Neural-network quantum Monte Carlo approaches for ab initio nuclear structure

Yilong Yang

Peking University

In collaboration with:

Evgeny Epelbaum, Chen Ji, Jie Meng, Lu Meng, and Pengwei Zhao

Email: yl_yang@pku.edu.cn

Outline

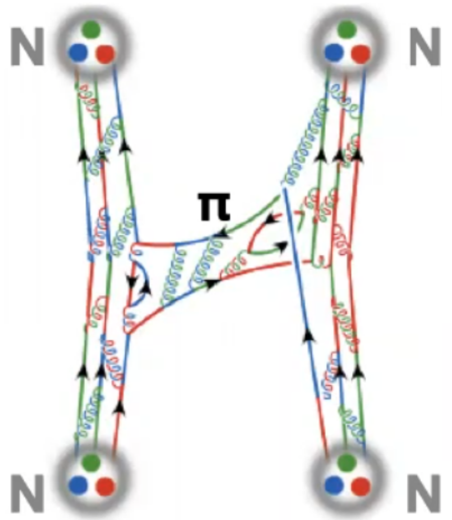
- Methodology
  - Introduction
  - Quantum Monte Carlo
  - Neural-network wave functions
- Applications
  - Probing long-range 3NF in peripheral $n\alpha$ scattering
  - Solving the Zemach radius puzzle in ^{6}Li
- Summary and outlooks

Ab initio nuclear theory

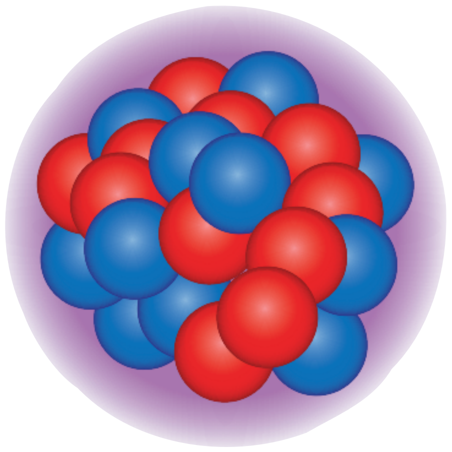
$$H = \sum_{i=1}^{A} \frac{\boldsymbol{p}_i^2}{2M_i} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk}$$

$$H\Psi(x_1, x_2, \ldots, x_A) = E\Psi(x_1, x_2, \ldots, x_A)$$

- **Generate the nuclear force**
  - Pionless EFT: contact interactions
  - Chiral EFT: contact + long-range π-exchange interactions



- **Solve the nuclear many-body problem *accurately***
  - 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

- Solves the many-body problem accurately and nonperturbatively
- Able to work with the bare EFT interactions
- Gives access to many nuclear properties, including spectra, transitions, form factors, responses, etc.

Lattice QCD

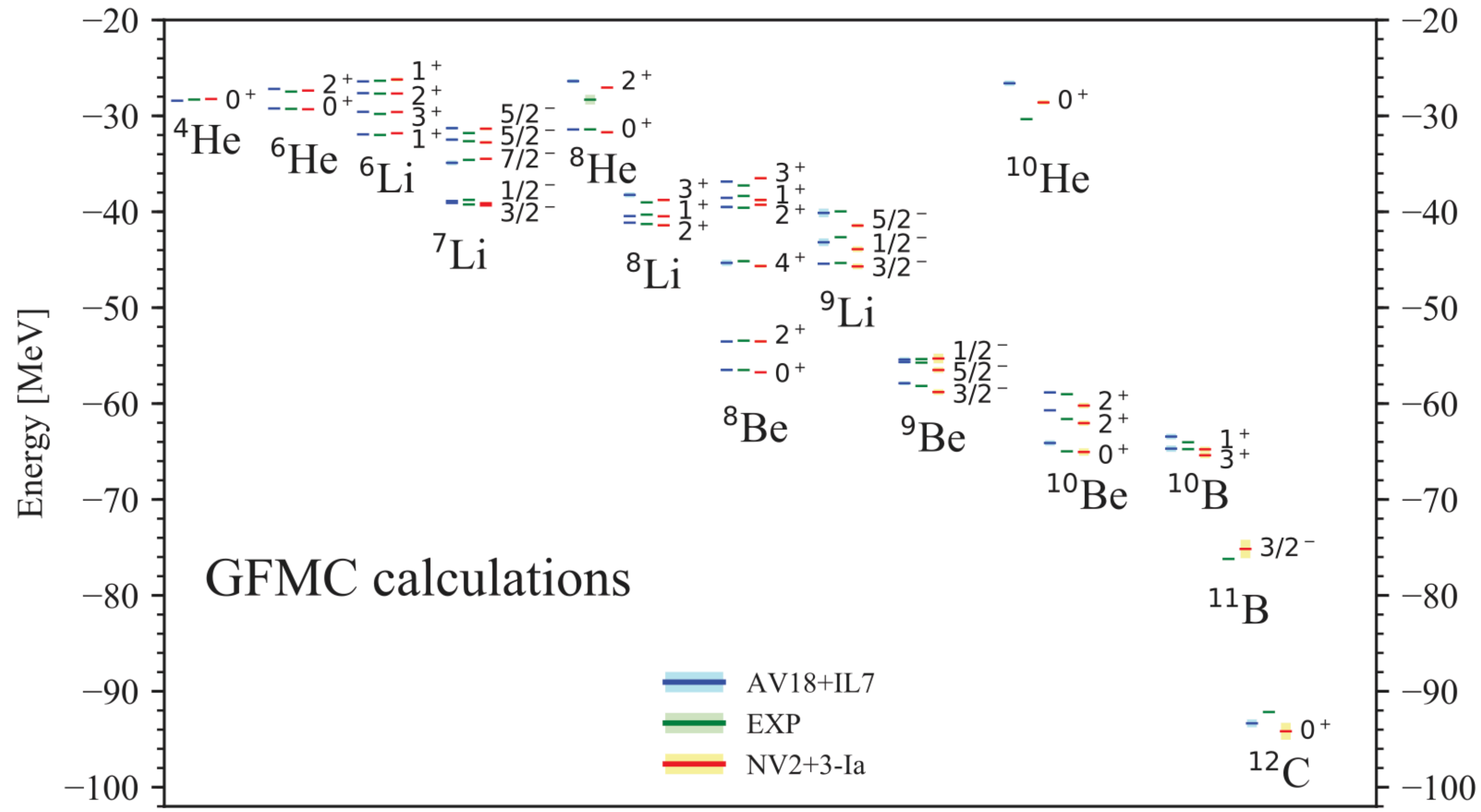
Nuclear Physics

Cold Atoms

Atoms/Molecules

Condensed Matter

…



Piraulli et al., Phys. Rev. Lett. 120, 052503 (2018)

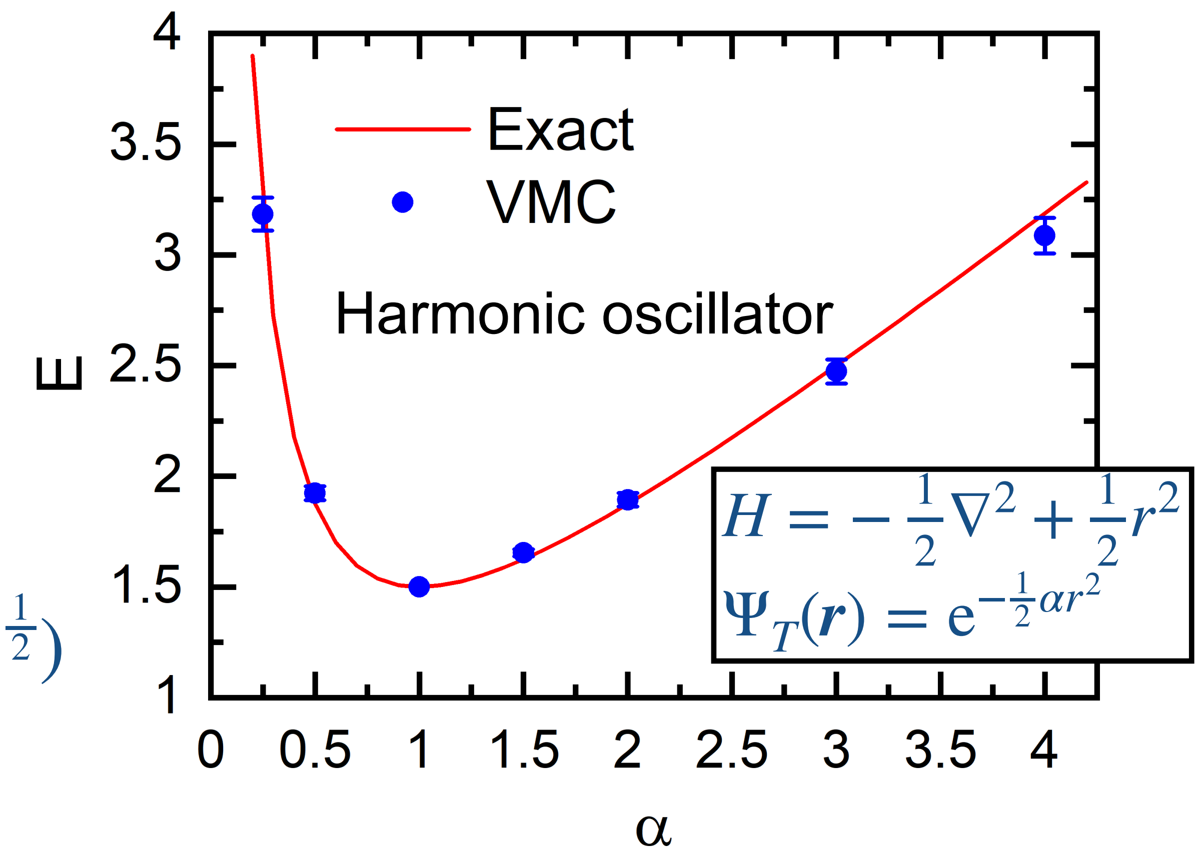
Quantum Monte Carlo methods

- Variational Monte Carlo (VMC)

