Neural-network variatonal Monte Carlo for atomic nuclei

Yilong Yang

Peking University

Supervisors: Jie Meng, Pengwei Zhao

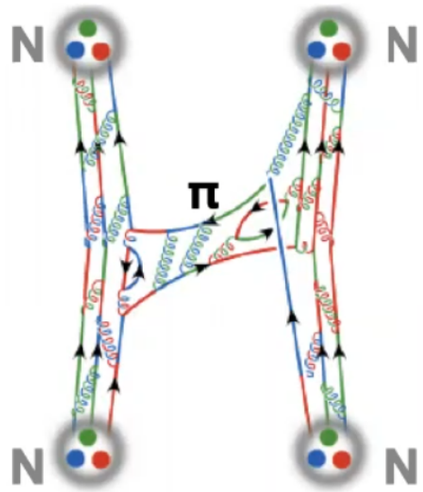
Outline

- Introduction
- Neural-network Variational Monte Carlo (VMC)
- Neural-network VMC for lattice EFT
- Summary and outlooks

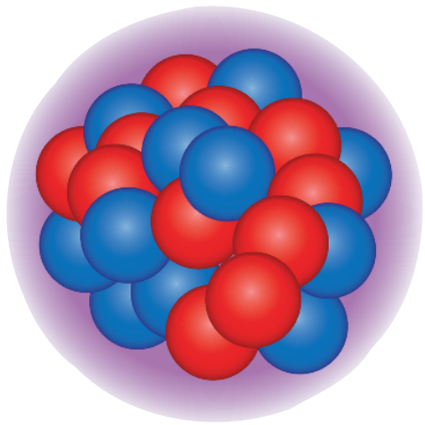
Ab initio nuclear theory

Overarching goal: Understand nuclear properties from a unified theory rooted in the forces among nucleons

- **Fix the nuclear force**
  - Phenomenological, meson-exchanges models
  - Effective field theories (EFTs)



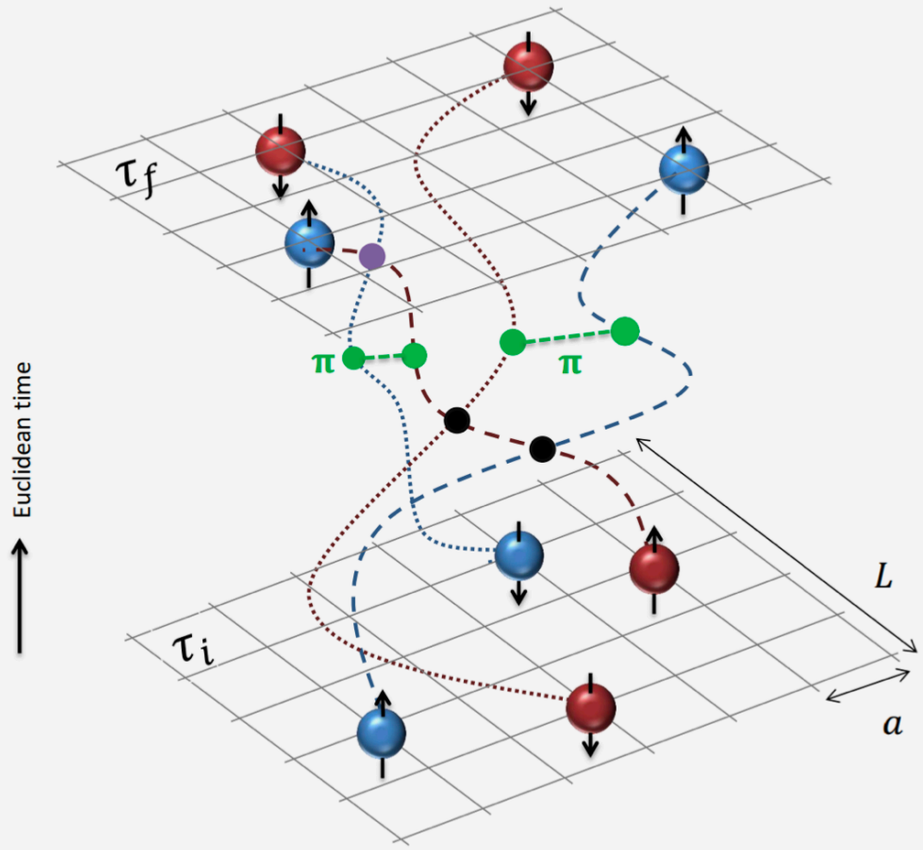
- **Solve the nuclear many-body problem**
  - Nuclear bulk properties: masses, radii, …
  - Nuclear spectra: energy levels, transitions, …
  - Nucleonic matter EoS: neutron stars, …
  - New physics: $0\nu\beta\beta$, electric dipole moments, …
  - …



Quantum Monte Carlo (QMC)

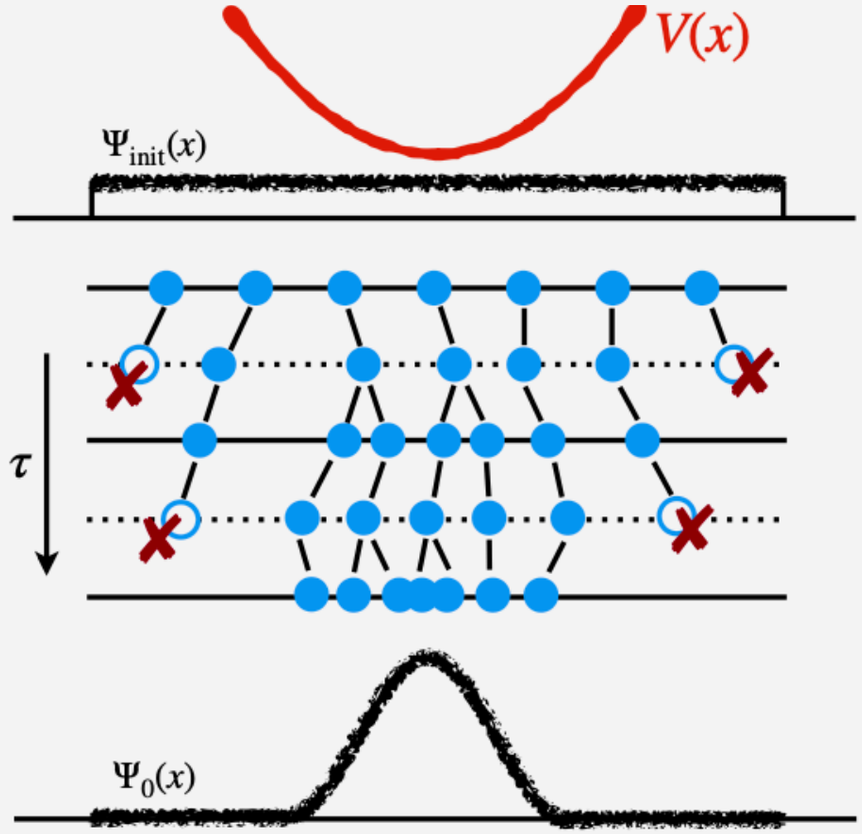
Solving quantum many-body problem with Monte Carlo techinques, widely used in Lattice QCD, Nuclear Physics, Quantum chemistry …

Lattice QMC



Dean Lee, Prog. Part. Nucl. Phys. 63, 117 (2009),
Lähde and Meißner, “Nuclear Lattice EFT”, Springer (2019)

Continuum QMC



Carlson et al., Rev. Mod. Phys. 87, 1067 (2015)
Gandolfi et al, Front. Phys. 8 (2020)

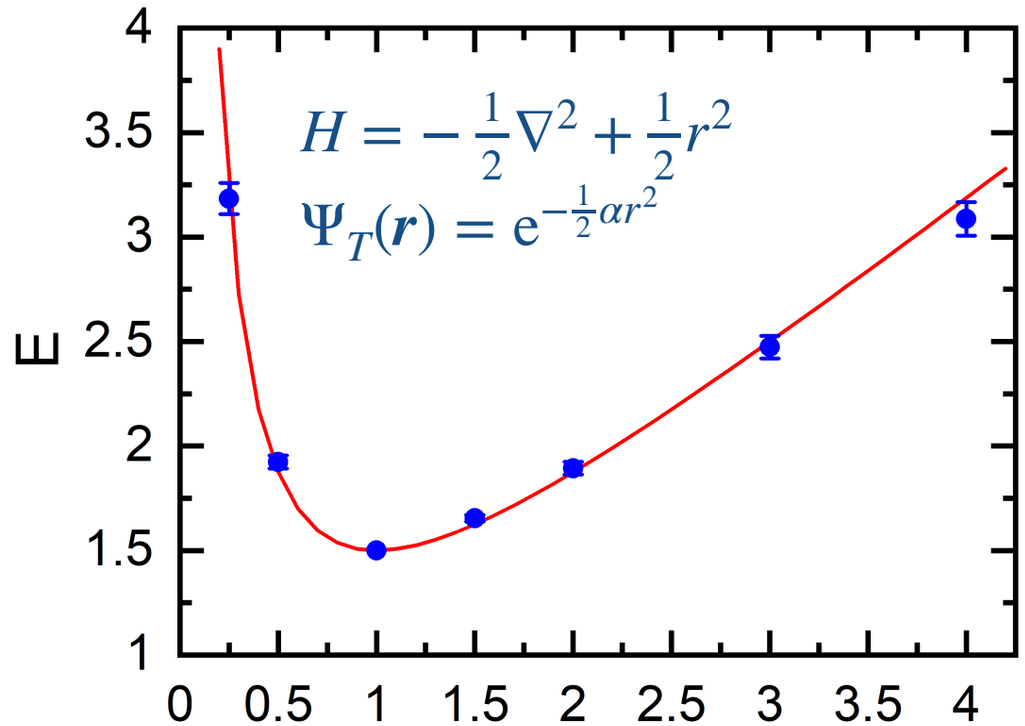
QMC methods

- Variational Monte Carlo (VMC)

  Accuracy depending on the ansatz

$$\min_{\Psi_T} E[\Psi_T] \geq E_0$$

$$E_T = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \int \mathrm{d}R \, |\Psi_T(R)|^2 E(R)$$

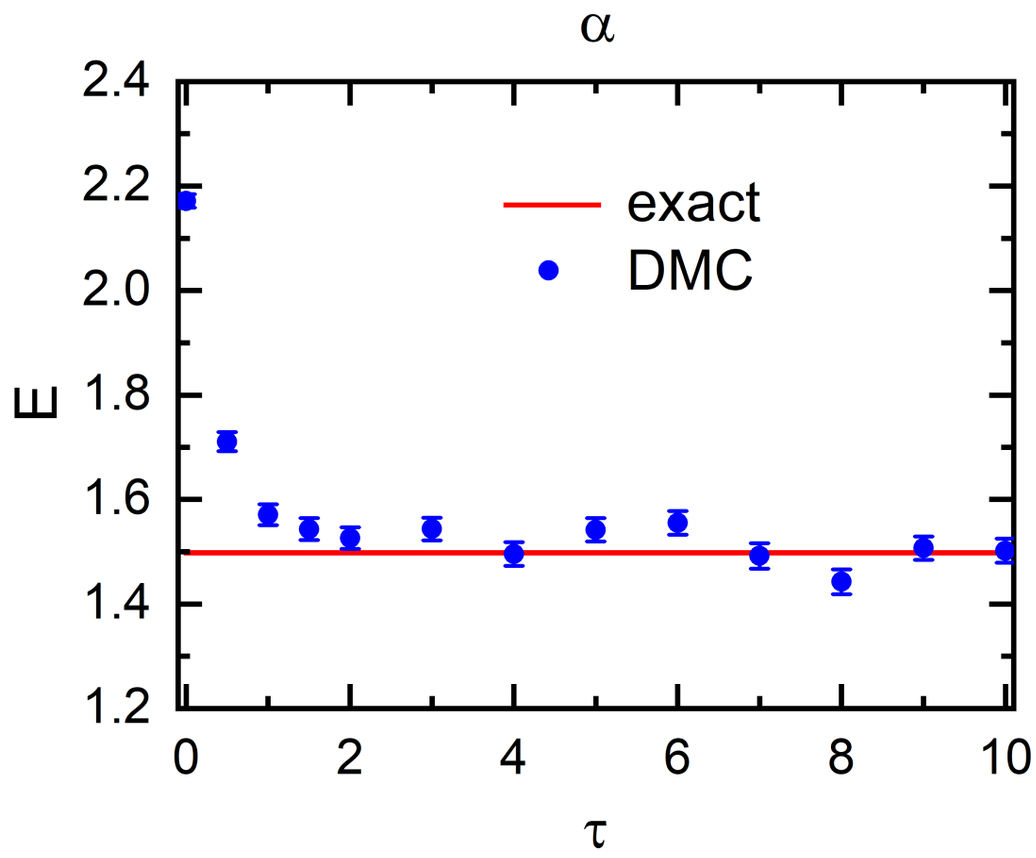


- Diffusion Monte Carlo (DMC)

  Accuracy depending on the sign problem

$$|\Psi(\tau)\rangle = \mathrm{e}^{-\hat{H}\tau} |\Psi_T\rangle \propto |\Psi_0\rangle + \mathrm{e}^{-E_1^*\tau} |\Psi_1\rangle + \cdots,$$

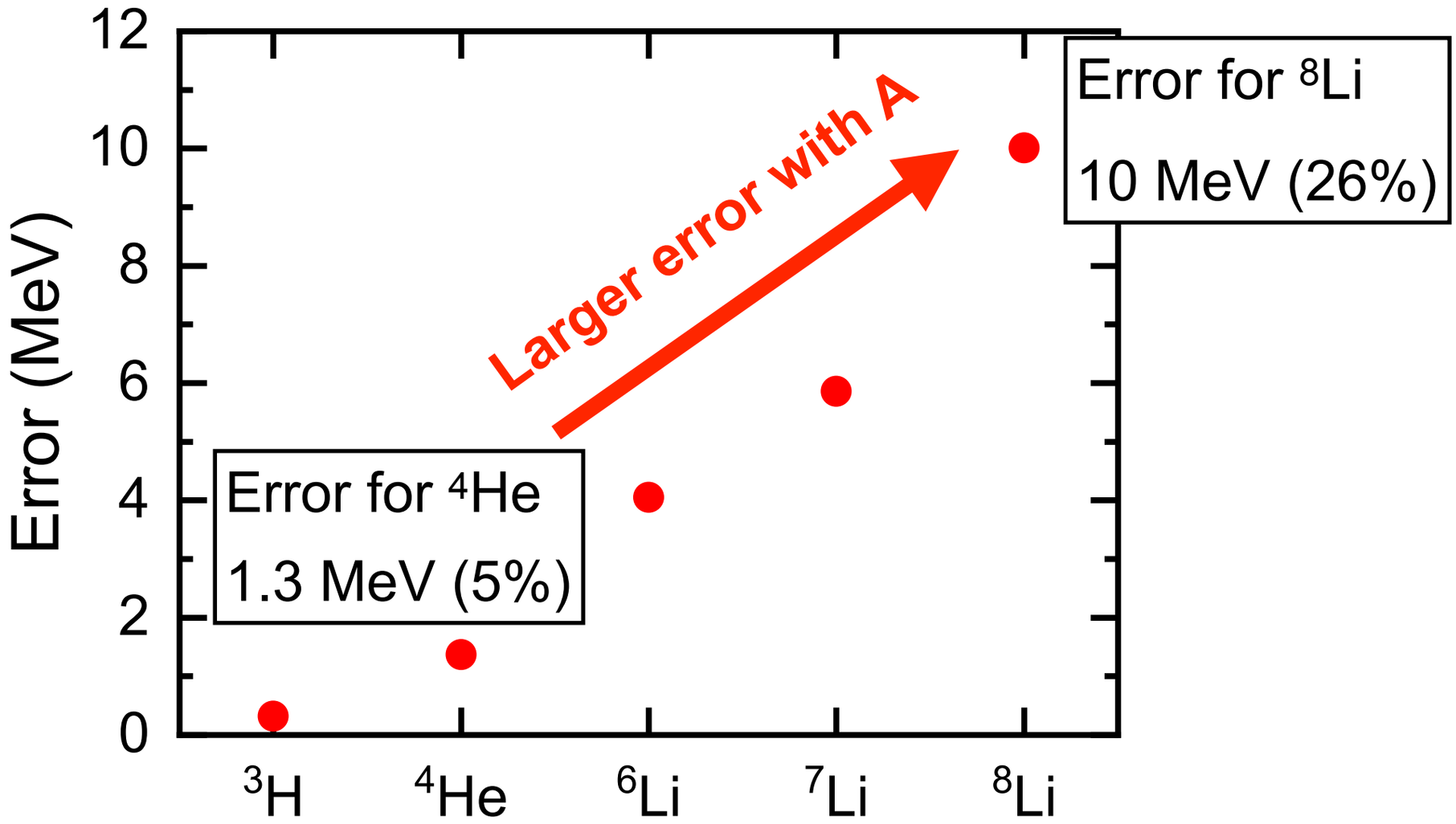
$$E_0 = \lim_{\tau\to\infty} \frac{\langle \Psi_T | H | \Psi(\tau)\rangle}{\langle \Psi_T | \Psi(\tau)\rangle},$$



