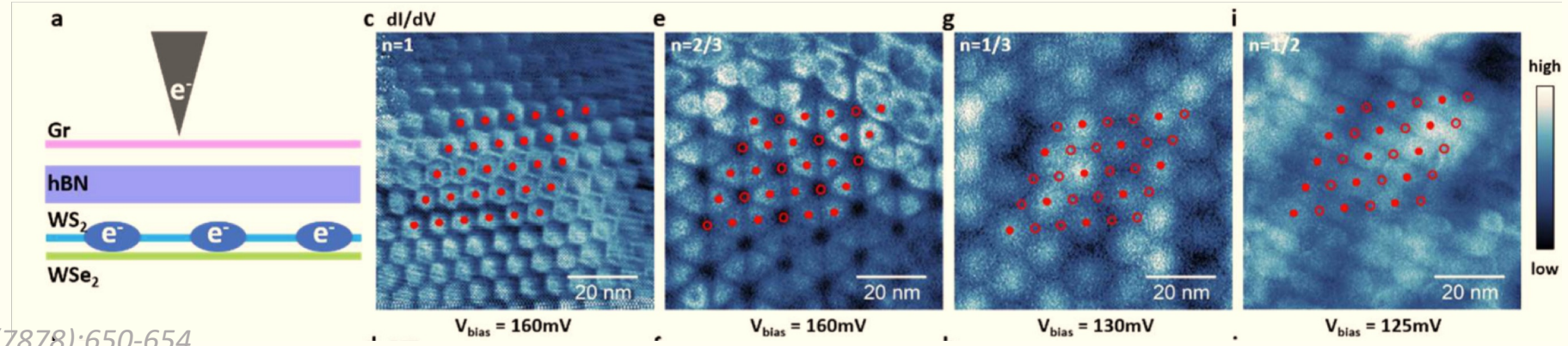
Generalized Wigner Crystal

WSe_2/WS_2

Triangular

Stripe

STM observation of GWC

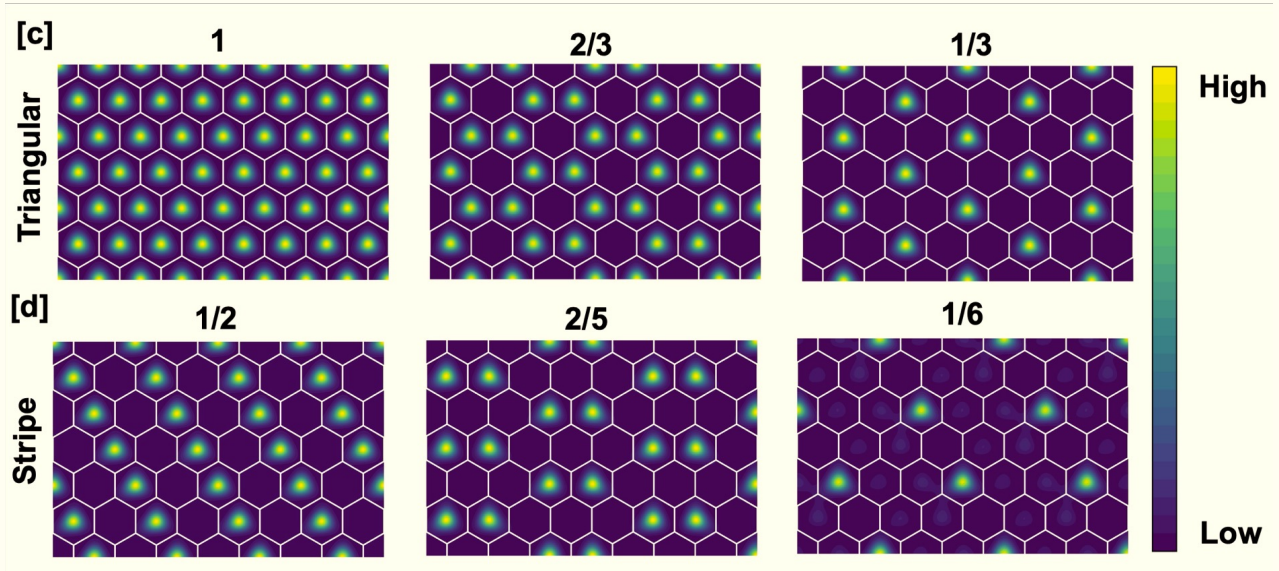


Li, Hongyuan, et al. Nature 597(7878):650-654

DeepSolid result

Our neural networks reproduced patterns in STM experiments!