Miminize the energy given trial function form

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

$$E_T = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \simeq \frac{1}{N} \sum_n E_L(R_n) + O(N^{-\frac{1}{2}})$$



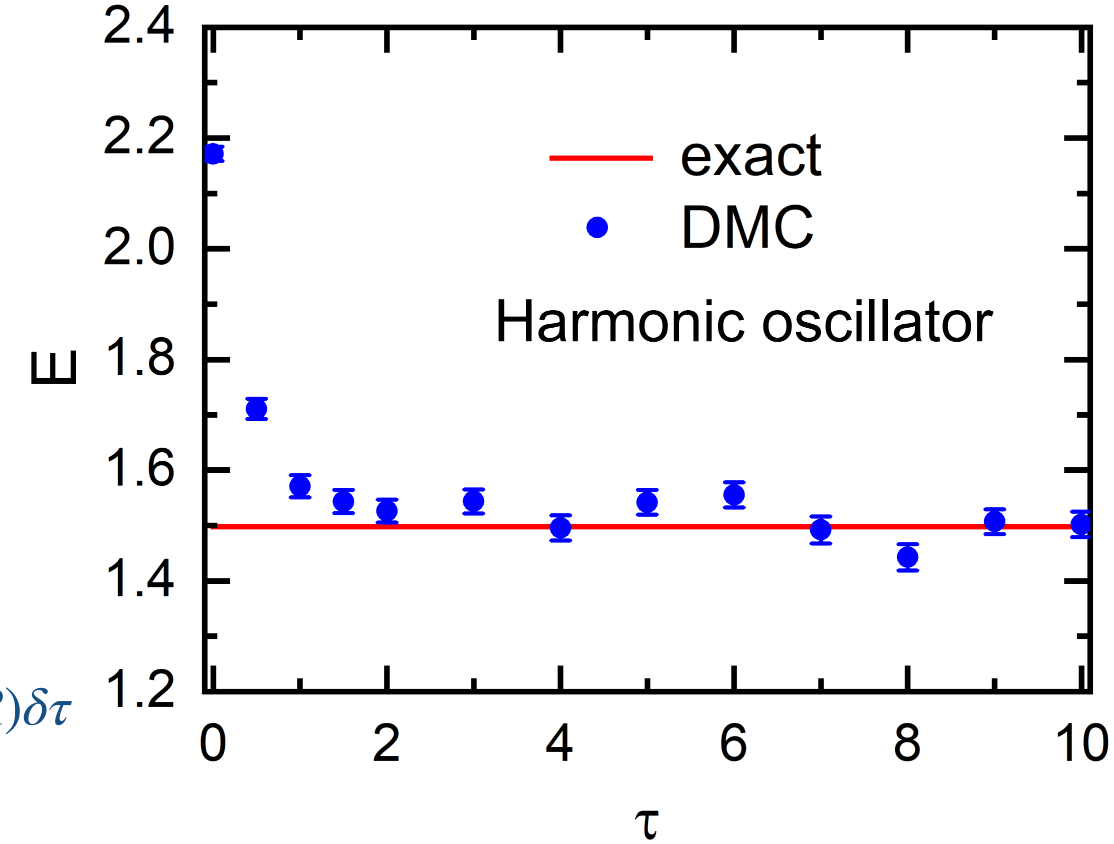
- Diffusion Monte Carlo (DMC)

Stochastic imaginary-time propagation

Initial wave function usually given by VMC

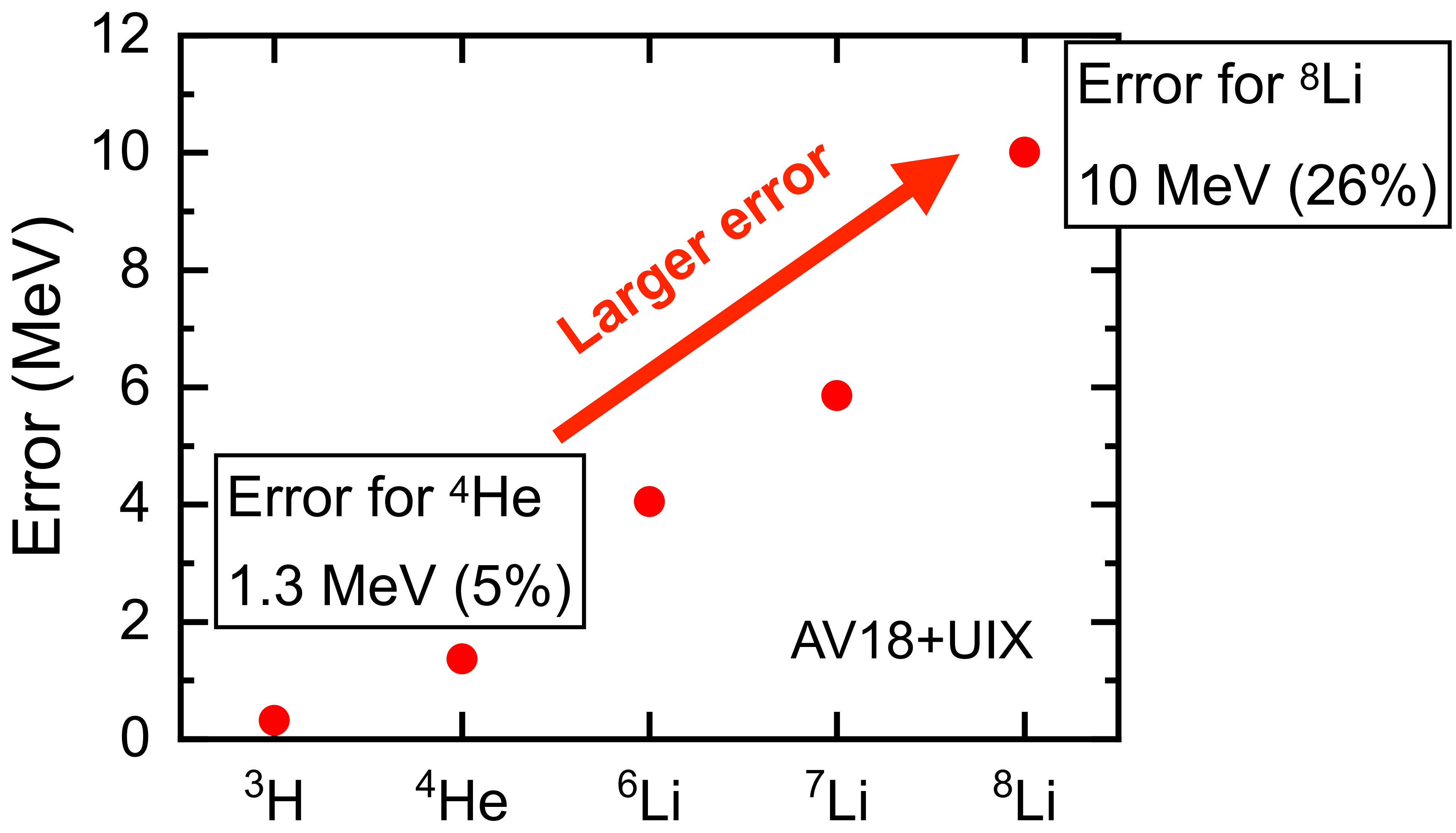
$$\lim_{\tau \to \infty} e^{-(\hat{H}-E)\tau} | \Psi_T \rangle = | \Psi_0 \rangle,$$

$$\langle R' | e^{-H\delta\tau} | R \rangle = \left(\frac{M}{2\pi\delta\tau} \right)^{\frac{3A}{2}} e^{-\frac{M}{2\delta\tau}(R'-R)^2} e^{-V(R)\delta\tau}$$