“Conventional” trial wave function

- “Conventional” trial wave function cannot reach ground states variationally and its quality deteriorate rapidly with increasing system size.

$$|\Psi_T\rangle = \prod_{i<j} \hat{F}_{ij} |\Phi\rangle, \quad \hat{F}_{ij} = f_c(r_{ij}) + f_\sigma(r_{ij}) \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j + \cdots$$

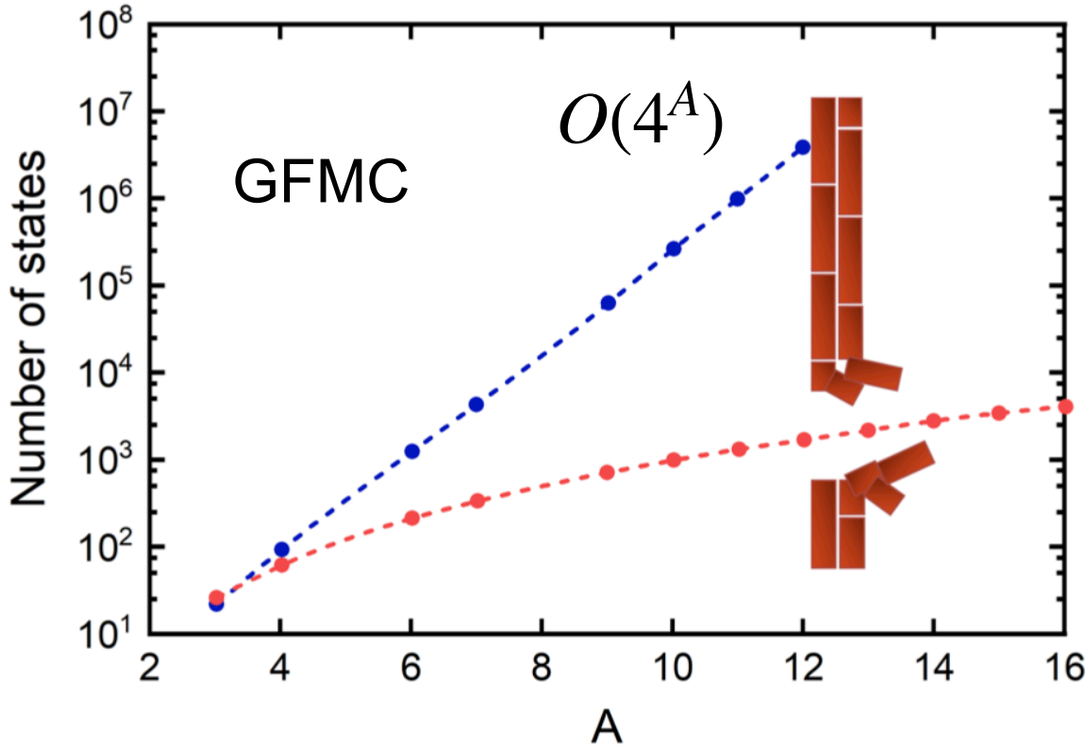


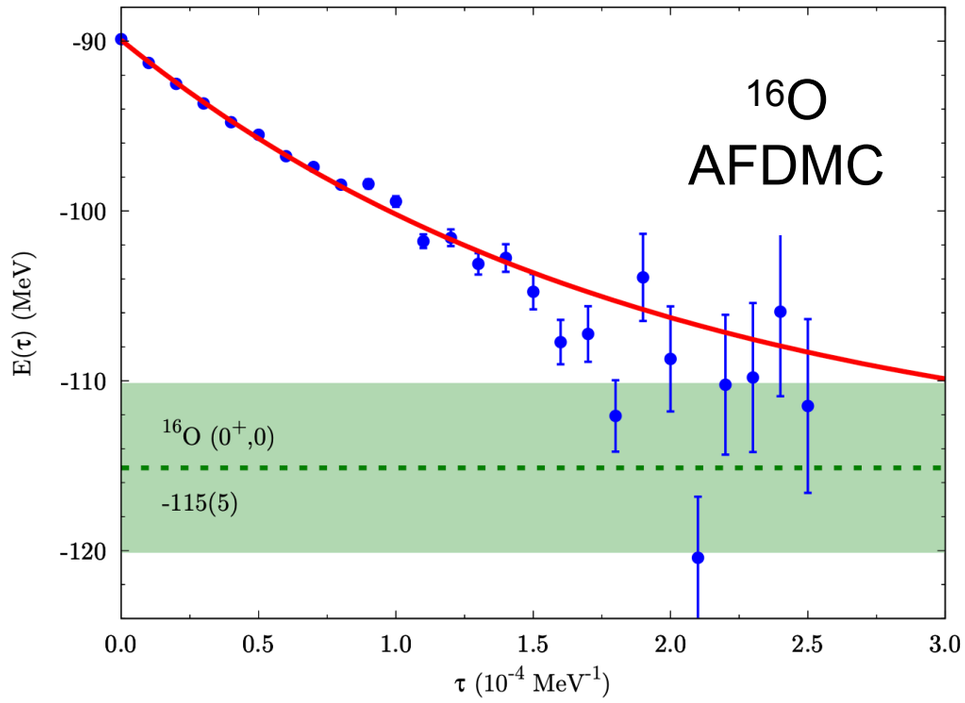
QMC calculations of $A \leq 8$ nuclei: Wiringa et al., PRC 62, 014001 (2000)

Exponential complexity in DMC

Curse of dimensionality: exponential increase of dimensions in spin-isospin space.

Fermion sign problem: an exponential increasing error-to-signal ratio in DMC.





Lonardoni et al., PRC 97, 044318 (2018)

➡ Devising a polynomial scaling and accurate trial wave function

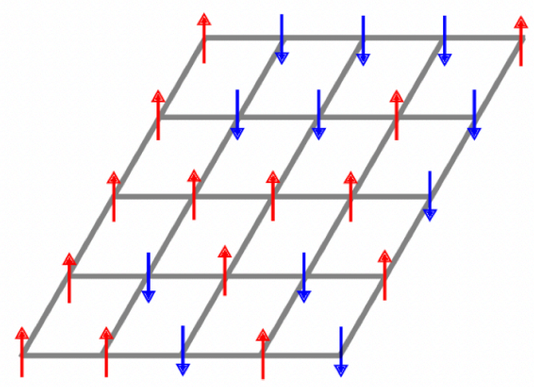
Neural-network Variational Monte Carlo

The quantum many-body problem

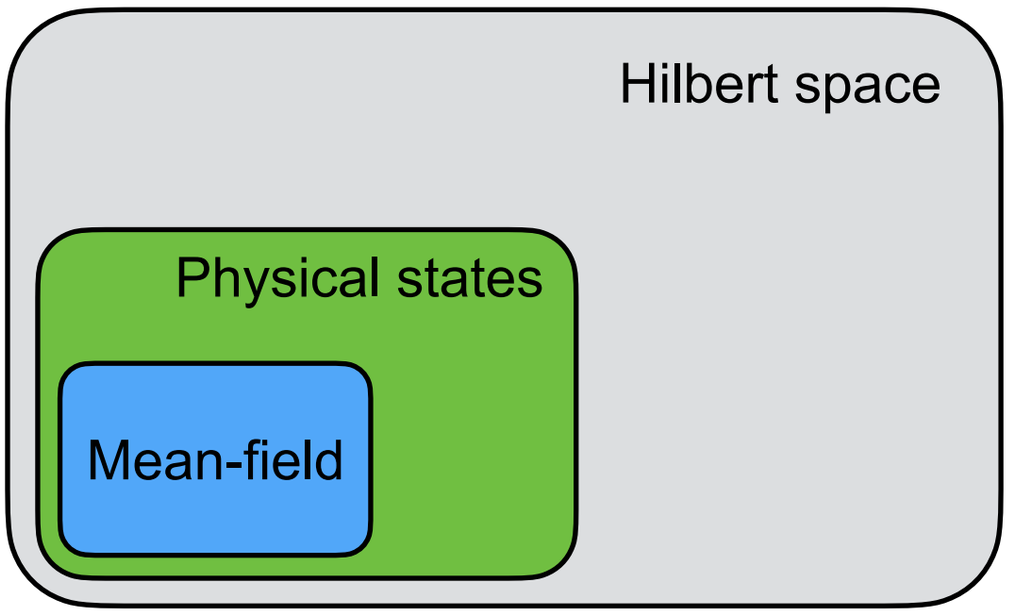
- Finding exact solutions of quantum many-body systems is, in principle, an exponentially hard problem.

$$H_{\mathrm{TFI}} = -h\sum_i \sigma_i^x - \sum_{\langle i,j\rangle} \sigma_i^z \sigma_j^z$$

$$|\Psi\rangle = c_{\uparrow\uparrow\ldots}|\uparrow\uparrow\cdots\rangle + c_{\downarrow\uparrow\ldots}|\downarrow\uparrow\cdots\rangle + \cdots + c_{\downarrow\downarrow\ldots}|\downarrow\downarrow\cdots\rangle$$



- However, the majority quantum states of physical interest have distinctive features and structures, which could allow an efficient polynomial solution



Solving quantum many-body problem with neural networks

RESEARCH ARTICLE

Science

MANY-BODY PHYSICS

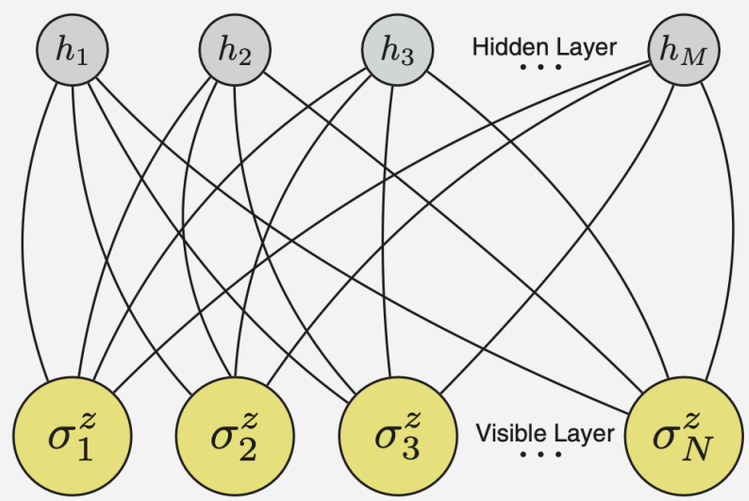
Solving the quantum many-body problem with artificial neural networks

Carleo and Troyer, Science 355, 602 (2017)

A reinforcement-learning scheme we demonstrate is capable of both finding the ground state and describing the unitary time evolution of complex interacting quantum systems.

$$c_{\uparrow\uparrow\ldots}|\uparrow\uparrow\cdots\rangle$$
$$c_{\downarrow\uparrow\ldots}|\downarrow\uparrow\cdots\rangle$$
$$\vdots$$
$$c_{\downarrow\uparrow\ldots}|\downarrow\uparrow\cdots\rangle$$

$$\longleftrightarrow \quad c_S \equiv \Psi(S;\boldsymbol{W})$$



Ab-initio quantum chemistry

FermiNet **PauliNet**

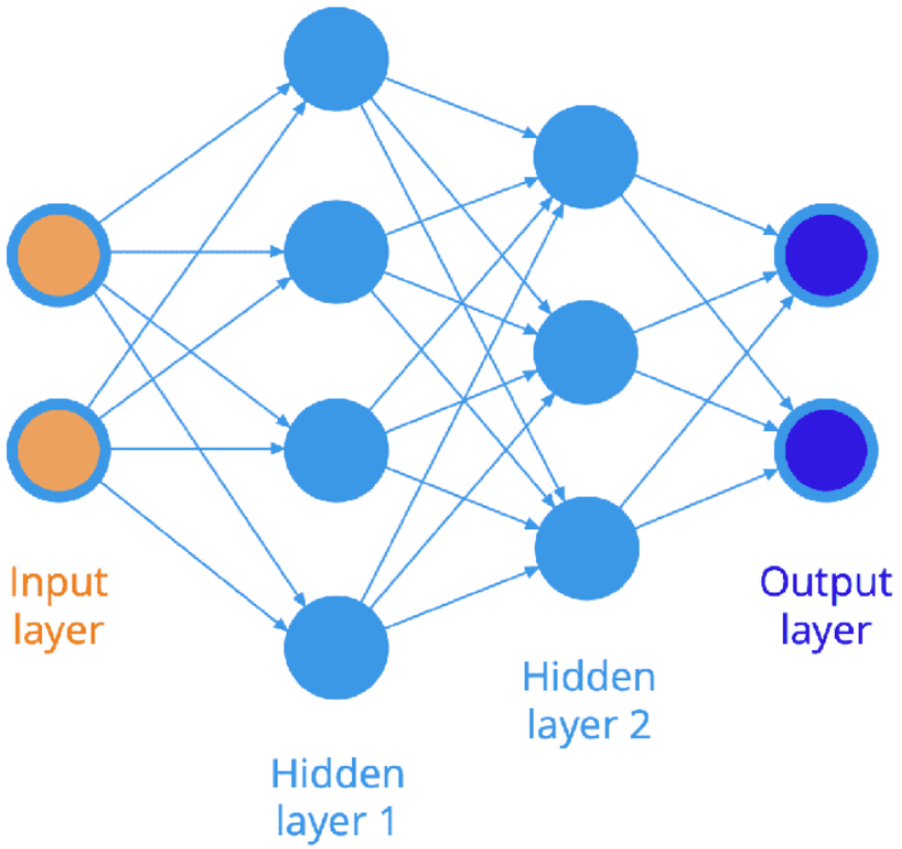
Pfau et al., Phys. Rev. Res. 2, 033429 (2020)
Hermann et al., Nat. Chem. 12, 891 (2020)

Neural networks

- Represents a function from inputs to outputs
- Nested sequence of linear and non-linear functions with variable parameters.

Universal Approximation Theorem: *existence / limit theorem*

a single-hidden-layer neural network can approximate any continuous function



outputs inputs

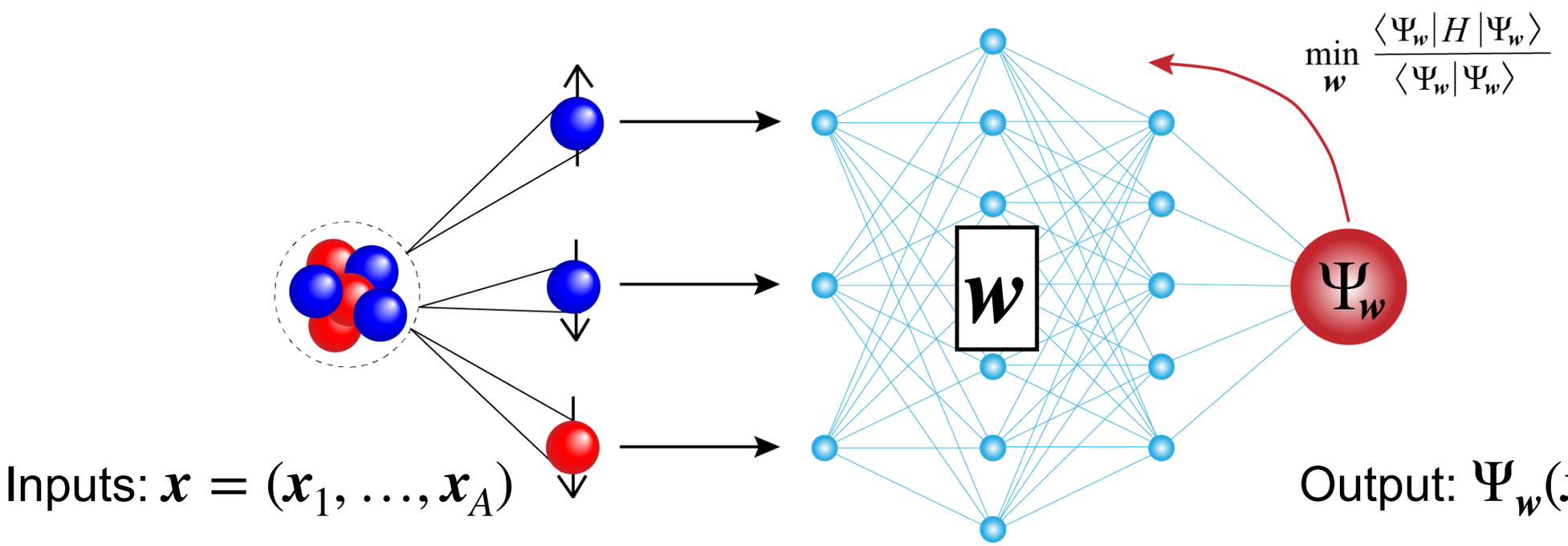
$$y_i = \sigma\left(\sum_{j=1}^{n} x_j \times w_{ij} + b_j\right)$$

w, b: adjustable weights (variational param.)