Li, Xiang, et, al. arXiv:2406.11134

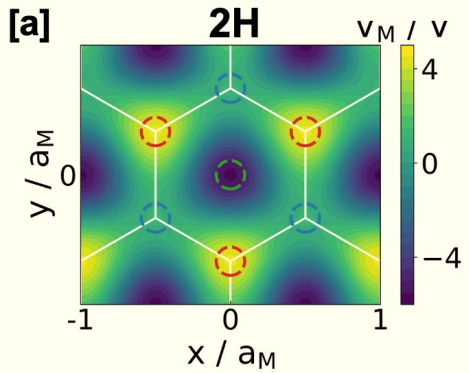


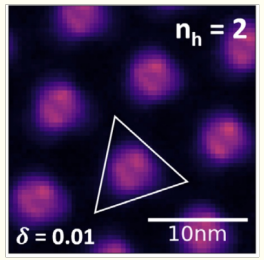
More interesting phases

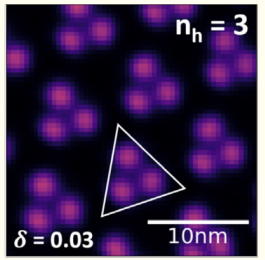
Wigner Molecular Crystal

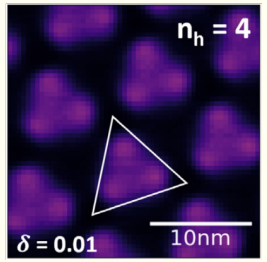
WS_2 homobilayer

2H config

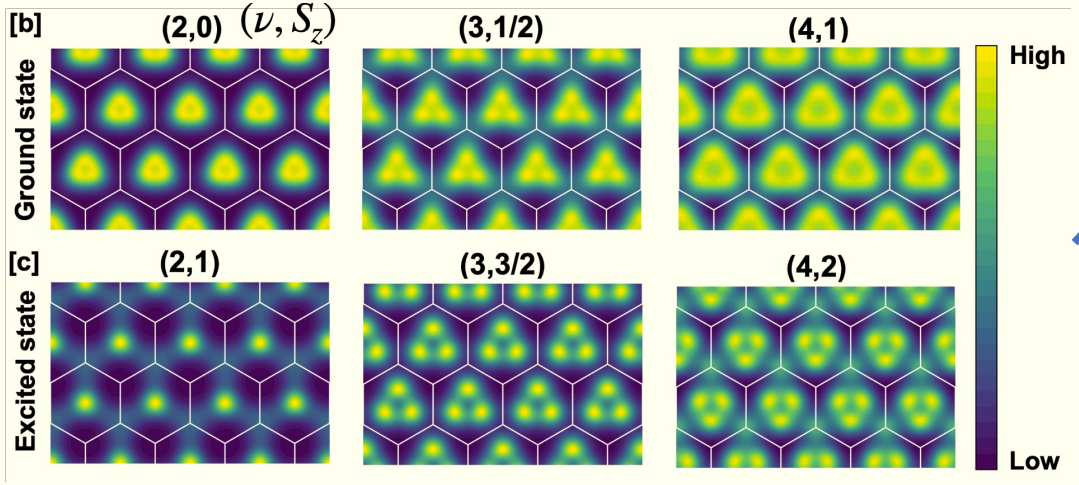








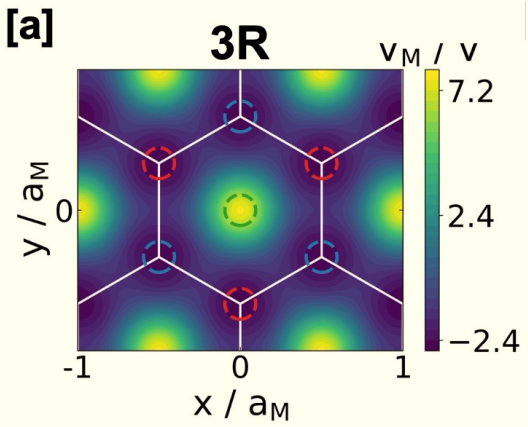
Reddy, Aidan, et al. PRL, 2465016, 2023
Li, Hongyuan, et al. Science 385 (6704), 86-91, 2024



Reproduced patterns
in STM experiments!