Challenge in Variational Monte Carlo

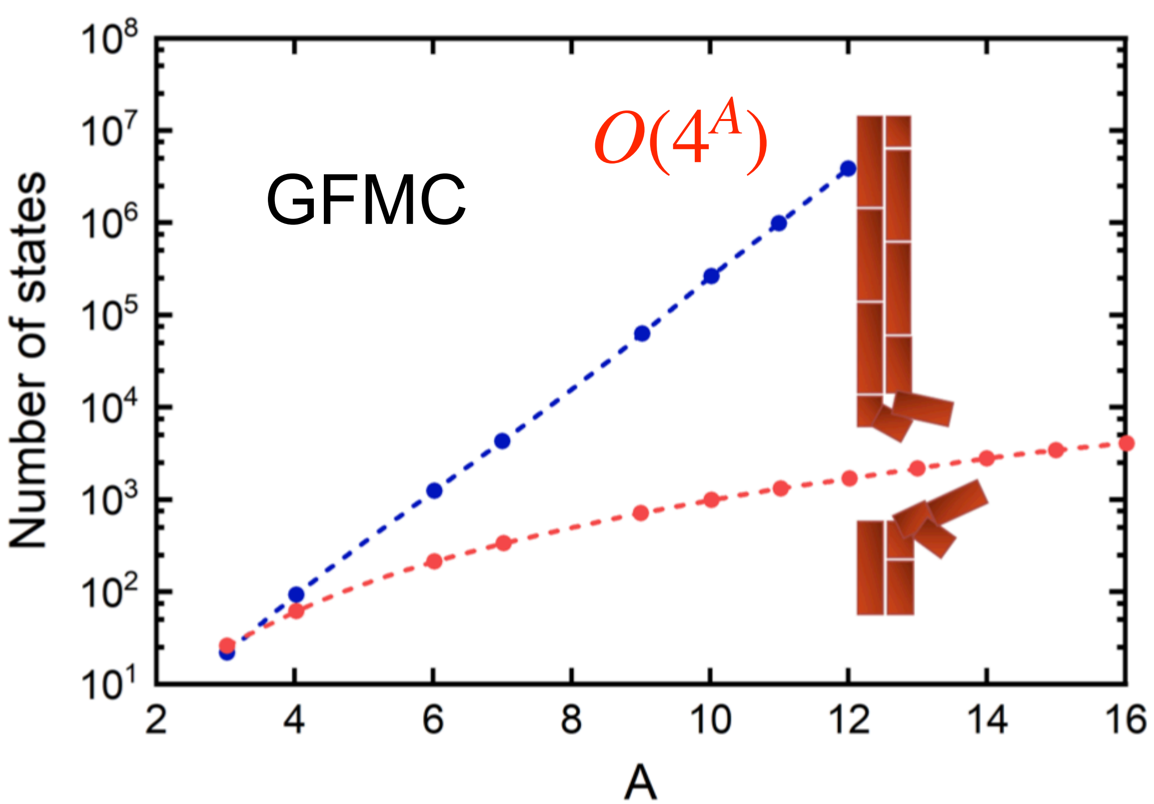
- Accuracy depends crucially on the quality of trial wave functions
- "Conventional" trial wave functions cannot reach the ground states variationally, and their quality deteriorates rapidly with increasing A

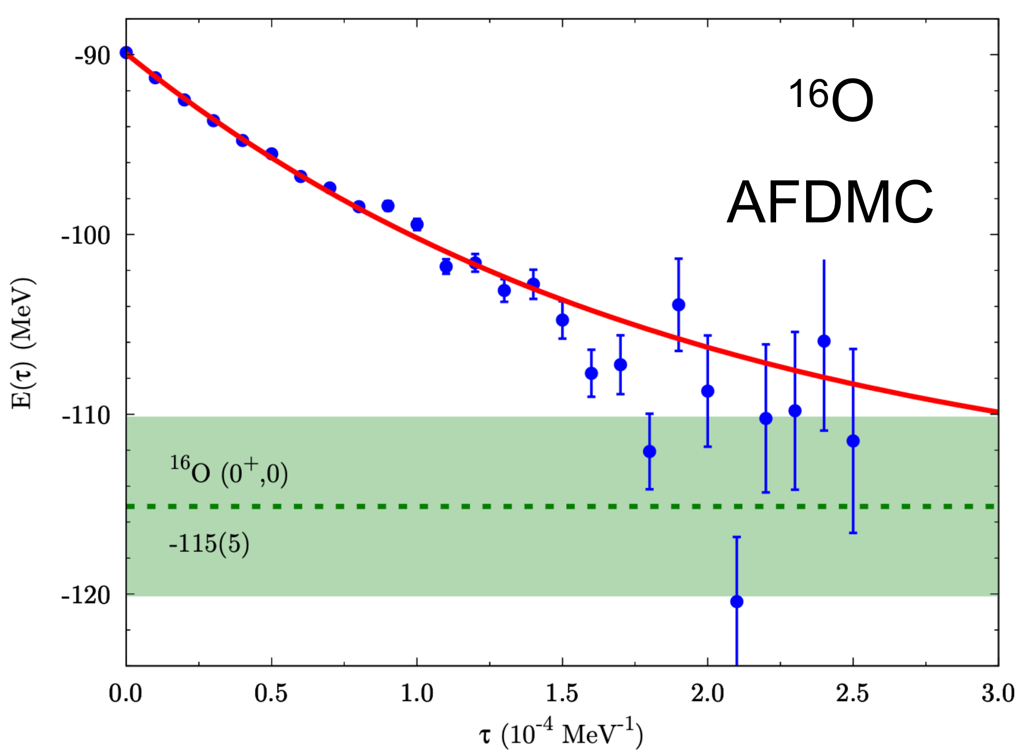


QMC calculations of $A \leq 8$ nuclei: Wiringa et al., PRC 62, 014001 (2000)

Challenge in Diffusion Monte Carlo

- Green's function Monte Carlo (GFMC) explicitly sums over (iso)spins, which leads to exponential scaling with A, limited to light $A \leq 12$ nuclei
- Auxiliary-field diffusion Monte Carlo (AFDMC) can adapt larger systems by sampling (iso)spins, but suffers from more severe sign problem



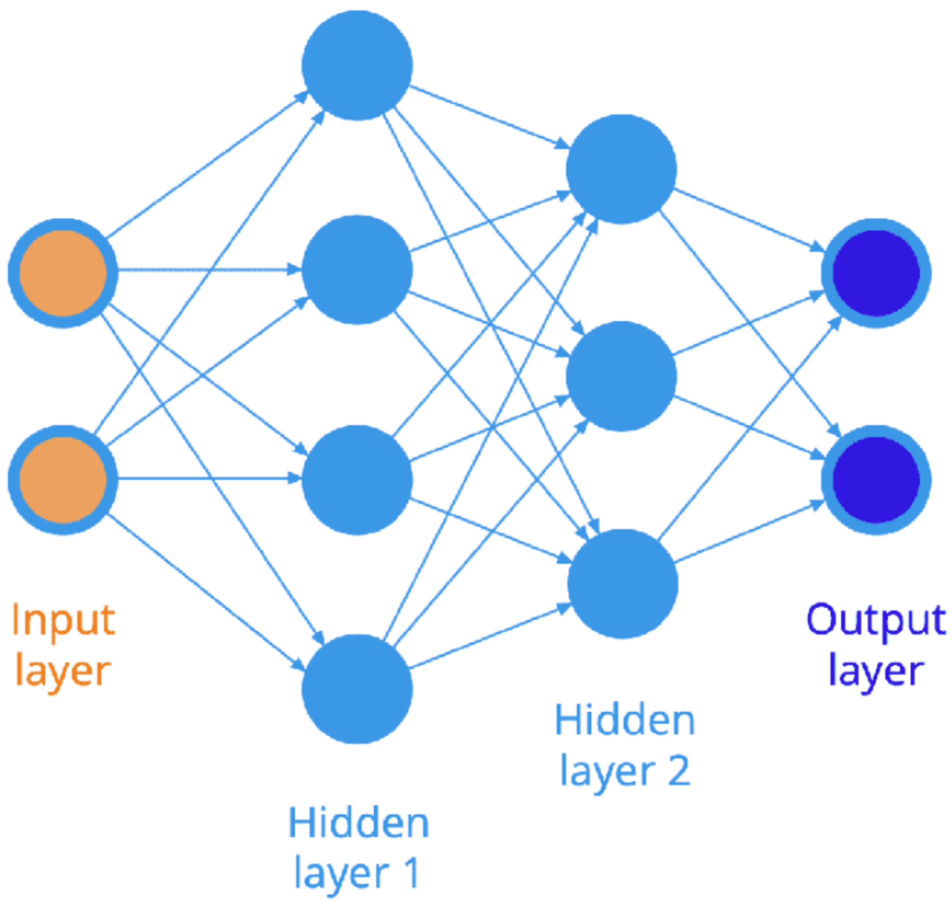


Lonardoni et al., PRC 97, 044318 (2018)

➡ Devising a polynomial scaling and accurate trial wave function

Neural networks

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



outputs — inputs

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

w, b: adjustable weights (variational param.)