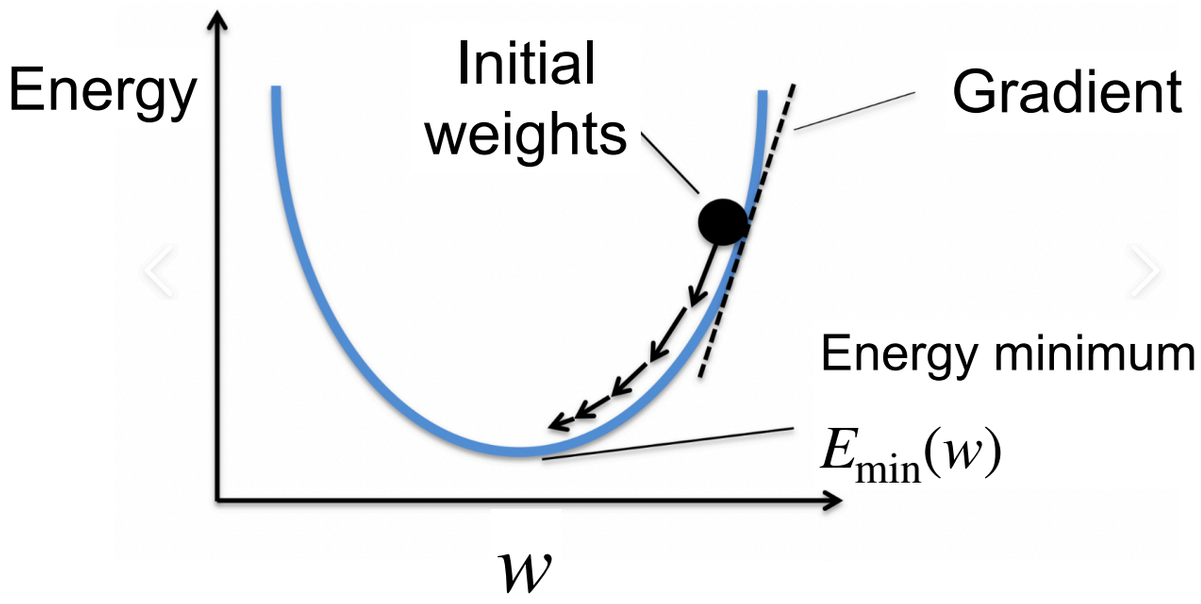
σ: nonlinear functions, e.g. $\tanh(x)$

Variational learning

- Neural networks: efficiently parametrize complicated wave functions



- Variational learning: train neural networks with variational principle (VMC)



Metropolis-Hastings Monte Carlo

- The energy is calculated stochastically on a set of samples

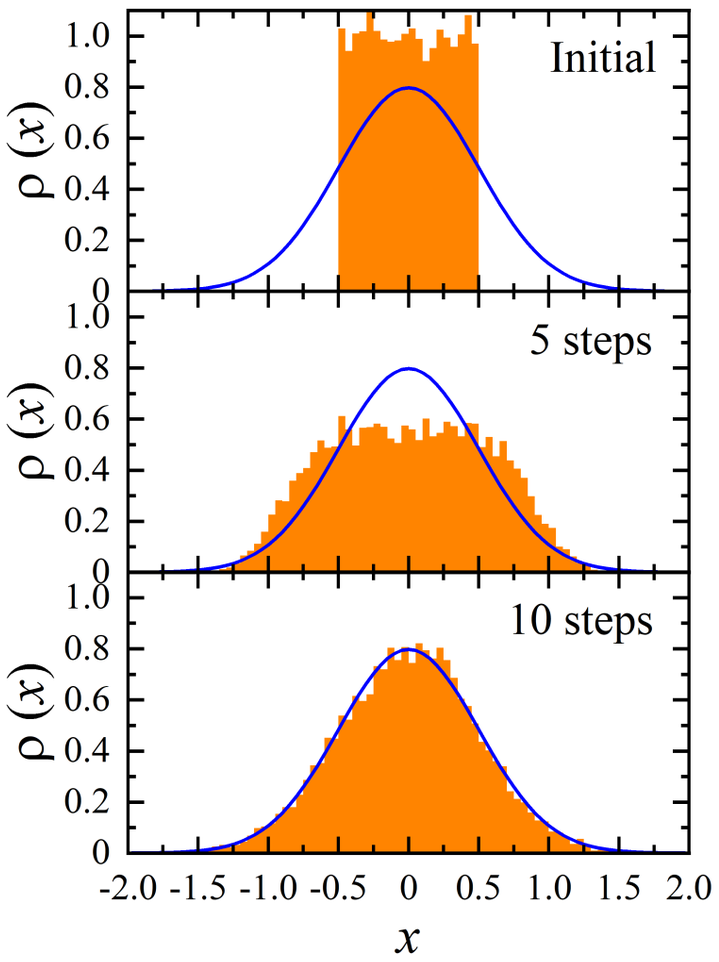
$$\langle O \rangle = \int \mathrm{d}R P(R) O(R) \simeq \frac{1}{N_{\mathrm{samp}}} \sum_{R \sim P(R)} O(R) \qquad P(R) = \frac{|\Psi(R)|^2}{\int \mathrm{d}R\, |\Psi(R)|^2}$$

- The Metropolis-Hastings algorithm perfoms random walk to obtain samples with the desired distribution $P(R)$.

1. For a given sample R_n, randomly propose a new position R_n', e.g., uniformly in $(R_n - d, R_n + d)$

2. Accept R' with the probability

$$P_{\mathrm{accept}}(R_n \to R_n') = \min\left\{1, \frac{P(R_n')}{P(R_n)}\right\}$$



Stochastic reconfiguration

- The neural network is trained iteratively by stochastic reconfiguration

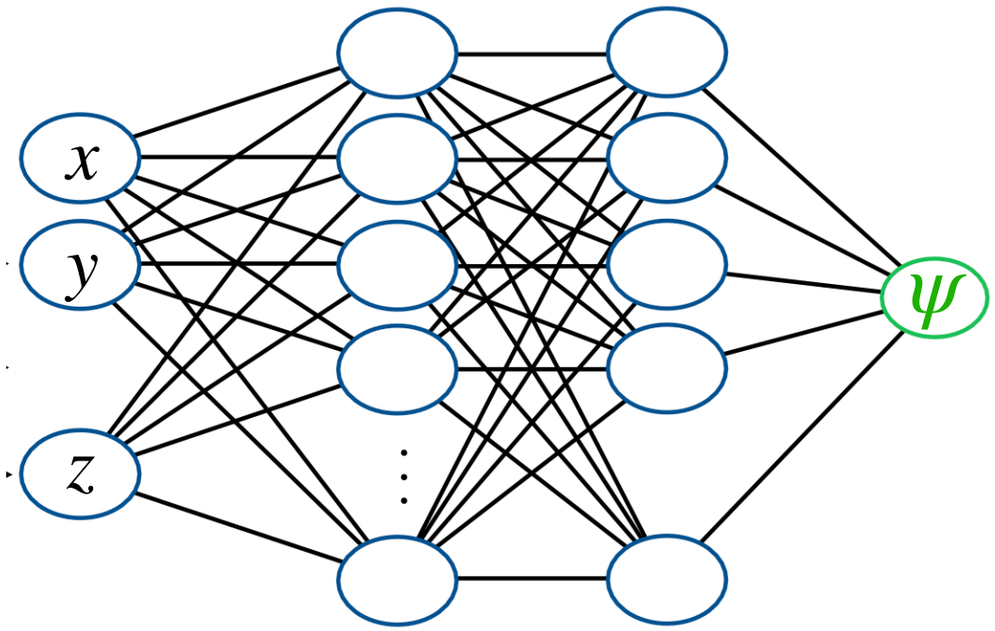
$$\boldsymbol{w}^{t+1} = \boldsymbol{w}^{t} - \gamma(\mathbf{S} + \epsilon I)^{-1}\boldsymbol{g}$$

Sorella, Phys. Rev. B 71, 241103(R) (2005)

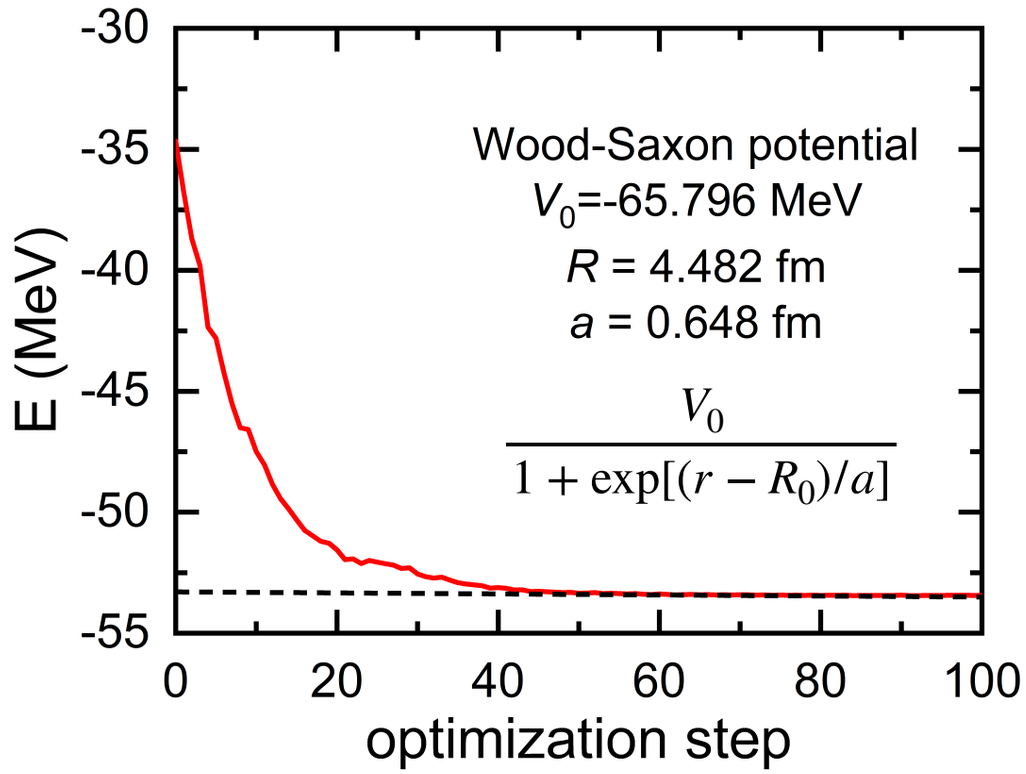
$$g_i = \partial_i E = 2\left(\frac{\langle \partial_i\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} - E\frac{\langle\partial_i\Psi|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right), \quad S_{ij} = \frac{\langle\partial_i\Psi|\partial_j\Psi\rangle}{\langle\Psi|\Psi\rangle} - \frac{\langle\partial_i\Psi|\Psi\rangle}{\langle\Psi|\Psi\rangle}\frac{\langle\Psi|\partial_j\Psi\rangle}{\langle\Psi|\Psi\rangle}$$

- It is equivalent to imaginary-time evolution in the variational manifold

$$|\Psi(w^{t+1})\rangle = |\Psi(w^{t})\rangle + \sum_i \Delta w_i \partial_i |\Psi(w^{t})\rangle \simeq (1 - \gamma H)|\Psi(w^{t})\rangle$$



$N = 16$



Leading-order Hamiltonian

- We consider a nuclear Hamiltonian derived in LO pionless EFT

$$H_{\mathrm{LO}} = \sum_{i=1}^{A} \frac{-\nabla_i^2}{2m_N} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk}$$

- NN potential: fit to S-wave scattering lengths, effective range, and deuteron binding energy

$$v_{ij} = C_{10}\delta_{R_0}(r_{ij})P_{10} + C_{01}\delta_{R_1}(r_{ij})P_{01} + v_{\mathrm{em}}$$

- 3N potential: adjusted to reproduce triton binding energy

$$V_{ijk} = \tilde{c}_E \sum_{\mathrm{cyc}} \delta_{R_3}(r_{ij})\delta_{R_3}(r_{ik})$$

Schiavilla et al., Phys. Rev. C 103, 054003 (2021)

Slater-Jastrow ansatz

- Nuclear many-body wave function must be antisymmetric

$$\Psi(\ldots, x_i, \ldots, x_j, \ldots) = -\Psi(\ldots, x_j, \ldots, x_i, \ldots)$$

- Mean-field wave function: Slater determinant, no correlations

$$\det[\phi(x)] = \begin{vmatrix} \phi_1(x_1), & \phi_1(x_2), & \cdots, & \phi_1(x_A) \\ \phi_2(x_1), & \phi_2(x_2), & \cdots, & \phi_2(x_A) \\ \vdots, & \cdots, & \ddots, & \vdots \\ \phi_A(x_1), & \phi_A(x_2), & \cdots, & \phi_A(x_A) \end{vmatrix} \qquad x_i = (\boldsymbol{r}_i, s_i, t_i)$$

- Neural-network Slater-Jastrow ansatz:

$$\Psi(x_1, \ldots, x_A) = J(x_1, \ldots, x_A) \det \phi(x) \qquad J(x_1, \ldots, x_A) = \rho\left(\sum_{i=1}^{A} \vec{\eta}(x_i)\right)$$

The exchange symmetry of Jastrow is ensured by the summation.