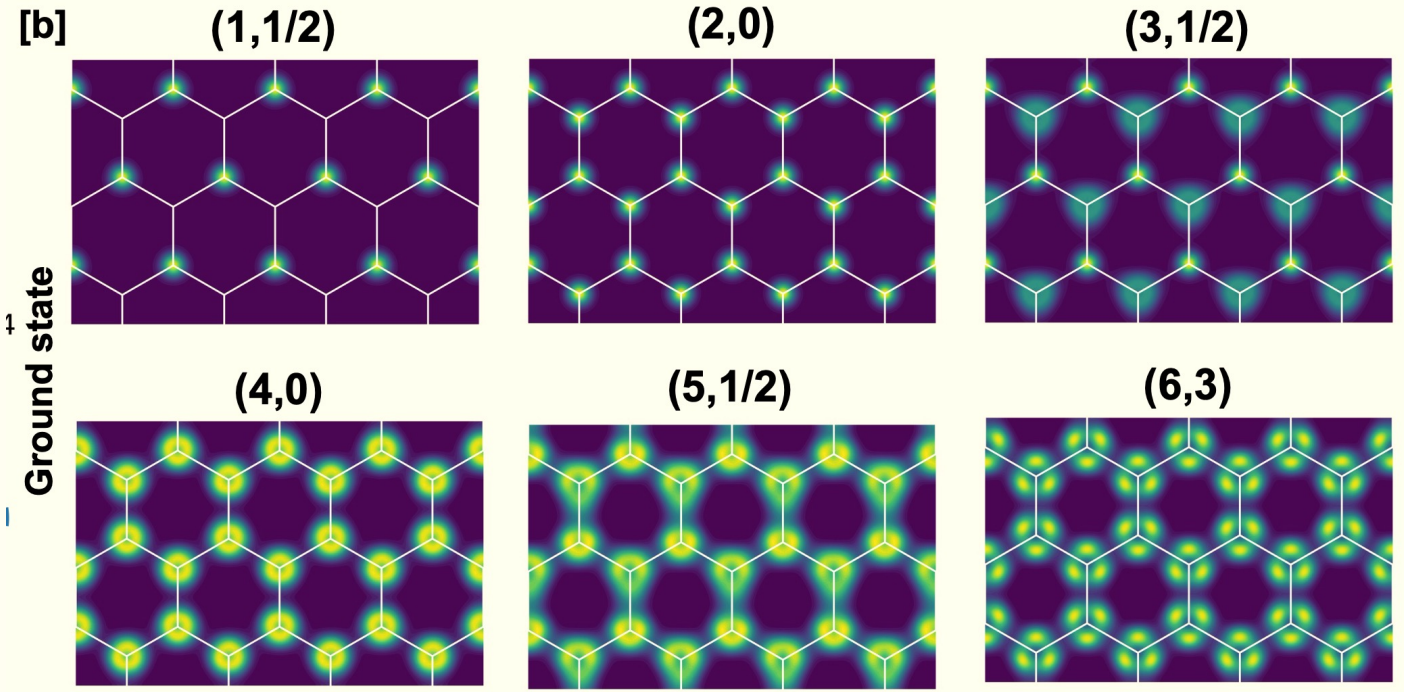
Li, Xiang, et, al. arXiv:2406.11134

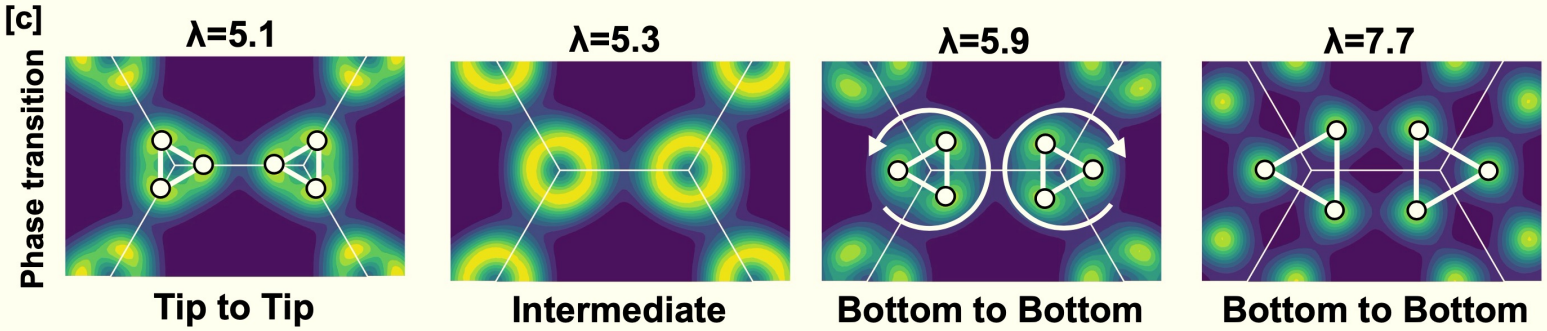
Potential with C_6 symmetry



WS_2 homobilayer

3R config



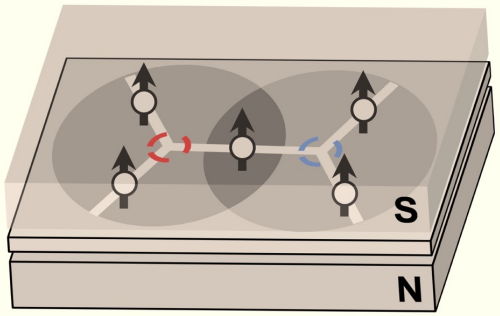


Larger λ means
Coulumb dominates more

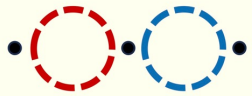
Wigner Covalent Crystal

Yet to be confirmed
in experiments !