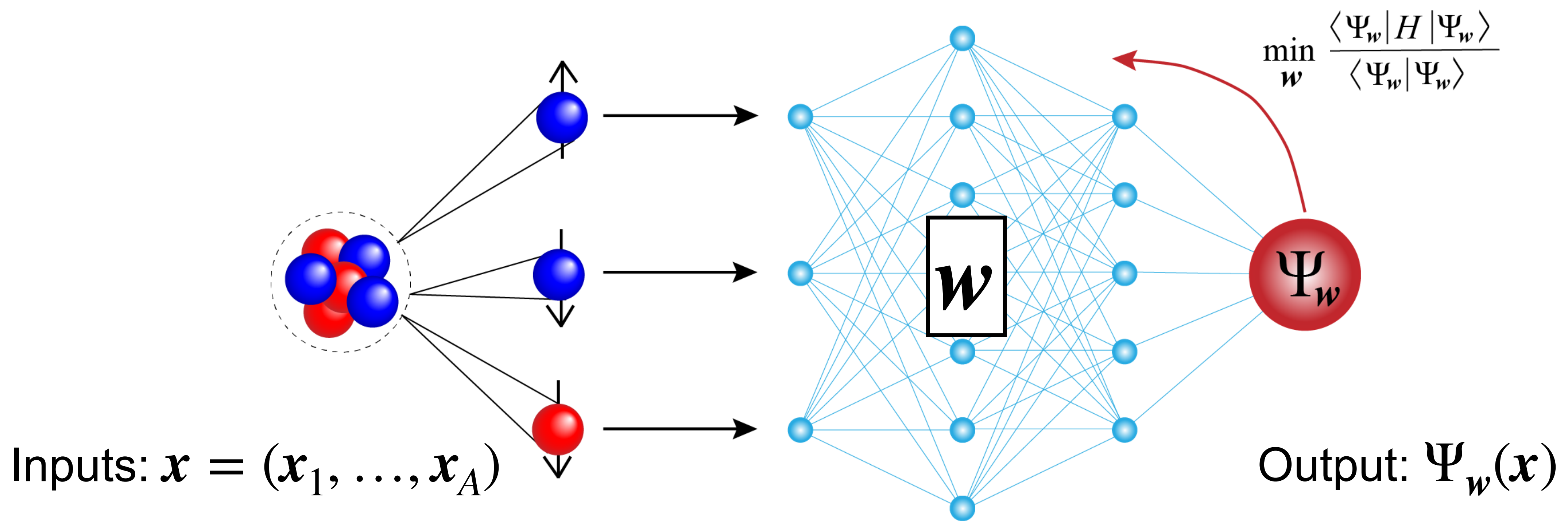
σ: nonlinear functions, e.g. $\tanh(x)$

Universal Approximation Theorem: *existence / limit theorem*

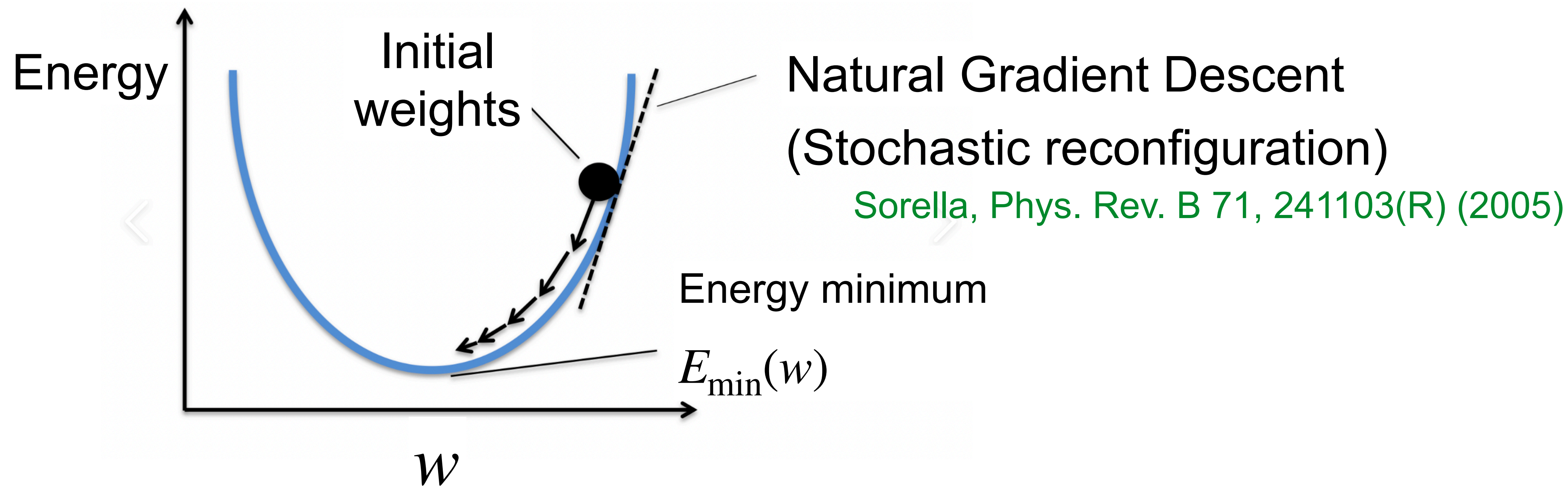
A single-hidden-layer neural network can approximate any continuous function given enough number of hidden neurons.

Neural-network variational Monte Carlo

- Neural networks: efficiently parametrize many-body wave functions



- Variational Monte Carlo: train neural networks with variational principle



LO pionless EFT Hamiltonian

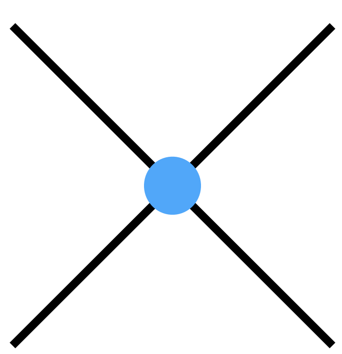
- We first consider a nuclear Hamiltonian derived in LO pionless EFT

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

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

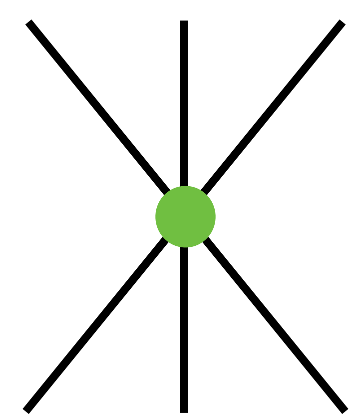
- NN potential: fit to S-wave scattering lengths, effective range, and deuteron binding energy

- 3N potential: adjusted to reproduce triton binding energy ()



$$v_{ij}^{\rm CI} = \sum_{p=1}^{4} v_p(r_{ij}) O_{ij}^{p}$$

$$O_{ij}^{p=1\text{-}4} = 1,\ \tau_{ij},\ \sigma_{ij},\ \sigma_{ij}\tau_{ij}$$



$$V_{ijk} = D \sum_{\rm cyc} e^{-(r_{ij}^2 + r_{jk}^2)/R_3^2}$$

r-space cutoffs: $R_{T=0} \simeq 1.5\ \text{fm},\ R_{T=1} \simeq 1.8\ \text{fm},\ R_3 = 1.0\ \text{fm}$

Neural-network wave functions

- 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

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

- Including many-body correlations 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

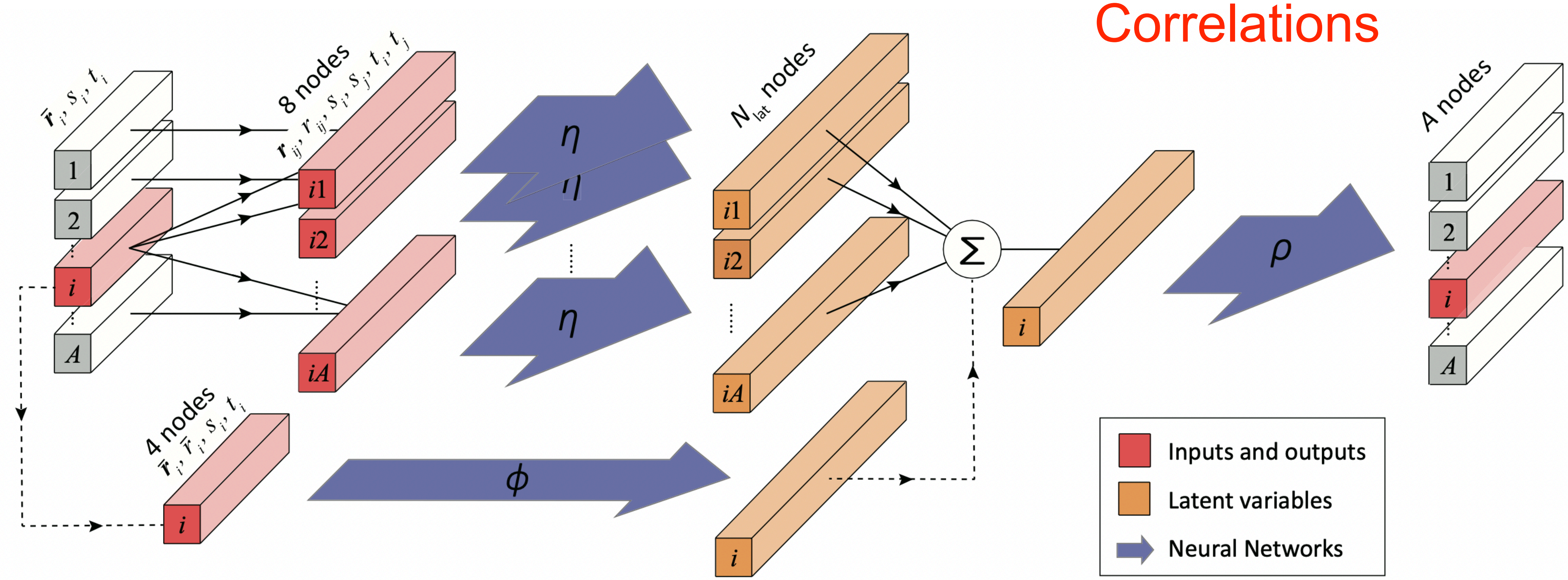
YLY and Pengwei Zhao, PRC 107, 034320 (2023); e-Print: 2211.13998