$$J(\ldots, x_i, \ldots, x_j, \ldots) = J(\ldots, x_j, \ldots, x_i, \ldots)$$

Adams et al., PRL 127, 022502 (2021); Deep-Sets: Zaheer et al., arXiv:1703.06114

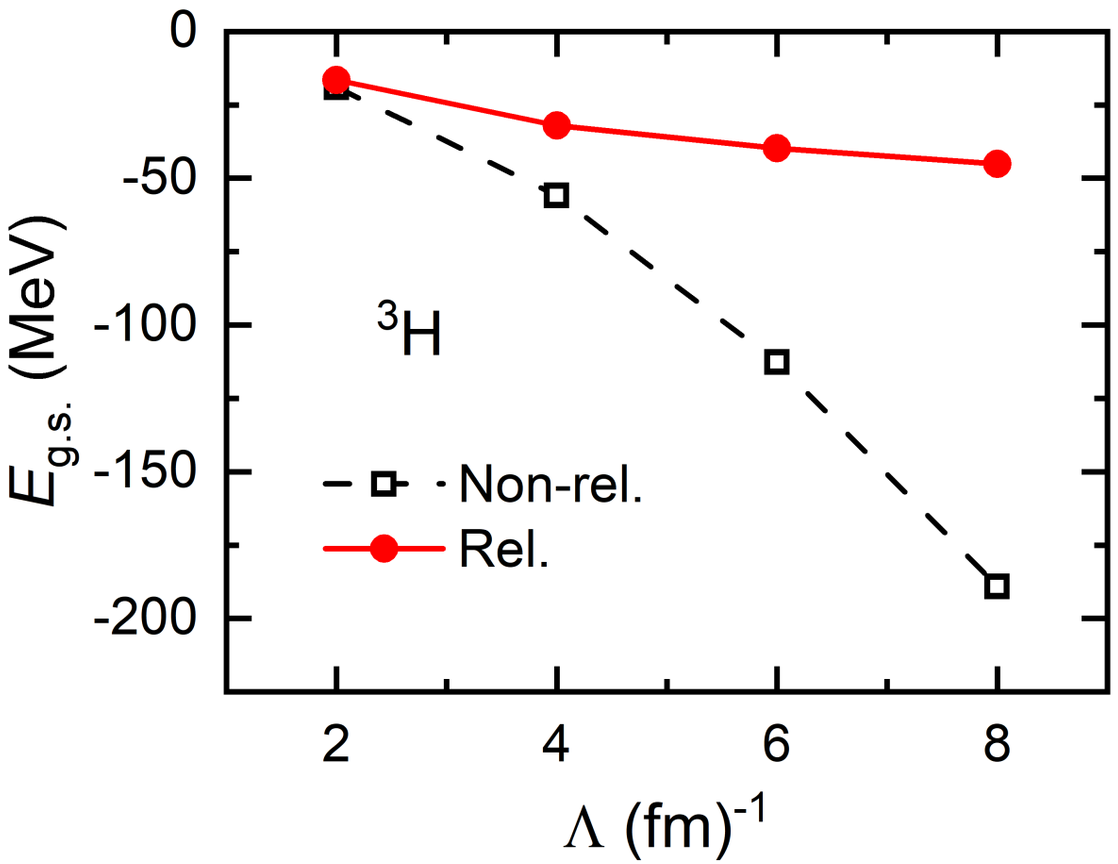
Relativistic effects in few-body systems

- Slater-Jastrow ansatz with rotational symmetry

$$J = \rho\left(\sum_{i<j}\vec{\eta}(|\boldsymbol{r}_i - \boldsymbol{r}_j|)\right)$$

- Relativistic corrections up to quadratic velocity dependence

$$\hat{H} = \sum_{i=1}^{A}\sqrt{\hat{\boldsymbol{p}}_i^2 + M_N^2} + \sum_{i<j}\delta_\Lambda(\boldsymbol{r}_{ij})\left(1 + v_t(\boldsymbol{r}_{ij}, \hat{\boldsymbol{p}}_{ij}^2) + v_b(\boldsymbol{r}_{ij}, \hat{\boldsymbol{P}}_{ij}^2)\right)$$



Thomas collapse avoided.

YLY and Zhao, PLB 835, 137587 (2022)

Backflow transformation

- The Slater-Jastrow ansatz is not “exact”, due to the incorrect nodal surface of the Slater determinant.

Results from Adams et al., PRL 127, 022502 (2021)

	Λ	VMC-ANN	VMC-JS	GFMC
^{2}H	4 fm^{-1}	$-2.224(1)$	$-2.223(1)$	$-2.224(1)$
	6 fm^{-1}	$-2.224(4)$	$-2.220(1)$	$-2.225(1)$
^{3}H	4 fm^{-1}	$-8.26(1)$	$-7.80(1)$	$-8.38(2)$
	6 fm^{-1}	$-8.27(1)$	$-7.74(1)$	$-8.38(2)$
^{4}He	4 fm^{-1}	$-23.30(2)$	$-22.54(1)$	$-23.62(3)$
	6 fm^{-1}	$-24.47(3)$	$-23.44(2)$	$-25.06(3)$

- Including many-body correlations directly in Slater determinant

“Hidden nucleons”

$$\begin{vmatrix} \boldsymbol{\phi}_{A\times A}(x), \boldsymbol{\phi}_{A\times A_h}(x_h) \\ \boldsymbol{\chi}_{A_h\times A}(x), \boldsymbol{\chi}_{A_h\times A_h}(x_h) \end{vmatrix}$$

Lovato et al., Phys. Rev. Research 4, 043178 (2022)

“Backflow transformation”

$$\begin{vmatrix} f_1(x_1;\{x_{1/}\}), & \cdots, & f_1(x_A;\{x_{A/}\}) \\ f_2(x_1;\{x_{1/}\}), & \cdots, & f_2(x_A;\{x_{A/}\}) \\ \vdots, & \ddots, & \vdots \\ f_A(x_1;\{x_{1/}\}), & \cdots, & f_A(x_A;\{x_{A/}\}) \end{vmatrix}$$

YLY and Zhao, PRC 107, 034320 (2023)

FeynmanNet

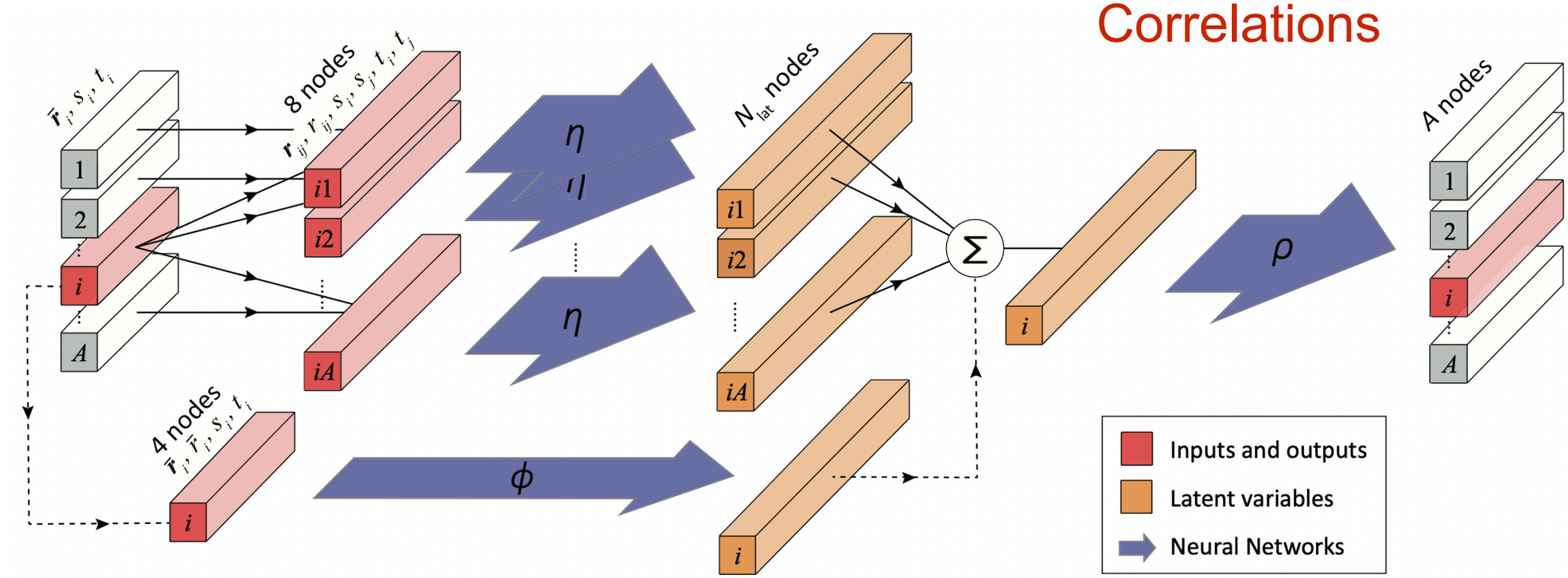
- We introduce spin-isospin dependent backflow with neural networks

YLY and Zhao, PRC 107, 034320 (2023)

$$\det[\Phi_{\mathrm{BF}}] = \begin{vmatrix} f_1(x_1;\{x_{1/}\}), & \cdots, & f_1(x_A;\{x_{A/}\}) \\ f_2(x_1;\{x_{1/}\}), & \cdots, & f_2(x_A;\{x_{A/}\}) \\ \vdots, & \ddots, & \vdots \\ f_A(x_1;\{x_{1/}\}), & \cdots, & f_A(x_A;\{x_{A/}\}) \end{vmatrix}$$

$$\phi_\mu(x_i) \Rightarrow f_\mu(x_i,\{x_{i/}\}) = \rho_\mu\left(\vec{\phi}(\bar{r}_i,s_i,t_i) + \sum_{j\neq i}\vec{\eta}(r_{ij},r_{ij},s_i,s_j,t_i,t_j)\right)$$

Correlations

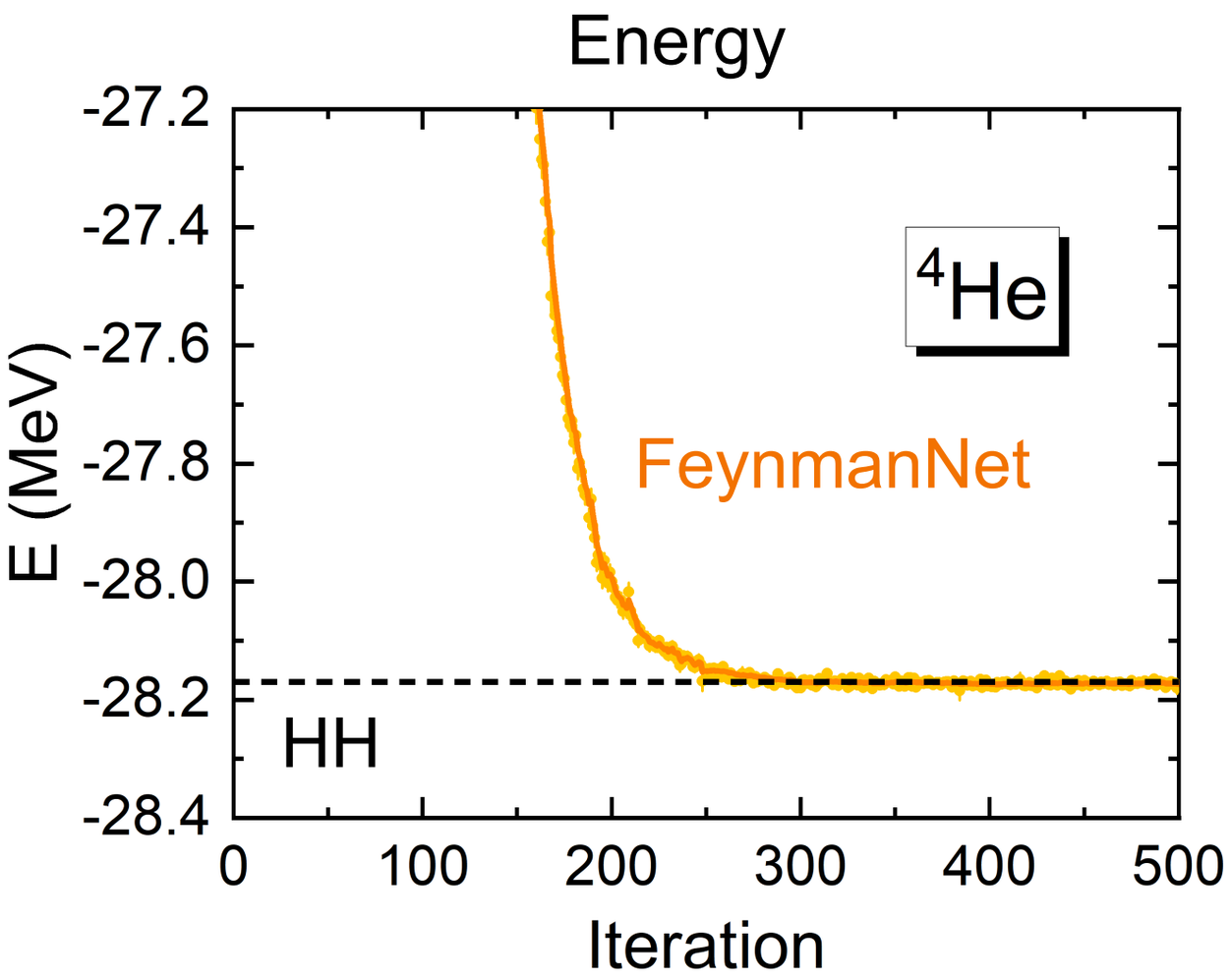


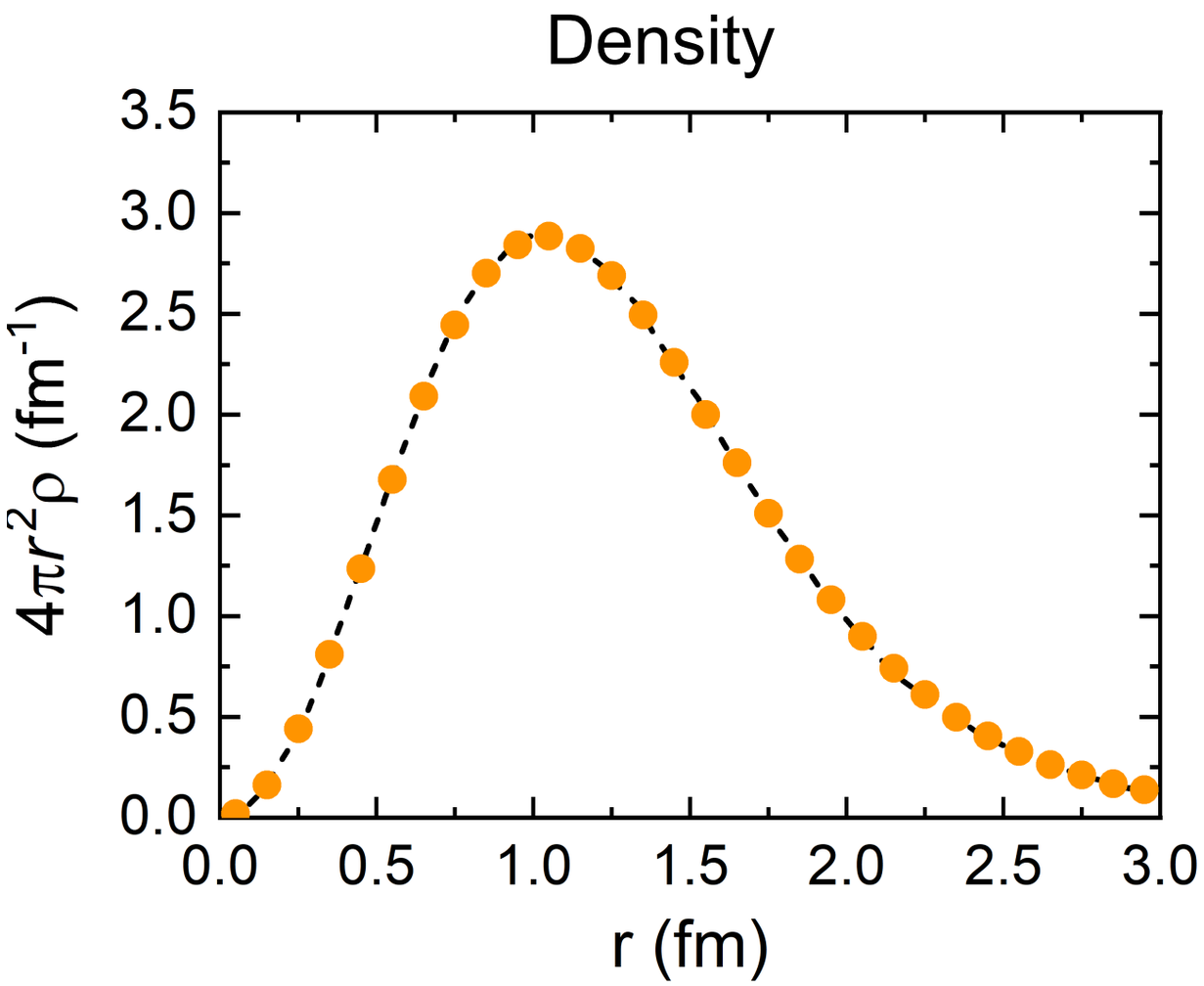
Backflow in ^{3}He liquid: Feynman and Cohen, Phys. Rev. 102, 1189 (1956)