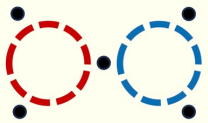
[a] Covalent bond



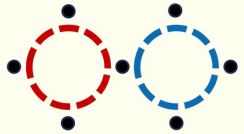
Monomer bond



Dimer bond

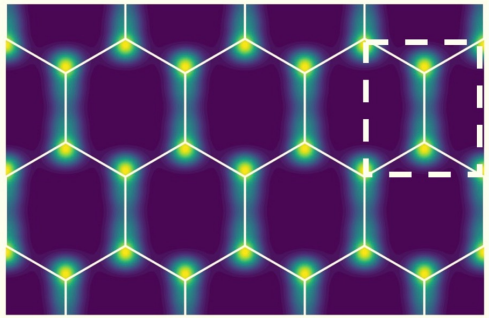


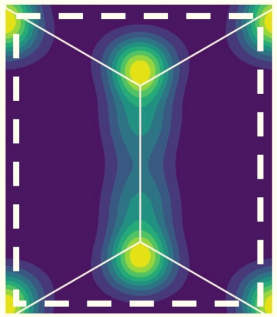
Trimer bond



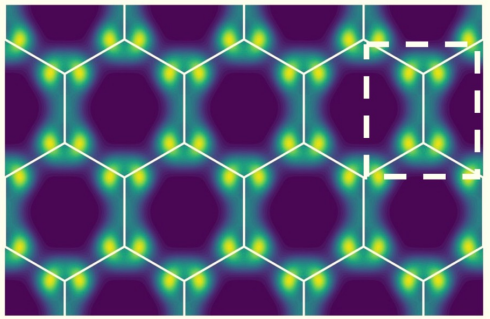
[b]

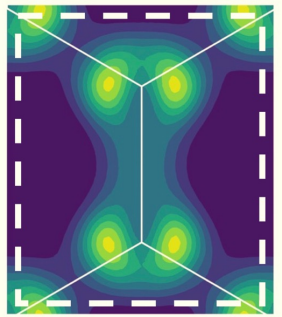
(3,3/2)



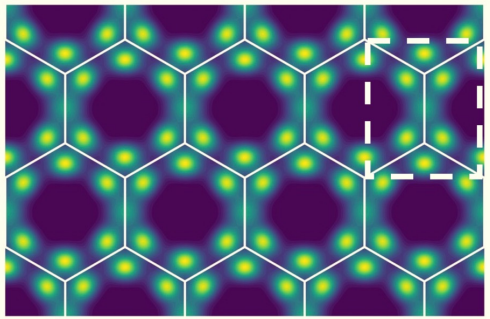


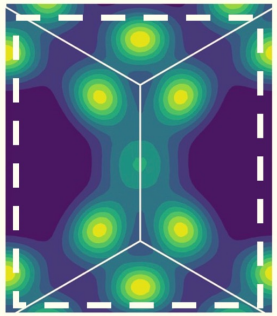
(5,5/2)



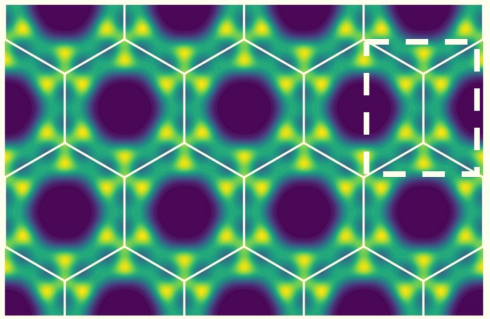


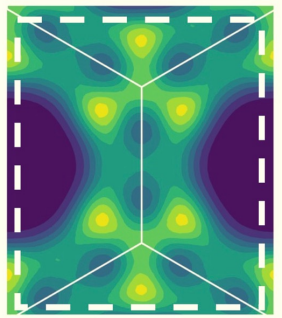
(7,7/2)





(9,9/2)





Li, Xiang, et, al. arXiv:2406.11134

Other applicable systems

- Electron gas
- Positron
- Superfluid
- ……

Outlook

- Better efficiency and accuracy.
- Expand the scope of application.
- Tackle challenging systems
  - Strong correlation
  - Multi-reference
  - ……

Thank you!

- https://github.com/bytedance/FermiNet_with_ECP
- https://github.com/bytedance/deepsolid
- https://github.com/bytedance/jaqmc
- https://github.com/bytedance/lapnet
- https://github.com/bytedance/netobs