- We introduce spin-isospin dependent backflow with neural networks

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

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

Correlations



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

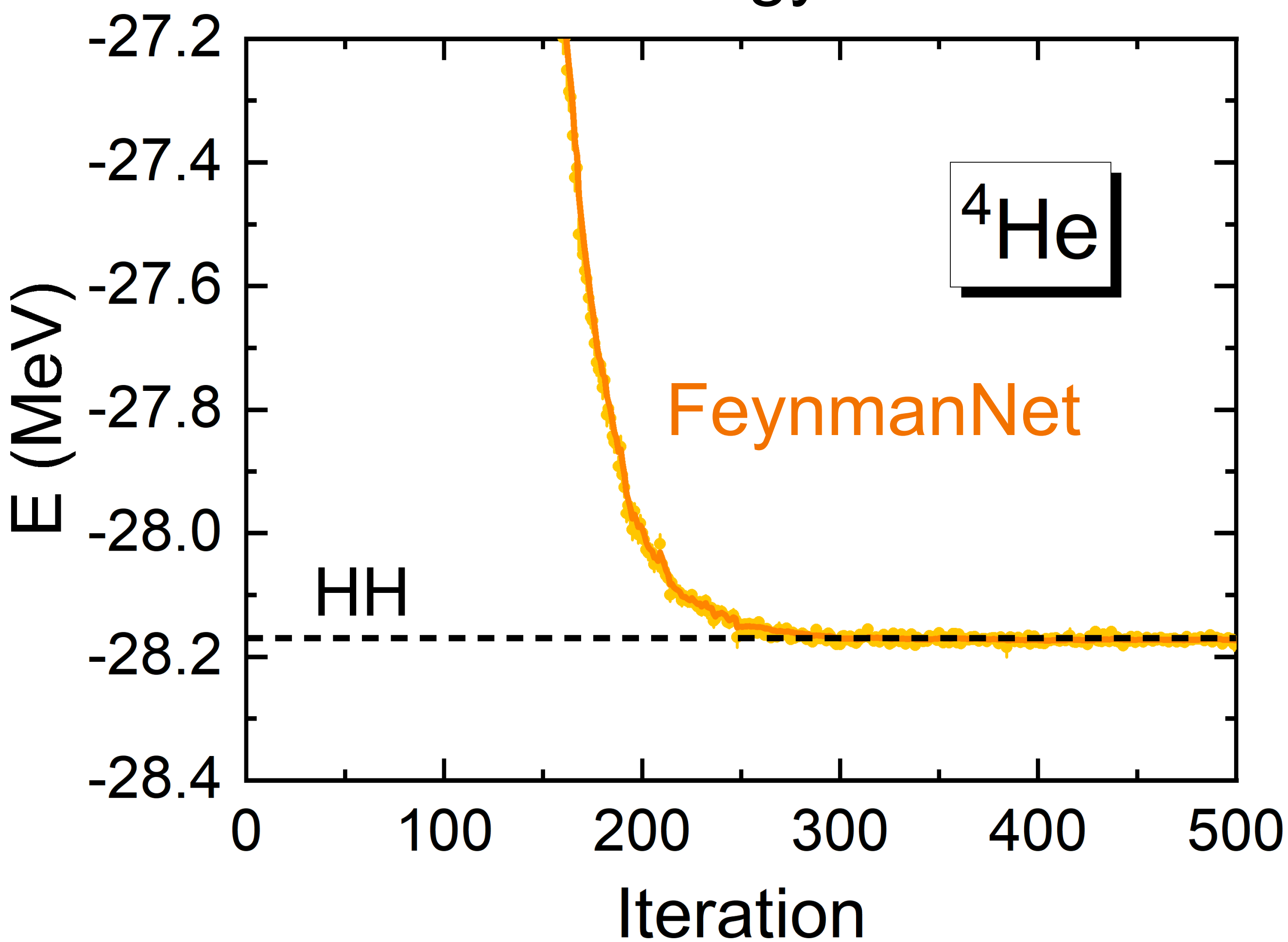
FeynmanNet: $A = 4$ nuclei

YLY and Pengwei Zhao, PRC 107, 034320 (2023); e-Print: 2211.13998

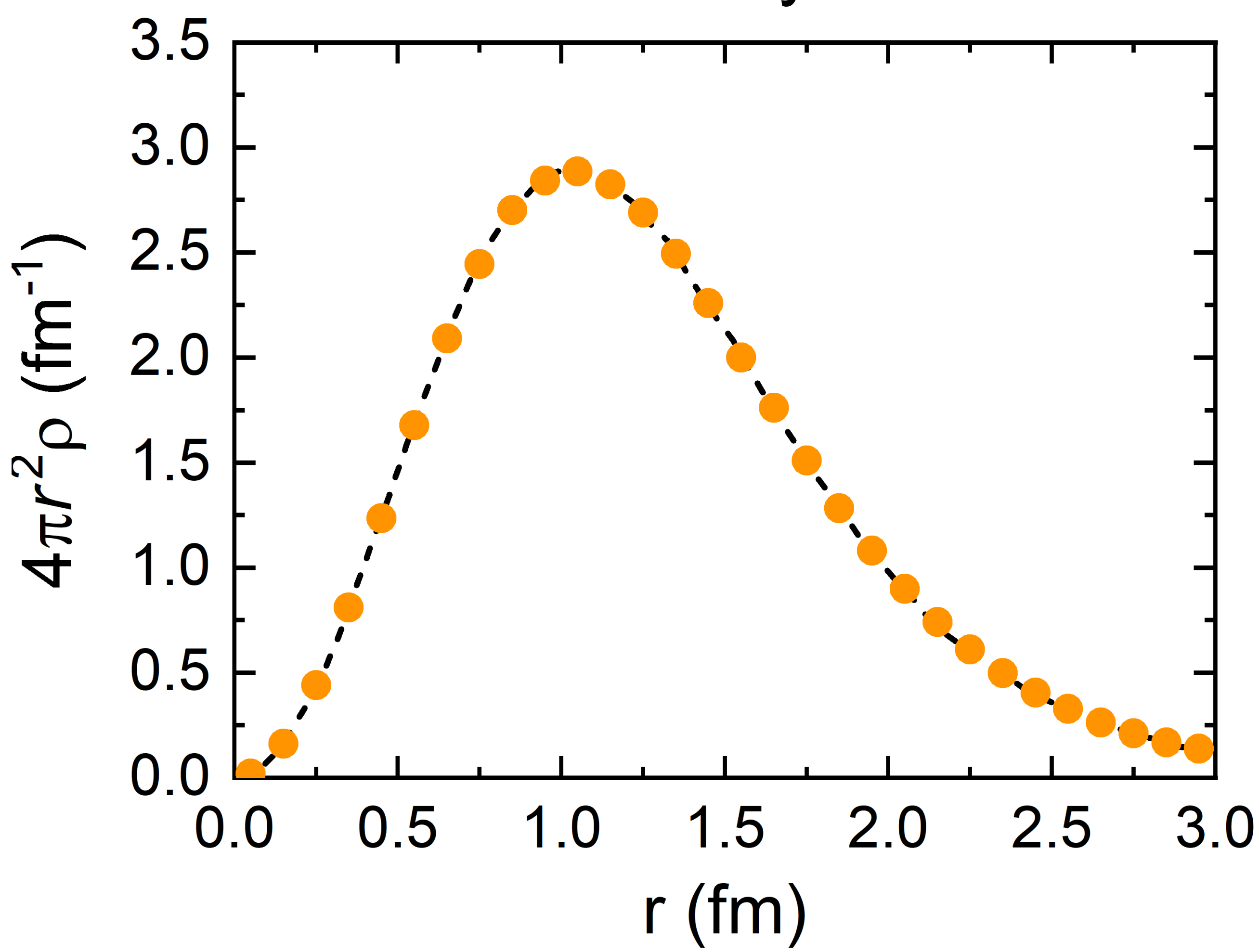
- Perfect agreement with the Hyperspherical Harmonics (HH) method