Benchmark for $A = 4$ nuclei

- Perfect agreement with the Hyperspherical Harmonics (HH) method

  HH method ~0.01 MeV accuracy for $A \leq 4$ nuclei



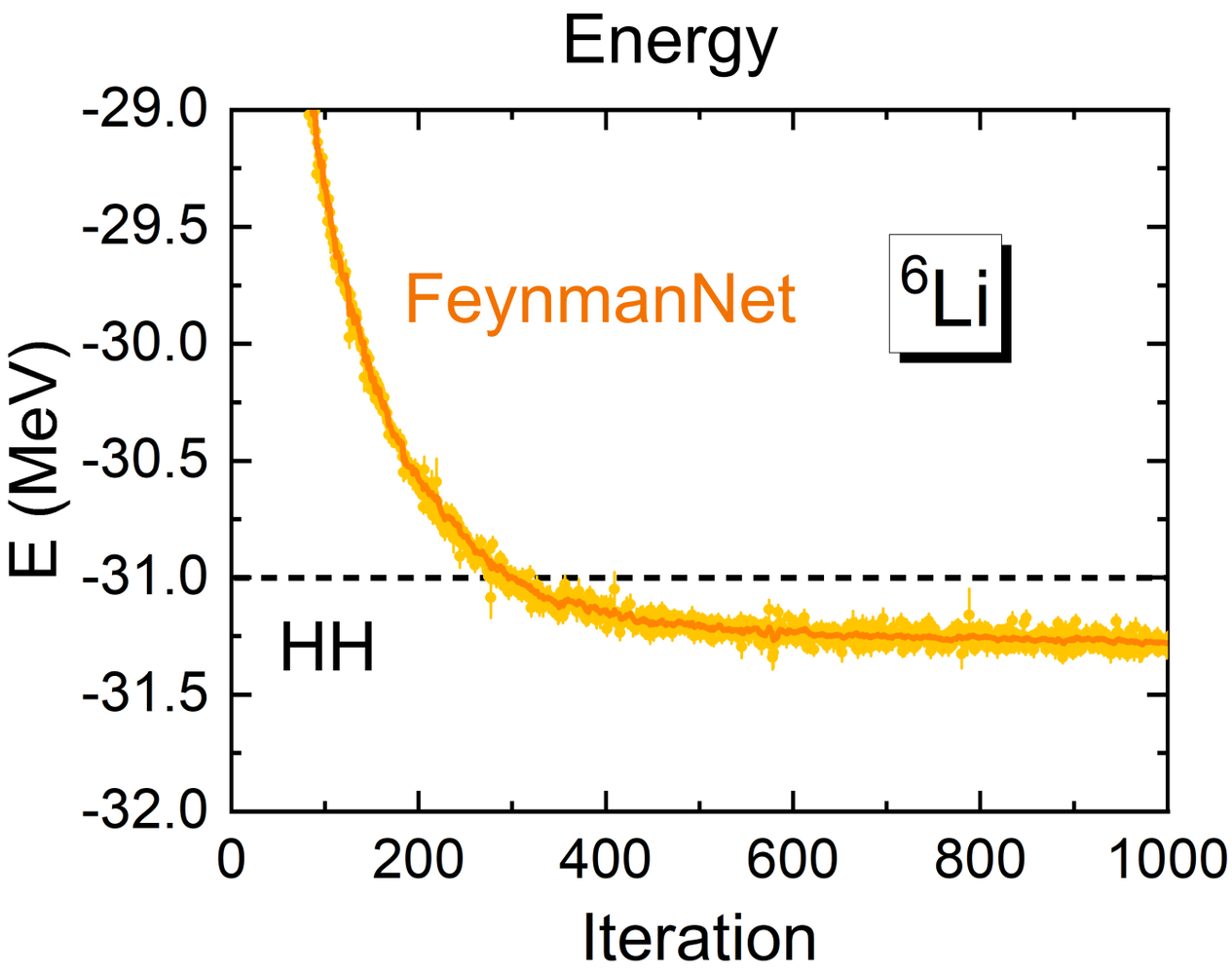


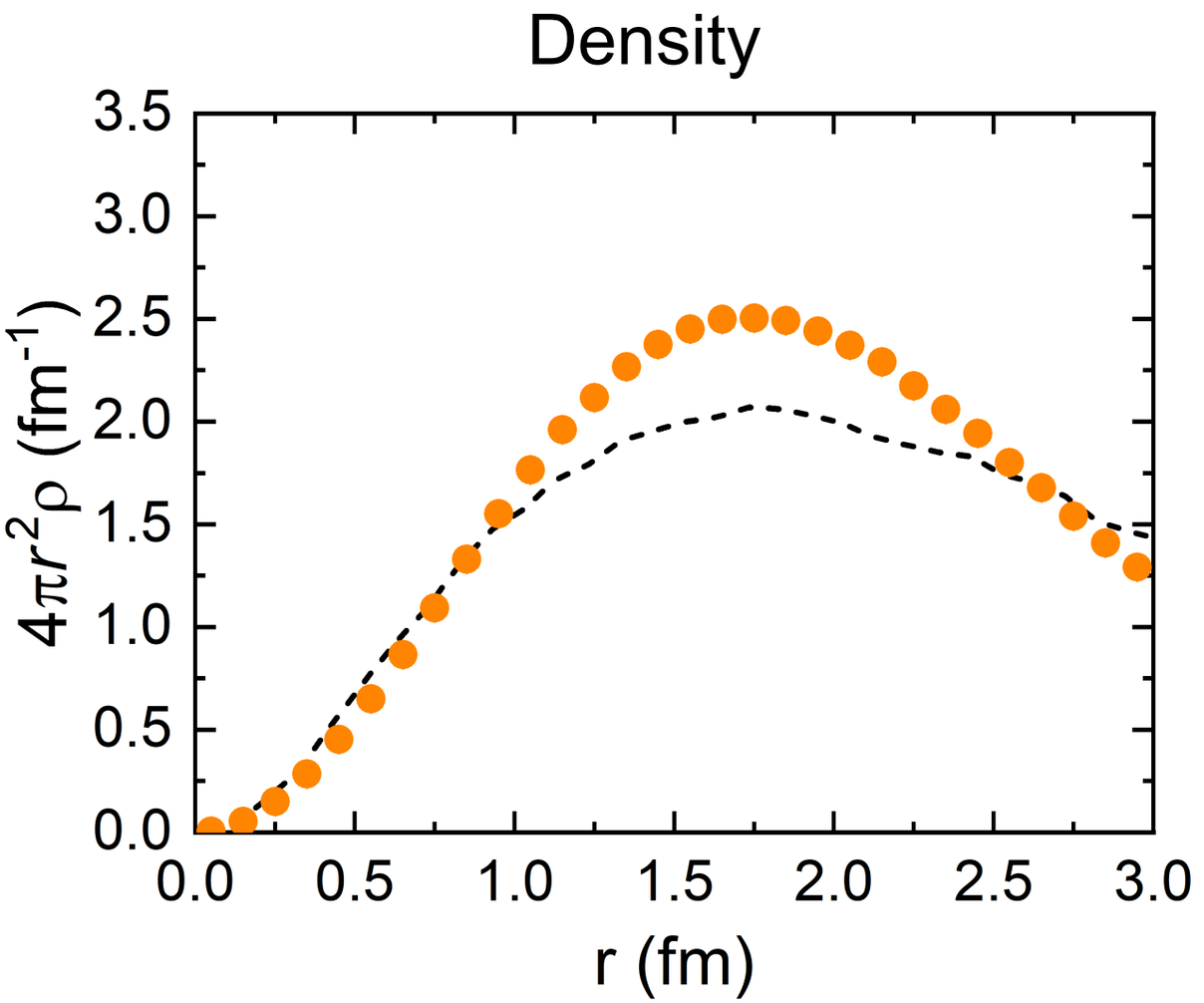
Hyperspherical Harmonics method: Gnech et al., Few-Body Syst. 63, 7 (2022)

Benchmark for $A = 6$ nuclei

- Better than the HH method; Lower the energy by 0.3 MeV.

  HH basis convergence is harder to reach for $A = 6$ nuclei

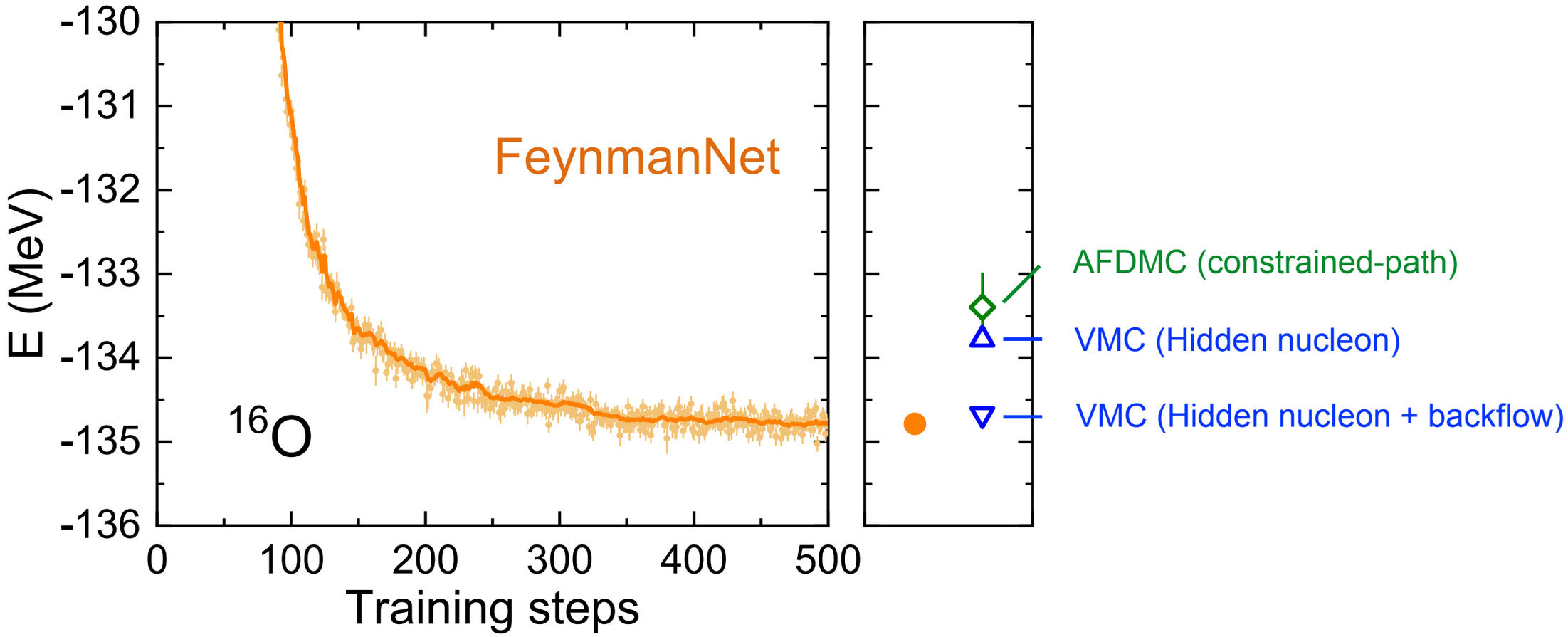




Hyperspherical Harmonics method: Gnech et al., Few-Body Syst. 63, 7 (2022)

Benchmark for $A = 16$ nuclei

- FeynmanNet provides the lowest energy among the QMC methods



VMC (FeynmanNet): YLY and Zhao, PRC 107, 034320 (2023)

VMC (Hidden nucleon): *Lovato et al., Phys. Rev. Research 127, 022502 (2022)*

VMC (Hidden nucleon + backflow): *Gnech et al., PRL 133, 142501 (2024)*

AFDMC (constrained-path): *Schiavilla et al., PRC 103, 054003 (2021)*

Towards solving excited states

PHYSICAL REVIEW C **107**, 034320 (2023)

Deep-neural-network approach to solving the *ab initio* nuclear structure problem

Y. L. Yang and P. W. Zhao*
State Key Laboratory of Nuclear Physics and Technology, School of Physics, Peking University, Beijing 100871, China

RESEARCH ARTICLE

QUANTUM CHEMISTRY

Accurate computation of quantum excited states with neural networks

David Pfau[1,2]*, Simon Axelrod[1,3,4], Halvard Sutterud[2], Ingrid von Glehn[1], James S. Spencer[1]

We present an algorithm to estimate the excited states of a quantum system by variational Monte Carlo, which has no free parameters and requires no orthogonalization of the states, instead transforming the problem into that of finding the ground state of an expanded system. Arbitrary observables can be calculated, including off-diagonal expectations, such as the transition dipole moment. The method works particularly well with neural network ansätze, and by combining this method with the FermiNet and Psiformer ansätze, we can accurately recover excitation energies and oscillator strengths on a range of molecules. We achieve accurate vertical excitation energies on benzene-scale molecules, including challenging double excitations. Beyond the examples presented in this work, we expect that this technique will be of interest for atomic, nuclear, and condensed matter physics.

Pfau et al., Science 385, 6711 (2024)

However, in recent years, advances in deep neural networks have led to their use as accurate ansätze for studying spin systems (6), electronic structure (7) and **nuclear systems (70), often reaching levels of accuracy rivaling** projector QMC methods. These advances have **led to a renewed interest** in VMC as a standalone method.

Neural network ansätze have already been applied to **ground state calculations in some of these domains (70, 71). We are excited to see how** NES-VMC and deep neural networks can be applied to many of the most challenging open problems in many-body quantum mechanics in the future.

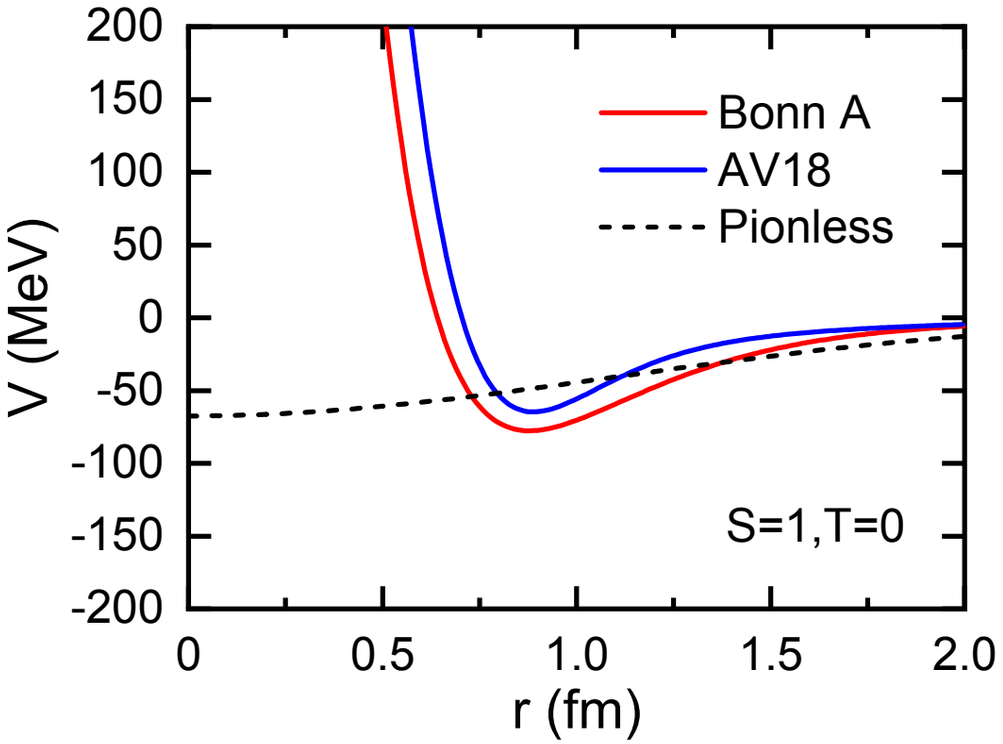
Ref. (70): YLY and Zhao, PRC 107, 034320 (2023)