Accuracy at the level of ~0.01 MeV for $A \leq 4$ *nuclei*

Energy



Density



Pionless EFT $H_{\rm LO}$ is used

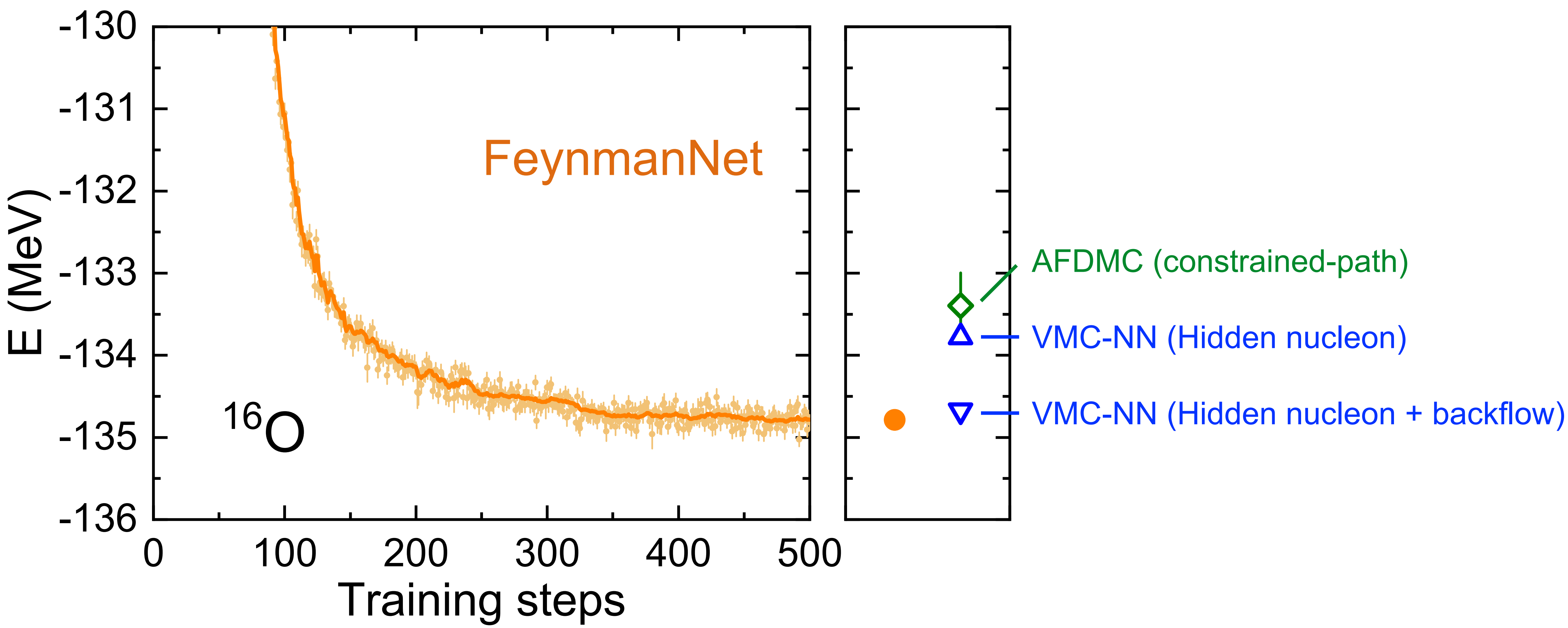
Schiavilla2012PRC

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

FeynmanNet: $A = 16$ nuclei

YLY and Pengwei Zhao, PRC 107, 034320 (2023); e-Print: 2211.13998

- 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)*

FeynmanNet: NLO pionless Hamiltonian

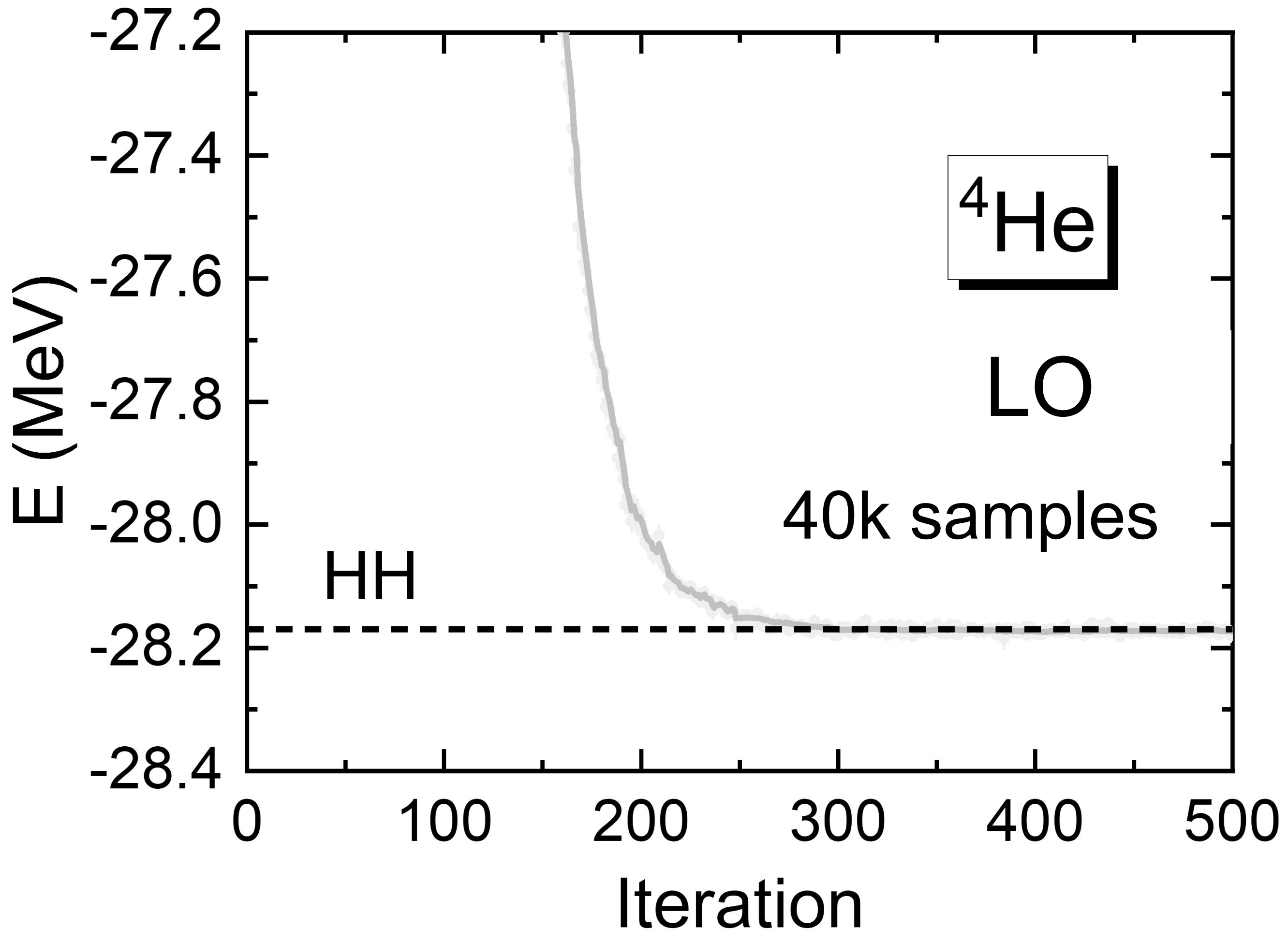
YLY and Pengwei Zhao, PRC 107, 034320 (2023); e-Print: 2211.13998

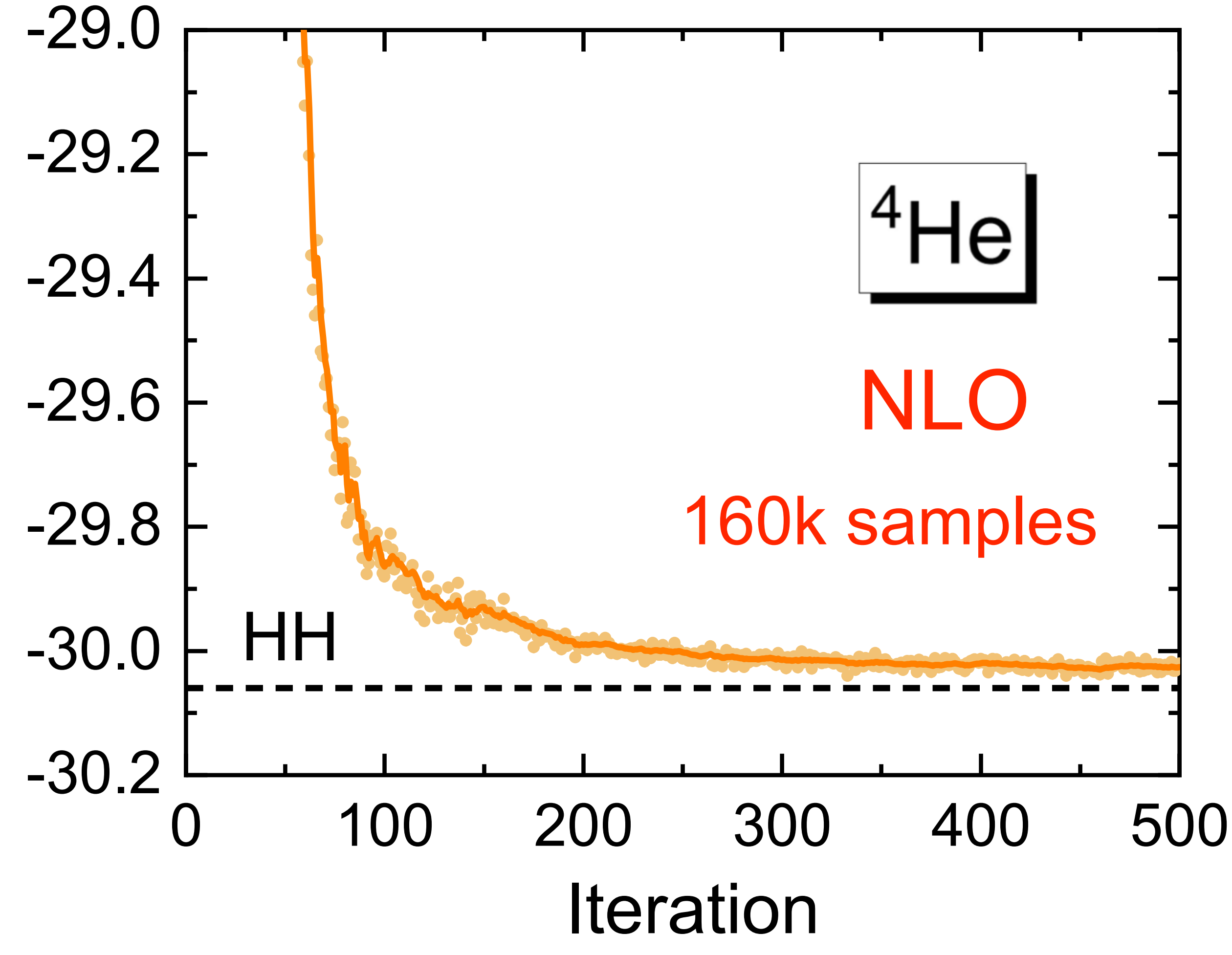
- NLO pionless Hamiltonian feasible for $A = 4$ nuclei, but requires much more number of MC samples and larger neural networks…

$$v_{ij}^{\mathrm{CI}} = \sum_{p=1}^{8} v_p(r_{ij}) O_{ij}^{p} \qquad O_{ij}^{p=1\text{-}8} \in \{1,\ \sigma_{ij},\ S_{ij},\ \boldsymbol{L}\cdot\boldsymbol{S}\} \otimes \{1,\ \tau_{ij}\}.$$

Noncentral coupling between spin and spatial d.o.f

$$3\sigma_i \cdot \hat{r}_{ij}\sigma_j \cdot \hat{r}_{ij} - \sigma_{ij}$$



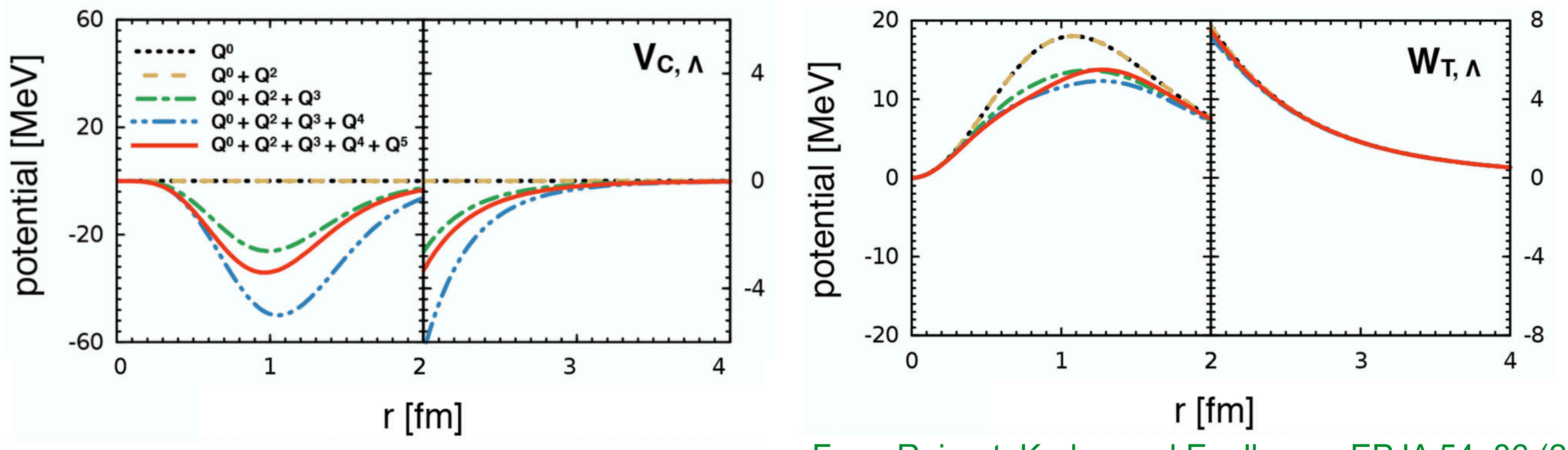


Realistic chiral interactions

- The tensor forces from π-exchange are essential in chiral EFT
- How to include tensor correlations in neural-network wave functions?
  - So far hard for NNs to directly learn the tensor correlations… YLY2023PRC
  - Use tensor operators explicitly in the w.f. $|\Psi(\mathscr{F}(R), \hat{O}^p)\rangle$

Neural networks — Spin-isospin operators

Central — Isovector tensor



From Reinert, Krebs, and Epelbaum, EPJA 54, 86 (2018)

Neural-network correlation functions

YLY and Pengwei Zhao, Chinese Phys. Lett. 42, 052101 (2025); e-Print: 2405.04203

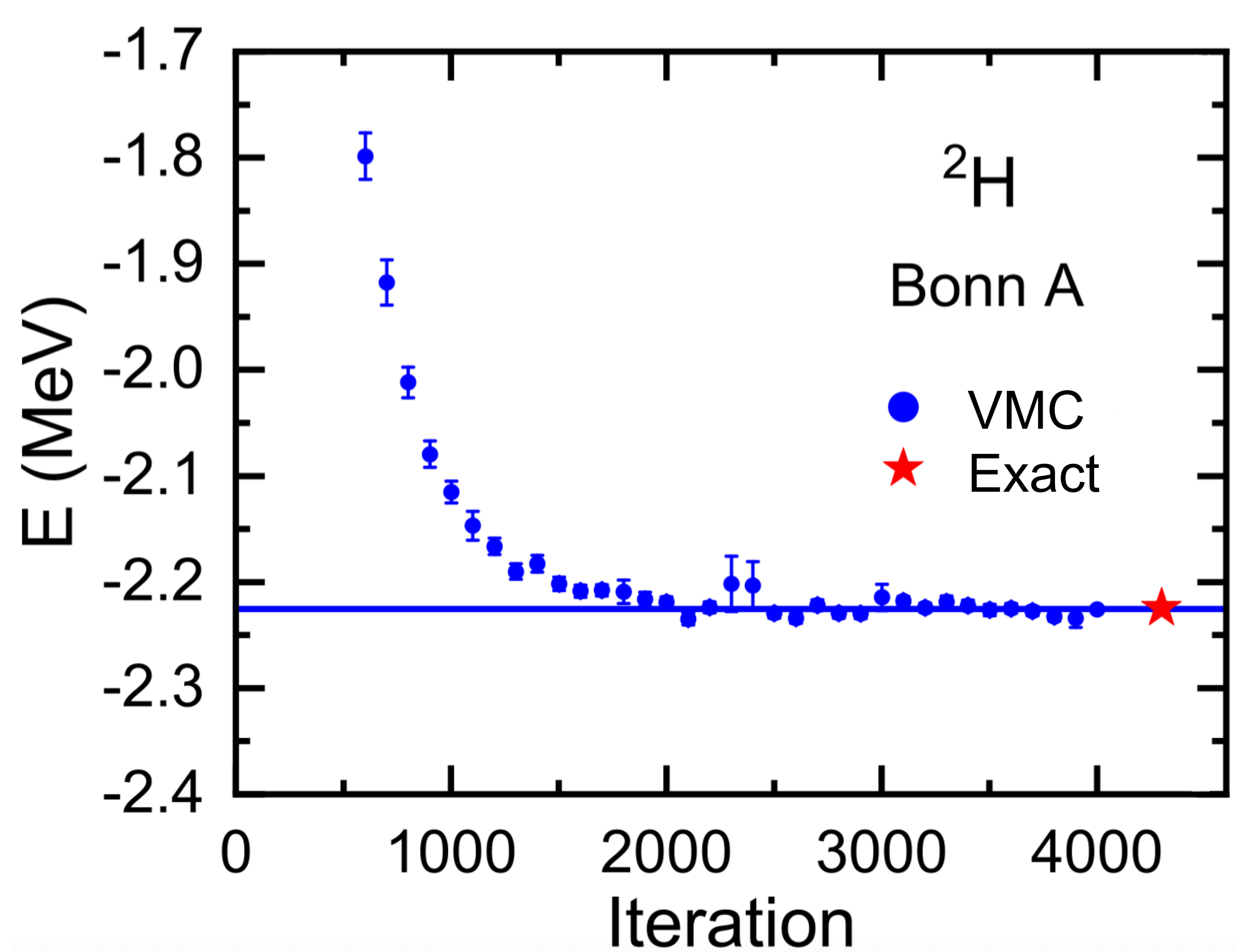
- Neural-network correlation functions + spin-isospin operators

$$|\Psi\rangle = \prod_{i<j} f(r_{ij}) \left[1 + \sum_{i<j} \sum_{p=2\text{-}6} u^p(r_{ij}) O_{ij}^p \right] |\Phi\rangle_{J^\pi,T}$$

f, u: neural networks

- Solves the deuteron exactly with Bonn potential

Similar form used in AFDMC calculations
Gandolfi et al., PRC 90, 061306(R) (2014)