Towards realistic forces

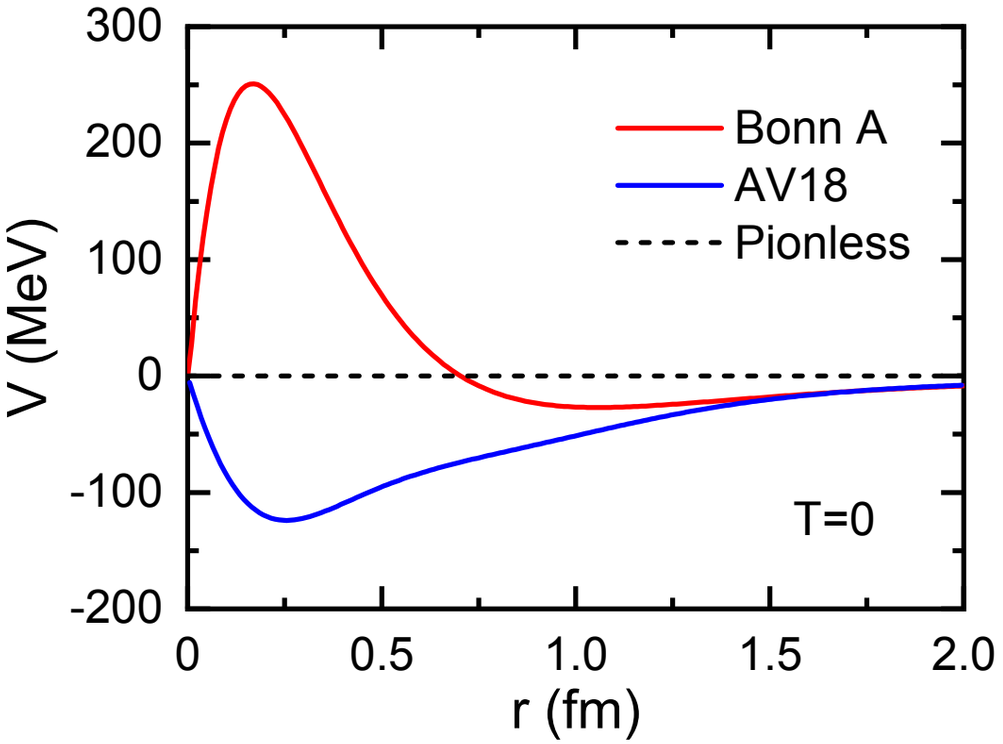
- Realistic nuclear forces are often characterized by strong repulsive core and tensor components

$$v_{ij} = \sum_p v_p(r_{ij}) O_{ij}^p \qquad O_{ij}^p \in \{1,\ \boldsymbol{\tau}_i \cdot \boldsymbol{\tau}_j,\ \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j,\ \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \boldsymbol{\tau}_i \cdot \boldsymbol{\tau}_j,\ S_{ij},\ S_{ij} \boldsymbol{\tau}_i \cdot \boldsymbol{\tau}_j, \ldots\}.$$

Central

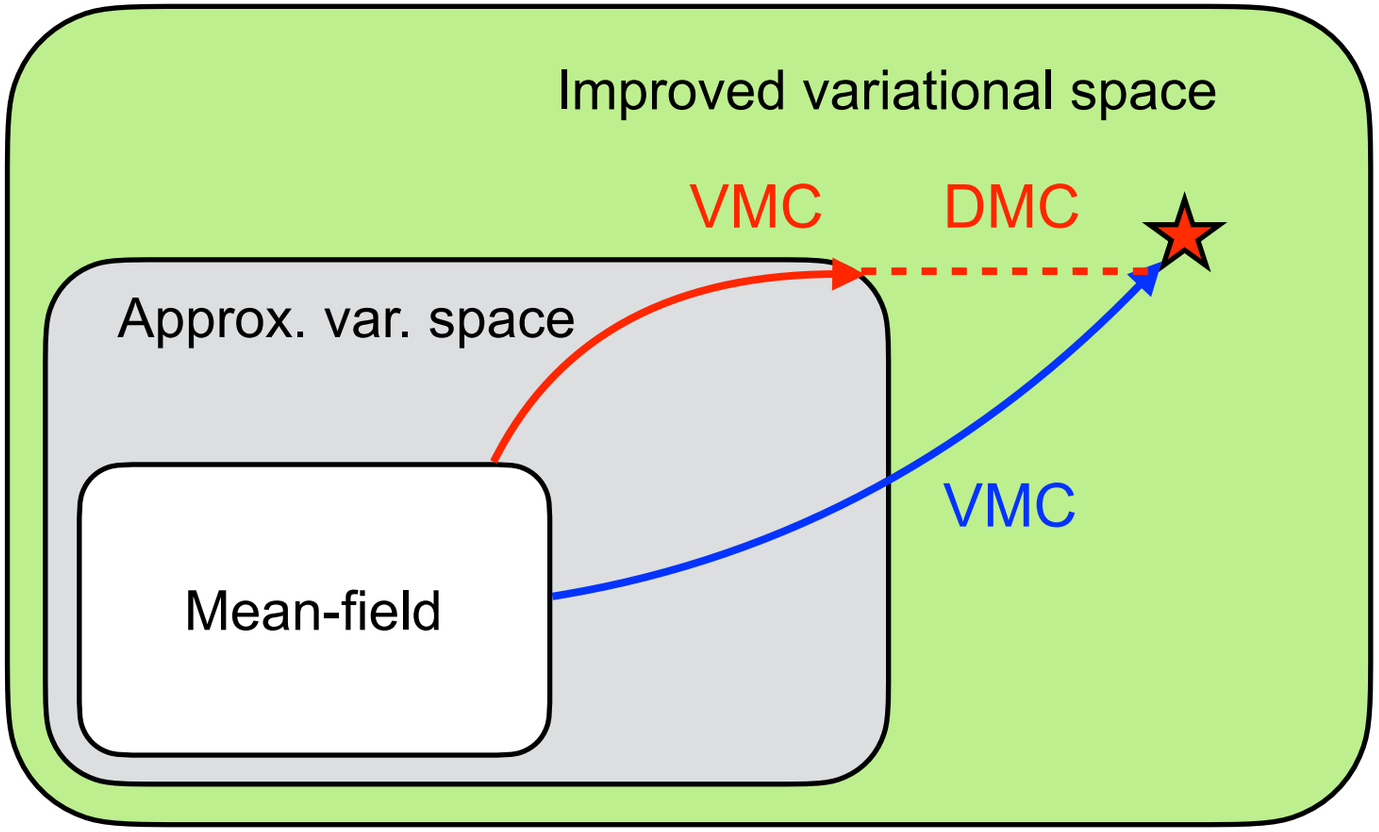


Tensor



Strategies for solving realistic forces

- Strategy 1: obtain a (good) approximate trial state from VMC calculations and perform subsequent DMC calculations to project out the true ground state.
- Strategy 2: including phsycial features and structures to improve the trial ansatz and perform VMC calculations to approach the true ground state



Neural-network ansatz for DMC calculations

- Neural-network Slater-Jastrow ansatz suitable for DMC caluclations: tensor correlations are built explicitly in the Jastrow correlator

YLY and Zhao, arXiv.2405.04203 (2024)

$$|\Psi_T\rangle = \mathcal{F}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A)\left[1 + \sum_{i<j}\sum_{p=2\text{-}6} \mathcal{U}_{ij}^{p}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A)\hat{O}_{ij}^{p}\right]|\Phi\rangle_{J^\pi},$$

Central correlations | Spin-isospin correlations | Mean-field part

- Spin-isospin correlations

spin-isospin independent backflow

$$\mathcal{U}_{ij}^{p}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A) = \rho_{\mathcal{U}}^{p}\left(\sum_{k\neq i,j}\phi_{\mathcal{U}}(r_{ij}, r_{ik}, r_{jk})\right) \qquad O_{ij}^{p} \in \{1, \boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j, S_{ij}\}\otimes\{1, \boldsymbol{\tau}_i\cdot\boldsymbol{\tau}_j\}.$$

Tensor operator

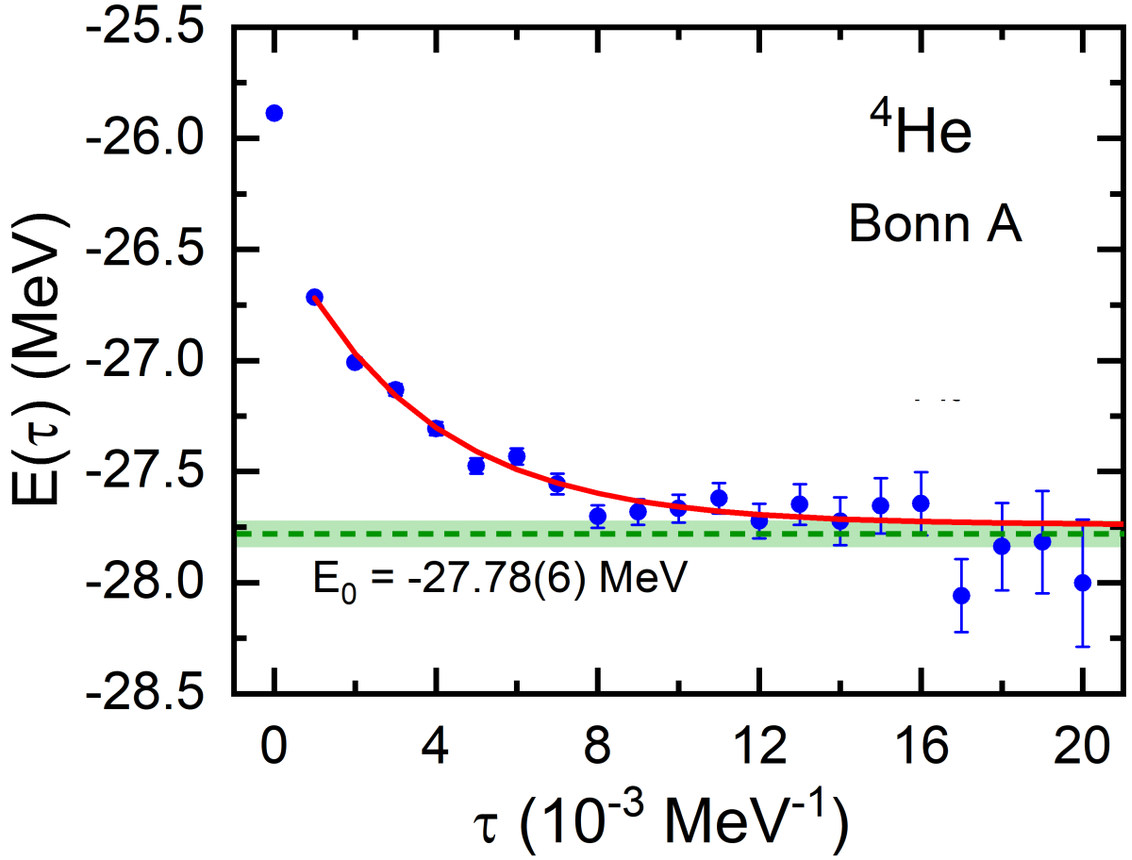
- Central correlations

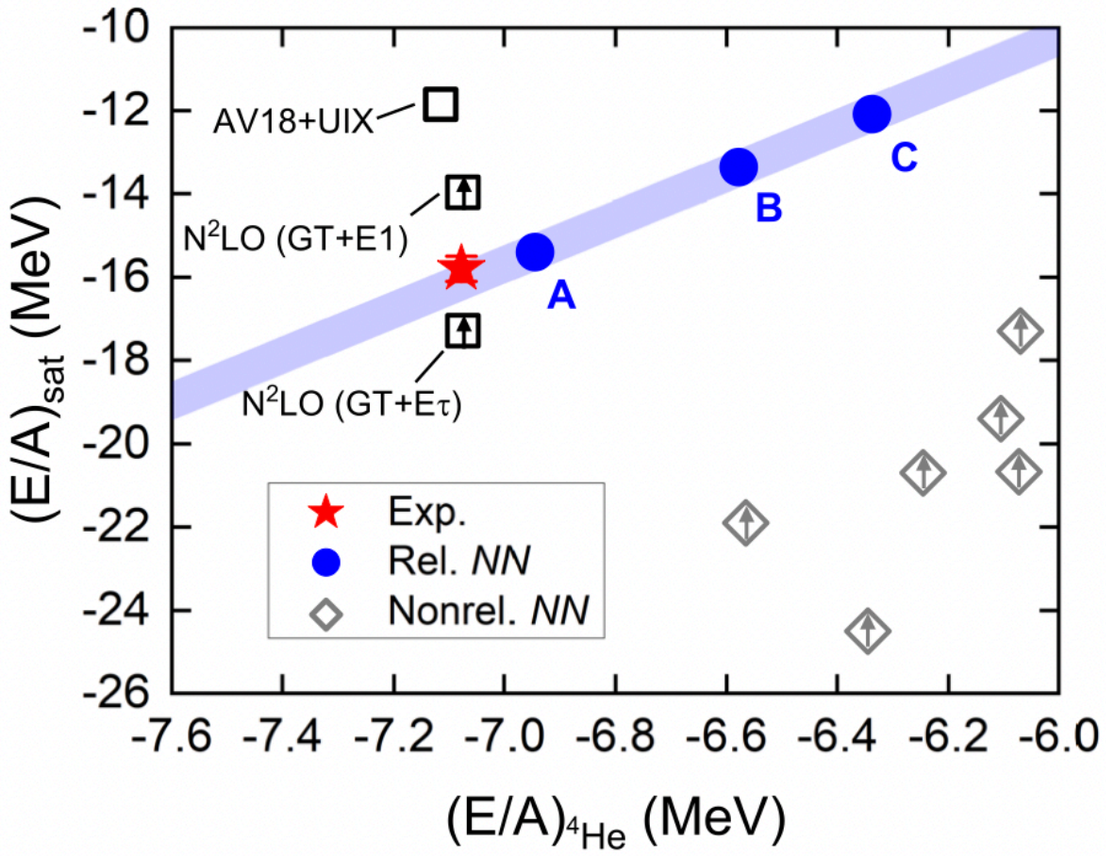
$$\mathcal{F}(\boldsymbol{r}_1, \ldots \boldsymbol{r}_A) = \rho_{\mathcal{F}}\left(\sum_{i<j}\phi_{\mathcal{F}}(r_{ij})\right)$$

$A \leq 4$ nuclei with Bonn force

- VMC calculations provide ~95% of binding energy; DMC calculations account for the rest ~5% of binding energy.
- We find that the Bonn A force can provide good description of light $A \leq 4$ nuclei and nuclear matter saturation simultaneously.

QMC caluclations of $A \leq 4$ nuclei: YLY and Zhao, arXiv:2405.04203 (2024)
RBHF with Bonn forces: Wang, Zhao, Ring, and Meng, PRC 103, 054319 (2021)





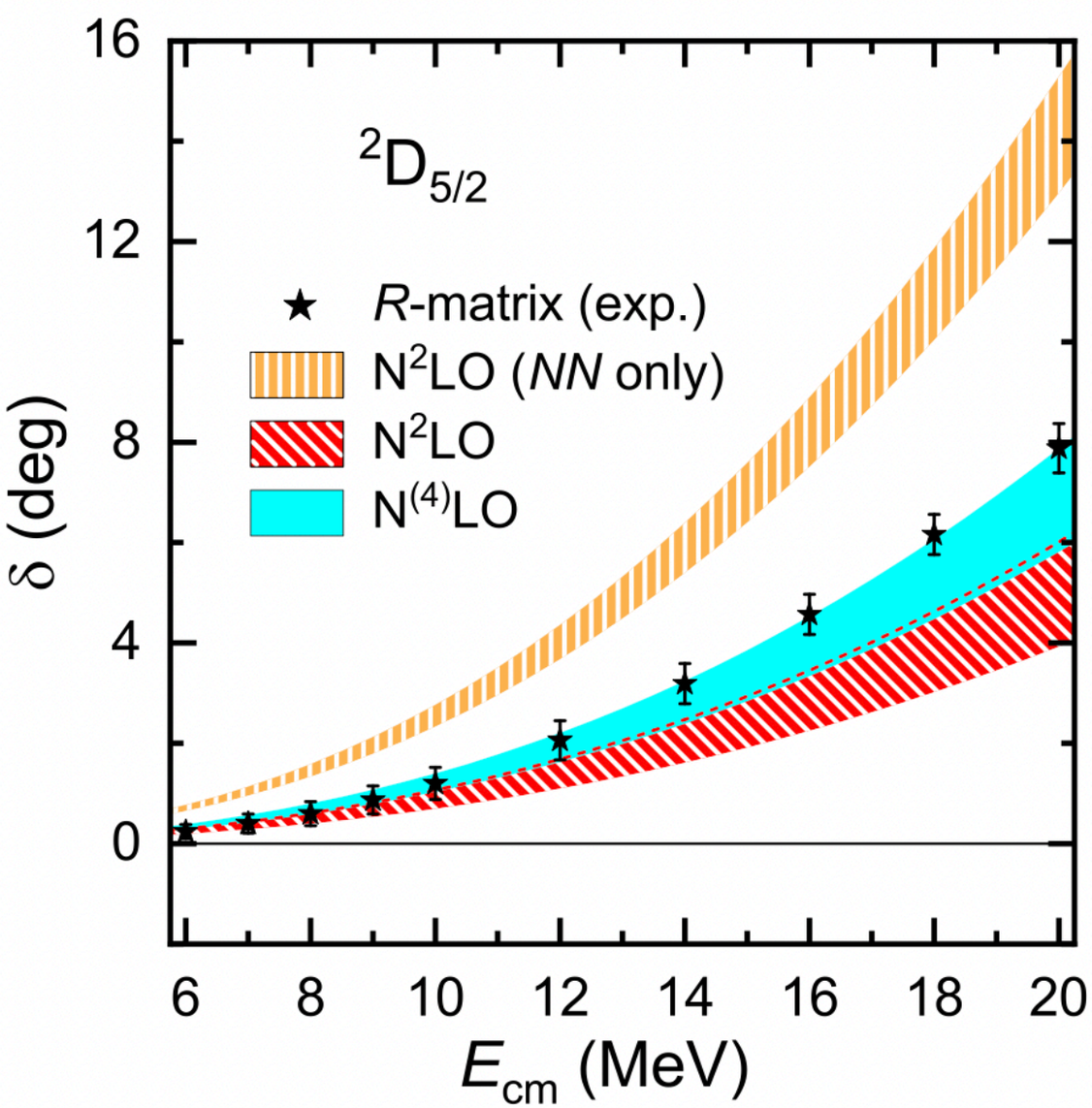
Neutron-α scattering with chiral forces

- We calculate D-wave $n\alpha$ scattering by placing the five-body system in a harmonic oscillator trap, using local chiral forces up to N^{2}LO

YLY, Epelbaum, Meng, Meng, and Zhao, arXiv.2502.09961 (2025)

$$H = T + V + V_{\rm HO}, \quad V_{\rm HO} = \sum_{i<j} \frac{1}{2}\frac{m_N}{A}\omega^2 r_{ij}^2$$

- Peripheral neutron-α scattering is a sensitive probe of long-range 3NF
- The 2π-exchange 3NF is crucial for describing D-wave $n\alpha$ phase shifts



N^{2}LO: $c_{1,3,4}$ from $O(Q^3)$ πN analysis Epelbaum2005NPA

N$^{(4)}$LO: $c_{1,3,4}$ with effective $O(Q^5)$ values Krebs2012PRC

Towards accurate VMC calculations with realistic forces

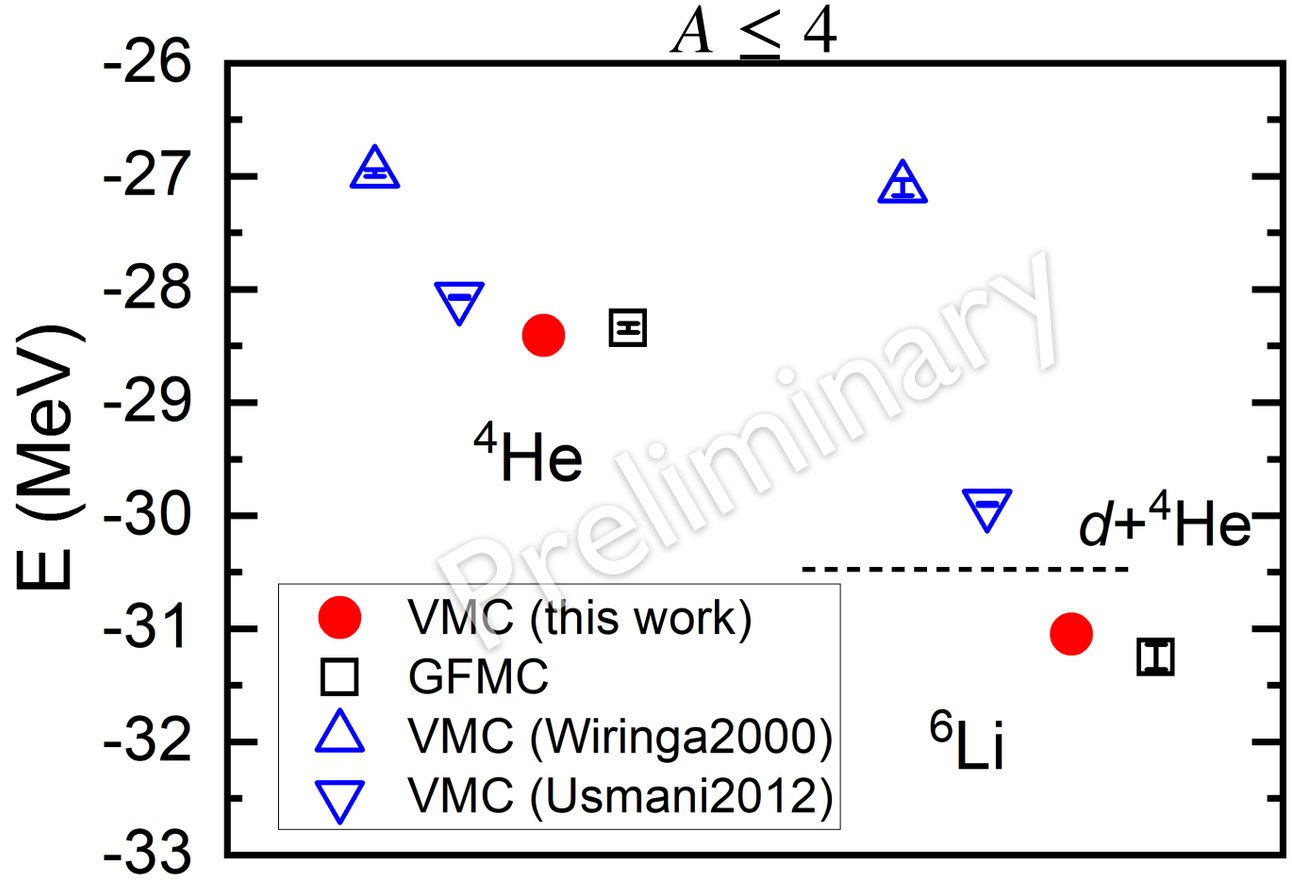
- With novel neural-network ansatz, the VMC energy can, for the first time, reach the GFMC energy for $A \leq 6$ systems with realistic Hamltonians.

YLY and Zhao, in preparation

Results for AV8′ + UIX′

^{4}He: ~ 0.03 MeV (0.1%)

^{6}Li: ~ 0.1 MeV (0.3%)



Outline

- Introduction
- Neural-network Variational Monte Carlo (VMC)
- Neural-network VMC for lattice EFT
- Summary and outlooks

From continuous space to lattice

- We consider the Hamiltonian limit (temporal spacing $\delta = 0$). Then, the ground state can be solved by variational principle.

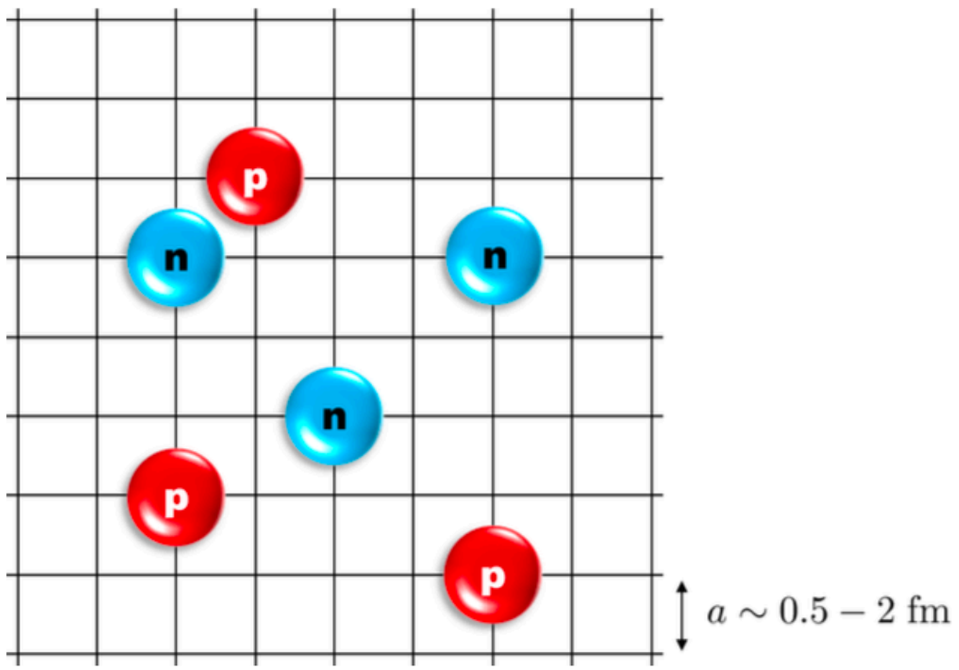
$$H = H_{\text{free}} + V_{2\text{N}} + \dots \qquad E_0 = \min_{\Psi} \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

- To calculate energy expectation, the Metropolis-Hasting sampling is done on discrete lattice, instead of continous space.

$$E = \frac{\sum_s \sum_{\boldsymbol{n}} \Psi(\boldsymbol{n}, s) \hat{H} \Psi(\boldsymbol{n}, s)}{\sum_s \sum_{\boldsymbol{n}} |\Psi(\boldsymbol{n}, s)|^2}$$

$$= \frac{1}{N_{\text{samp}}} \sum_{(\boldsymbol{n},s)\sim|\Psi|^2} \frac{\hat{H}\Psi(\boldsymbol{n}, s)}{\Psi(\boldsymbol{n}, s)}$$

$$\boldsymbol{n} = (\boldsymbol{n}_1, \dots \boldsymbol{n}_N) \in \mathbb{Z}^{3N}$$



Dilute neutron matter

- We study N neutrons in a periodic box with length L, to simulate the dilute neutron matter

$$H = H_{\text{free}} + C\sum_{\boldsymbol{n}} \rho_\uparrow(\boldsymbol{n})\rho_\downarrow(\boldsymbol{n}) + C'\sum_{\boldsymbol{n}}\sum_{l=1,2,3} [\rho_\uparrow(\boldsymbol{n})\rho_\downarrow(\boldsymbol{n}+\hat{l}) + \rho_\downarrow(\boldsymbol{n}+\hat{l})\rho_\uparrow(\boldsymbol{n})]$$

- The LECs are fixed by $a_0 = \infty, r_0 = 0$ in S wave.

Bour, Li, Lee, Meißner, and Mitas, PRA 83, 063319 (2011)

$$C = -3.7235, \quad C_2' = -0.3008$$

- Pairing is important in the dilute neutron matter. So, we use a pfaffian-backflow ansatz

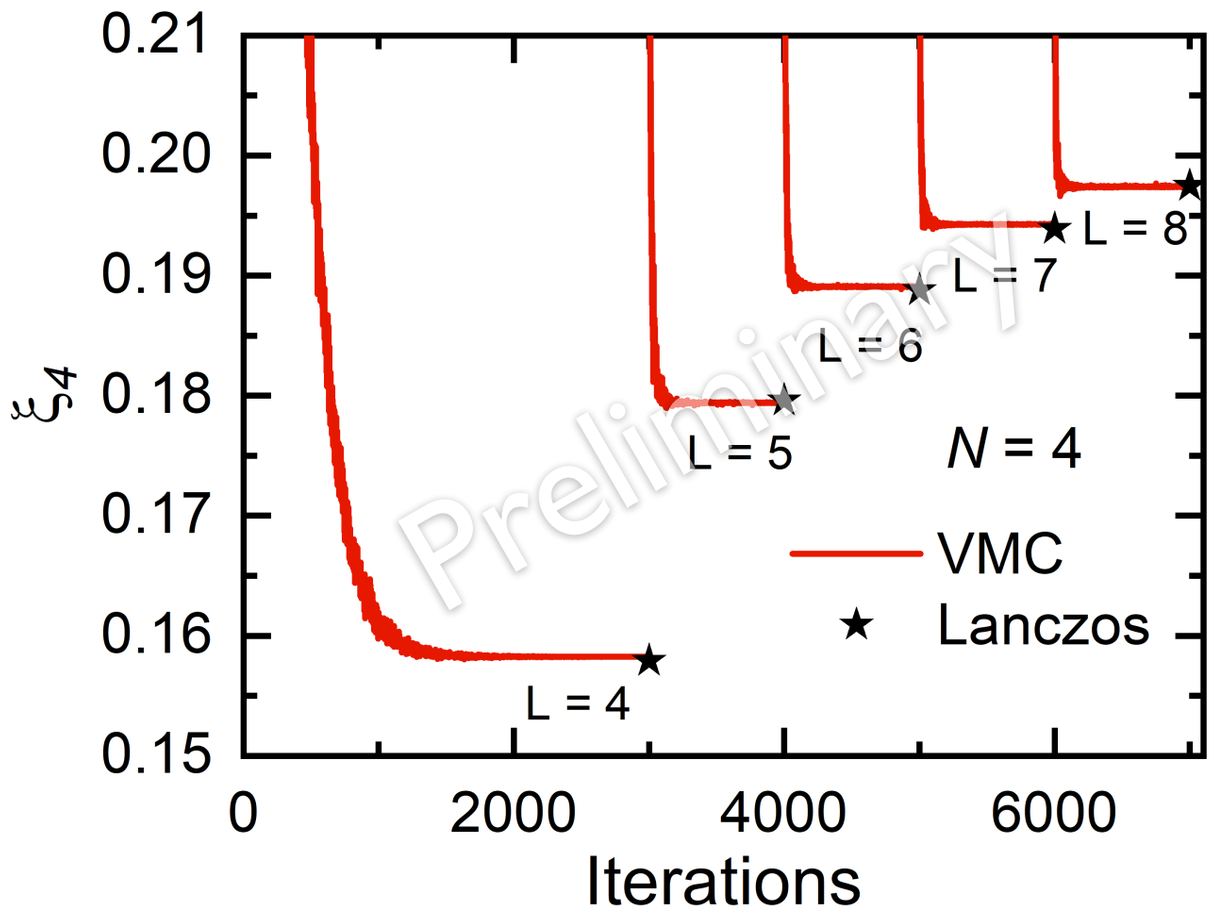
Kim et al., Commun. phys. 7: 148 (2024)

$$\Psi(\boldsymbol{n}_1, s_1, \ldots, \boldsymbol{n}_N, s_N) = \text{pf}\begin{pmatrix} 0 & \phi(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2) & \cdots & \phi(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_N) \\ \phi(\tilde{\boldsymbol{x}}_2, \tilde{\boldsymbol{x}}_1) & 0 & \cdots & \phi(\tilde{\boldsymbol{x}}_2, \tilde{\boldsymbol{x}}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\tilde{\boldsymbol{x}}_N, \tilde{\boldsymbol{x}}_1) & \phi(\tilde{\boldsymbol{x}}_N, \tilde{\boldsymbol{x}}_2) & \cdots & 0 \end{pmatrix}$$

$\boldsymbol{n}_i \in \mathbb{Z}^3, s_i = \pm 1/2$: single-neutron variables $\quad$ $\tilde{\boldsymbol{x}}_i$: backflow-transformed variables

Benchmark for $N = 4$

- Reproduces the exact ground-state energies for $N = 4$ systems
- Transfer learning: The wave function trained at one box size is used as a good starting point for calculations at other box size.



$$\xi_N = E_N / E_N^{\text{free}}$$

Lanczos results from Bour, Li, Lee, Meißner, and Mitas, PRA 83, 063319 (2011)

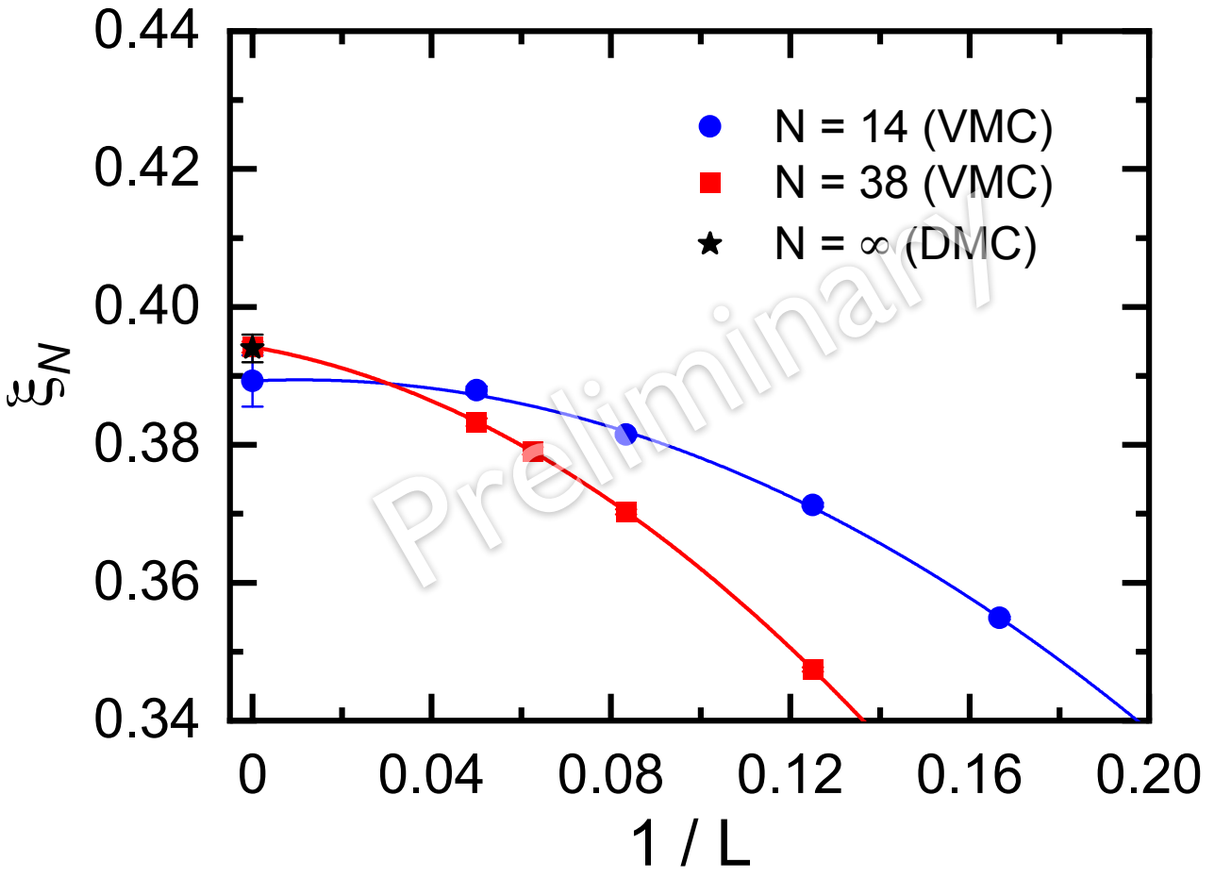
Results for larger systems

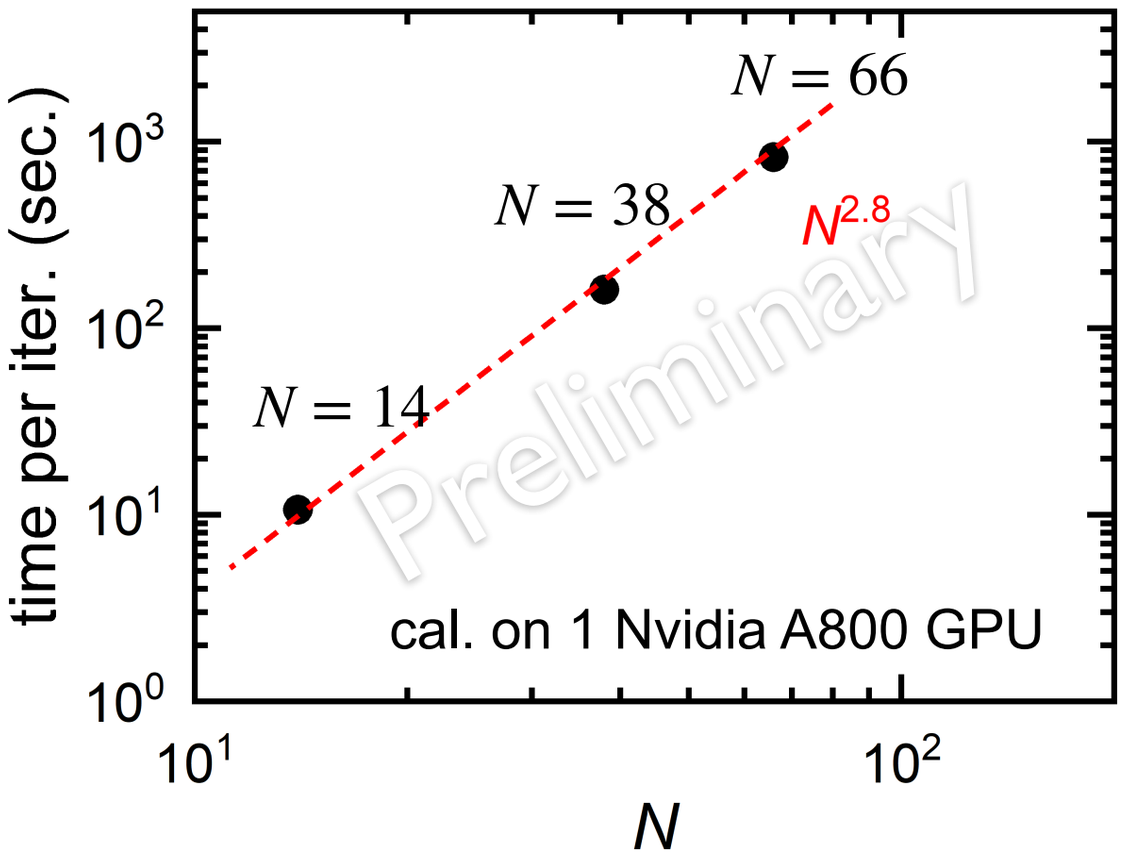
- After infinite-volume extrapolation, the present VMC calculations reproduce the thermodynamic limit from fix-node DMC calculations.

Forbes, Gandolfi, and Gezerlis, PRL 106, 235303 (2011)

- Computational scaling $O(N^3)$; no explicit dependence on box size L.

- Once the wave function is solved, it can be saved for calculating other observables without repeating VMC calculations.





Summary and outlooks

- **Neural-network Variational Monte Carlo is a powerful method for solving the *ab initio* nuclear many-body problem.**

  - Starting from a LO pionless EFT Hamiltonian, FeynmanNet can provide accurate polynomial-scaling solutions for nuclei.
  - To solve realistic nuclear forces, one can either improve the neural-network ansatz or perform subsequent Diffusion Monte Carlo calculations.
  - The extension to lattice systems is demonstrated by solving up to $N = 38$ neutrons.

Future developments towards excited states, accurate solution of high-precision nuclear forces, nuclear lattice EFT, …

Thank you for your attention!

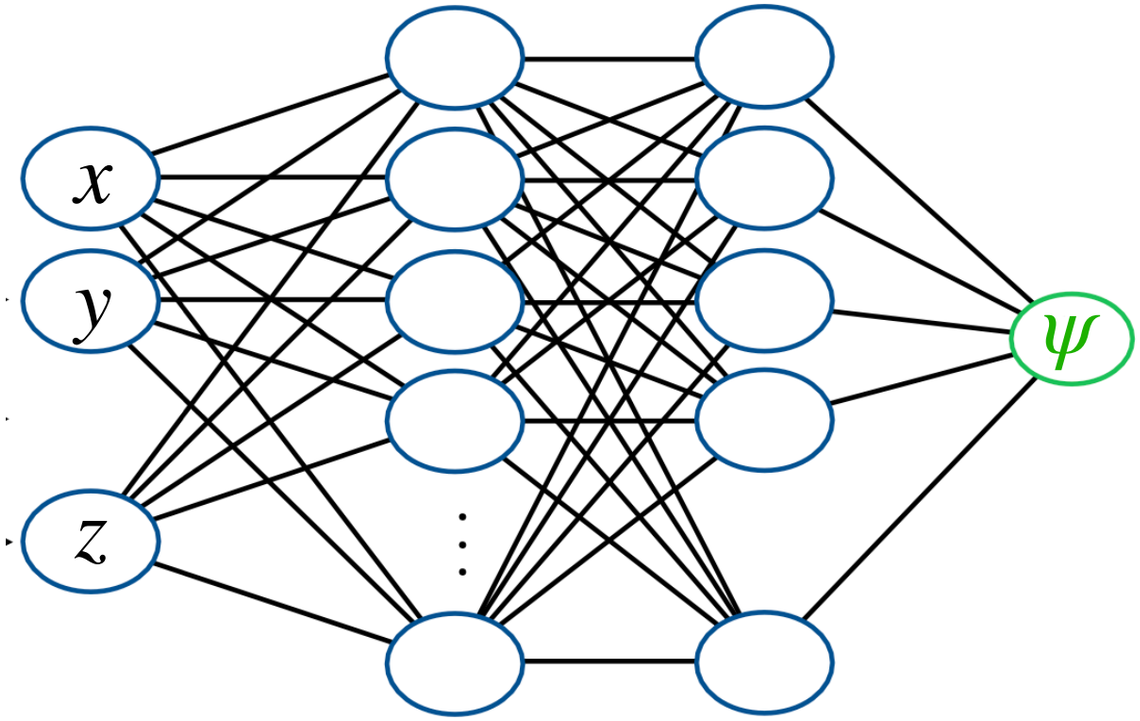
Example: an one-body problem

- Schrödinger equation with a Woods-Saxon potential

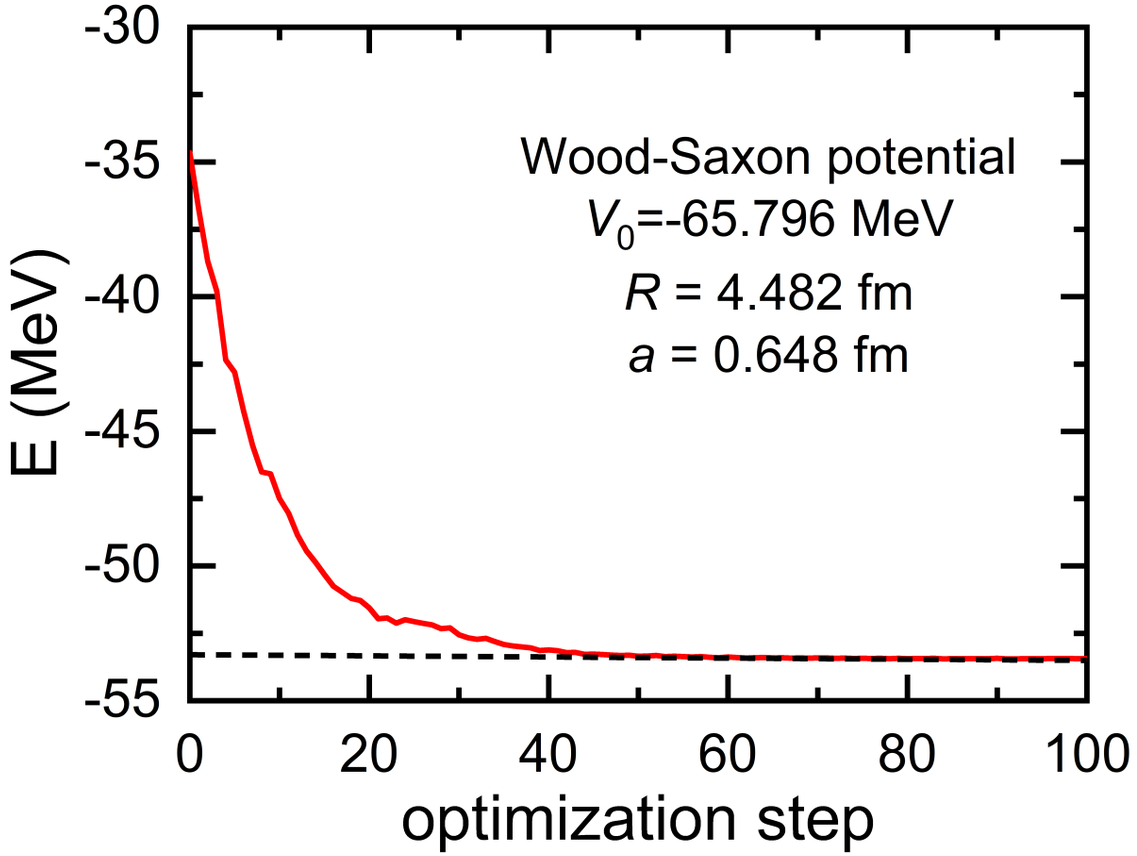
$$-\frac{\nabla^2}{2M_N}\psi(\boldsymbol{r};w)+\frac{V_0}{1+\exp[(r-R_0)/a]}\psi(\boldsymbol{r};w)=E\psi(\boldsymbol{r};w)$$

- Gradient descent

$$\boldsymbol{w}^{(t+1)}=\boldsymbol{w}^{(t)}-\gamma\nabla_w E(\boldsymbol{w})$$



$N=16$



Slater-Jastrow Ansatz

- Nuclear many-body wave function
  - *Spatial and spin-isospin d.o.f.:* $x_i = (\boldsymbol{r}_i, s_i, t_i)$
  - *Exchange antisymmetry* $\Psi(\ldots, x_i, \ldots, x_j, \ldots) = -\Psi(\ldots, x_j, \ldots, x_i, \ldots)$
- Slater determinants: antisymmetric, but no correlations

$$\det[\phi(x)] = \begin{vmatrix} \phi_1(x_1), & \phi_1(x_2), & \cdots, & \phi_1(x_A) \\ \phi_2(x_1), & \phi_2(x_2), & \cdots, & \phi_2(x_A) \\ \vdots, & \cdots, & \ddots, & \vdots \\ \phi_A(x_1), & \phi_A(x_2), & \cdots, & \phi_A(x_A) \end{vmatrix}$$

- Slater-Jastrow ansatz with neural-network Jastrow correlator

$$\Psi(x_1, \ldots, x_A) = J(x_1, \ldots, x_A) \det \phi(x) \qquad J(x_1, \ldots, x_A) = \rho\left(\sum_{i=1}^{A} \phi(x_i)\right)$$

The exchange symmetry of Jastrow is ensured by the summation.

Comparison with DMC calculations

- The neural-network Jastrow outperforms conventional ones.
- Considering symmetries can augment the performance of neural networks.

The remaining difference is due to the approximate factorization of wave function ansatz into a Slater and a Jastrow part.

$$\Psi(x_1, \ldots, x_A) = J(x_1, \ldots, x_A) \det \phi(x)$$

FeynmanNet

$$\Psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_A) = \mathrm{e}^{U(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_A)} \sum_{k=1}^{N_{\mathrm{det}}} w_n \det \left[\Phi_{\mathrm{BF}}^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_A) \right]$$

YLY and Zhao, PRC 107, 034320 (2023)

- Jastrow factor

$$U(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_A) = \rho \left(\sum_{i<j} \eta(\boldsymbol{x}_{ij}) \right)$$

$$\boldsymbol{x}_{ij} = (\boldsymbol{r}_{ij}, r_{ij}, s_i, s_j, t_i, t_j)$$

Universal representation of *symmetric* function

DeepSets: Zaheer et al., arXiv.1703.06114 (2017)

- Determinants with backflow

$$\begin{vmatrix} f_1(\boldsymbol{x}_1; \{\boldsymbol{x}_{1/}\}), & \cdots, & f_1(\boldsymbol{x}_A; \{\boldsymbol{x}_{A/}\}) \\ f_2(\boldsymbol{x}_1; \{\boldsymbol{x}_{1/}\}), & \cdots, & f_2(\boldsymbol{x}_A; \{\boldsymbol{x}_{A/}\}) \\ \vdots, & \ddots, & \vdots \\ f_A(\boldsymbol{x}_1; \{\boldsymbol{x}_{1/}\}), & \cdots, & f_A(\boldsymbol{x}_A; \{\boldsymbol{x}_{A/}\}) \end{vmatrix}$$

Universal representation of *anti-symmetric* function

Pfau et al., Phys. Rev. Research (2020)