Neural-network correlation functions

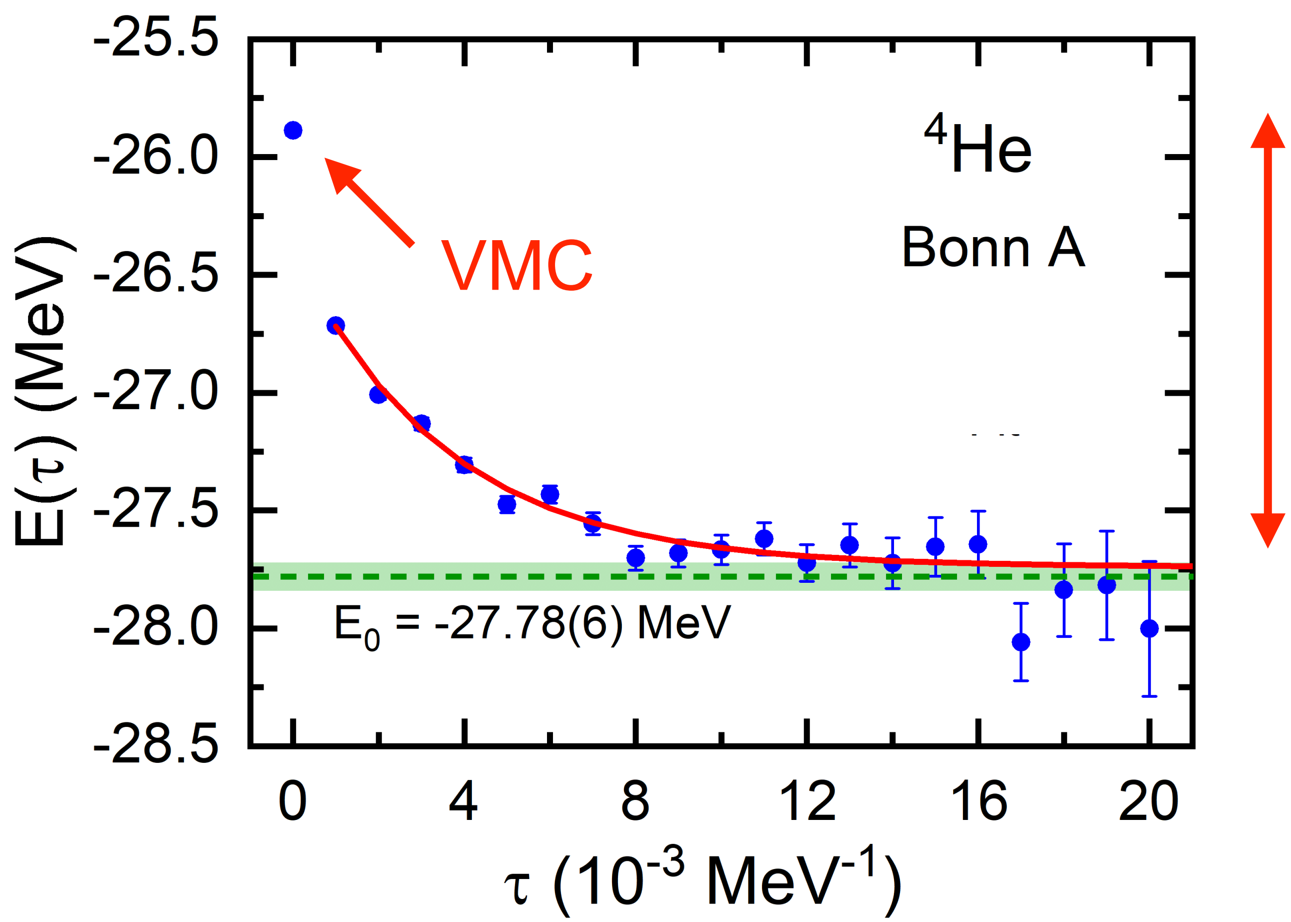
YLY and Pengwei Zhao, Chinese Phys. Lett. 42, 052101 (2025); e-Print: 2405.04203

- Neural-network correlation functions + spin-isospin operators

$$|\Psi\rangle = \prod_{i<j} f(r_{ij}) \left[1 + \sum_{i<j} \sum_{p=2\text{-}6} u^p(r_{ij}) O_{ij}^p \right] |\Phi\rangle_{J^\pi,T}$$

f, u: neural networks

- For few-body nuclei, Diffusion Monte Carlo is still needed

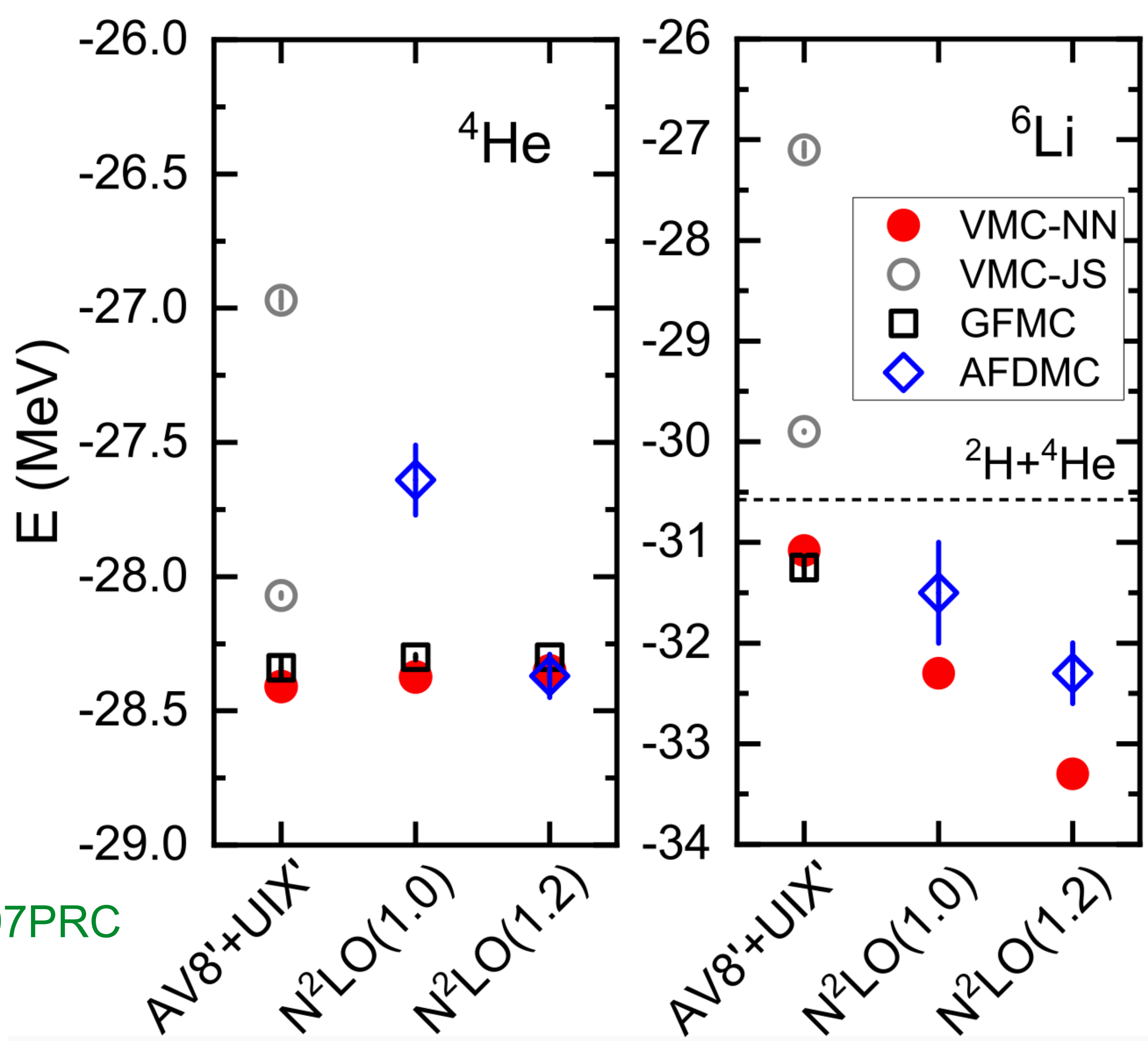


Shortage:

Only include two-body correlations

Accurate neural-network wave function for solving realistic forces

YLY, Evgeny Epelbaum, Ji Chen, and Pengwei Zhao, in preparation



Argonne+Urbana forces

Wiringa1995PRC, Pudliner1997PRC

Local χEFT forces

Gezerlis2014PRC, Lynn2017PRC

VMC-JS: Wiringa2000PRC
UsmaniPRC2012

GFMC: Lynn2017PRC

AFDMC: Lonardoni2018PRC

New trial wave function $|\Psi(\mathscr{F}(R), \hat{O}^p)\rangle$ with many-body correlations

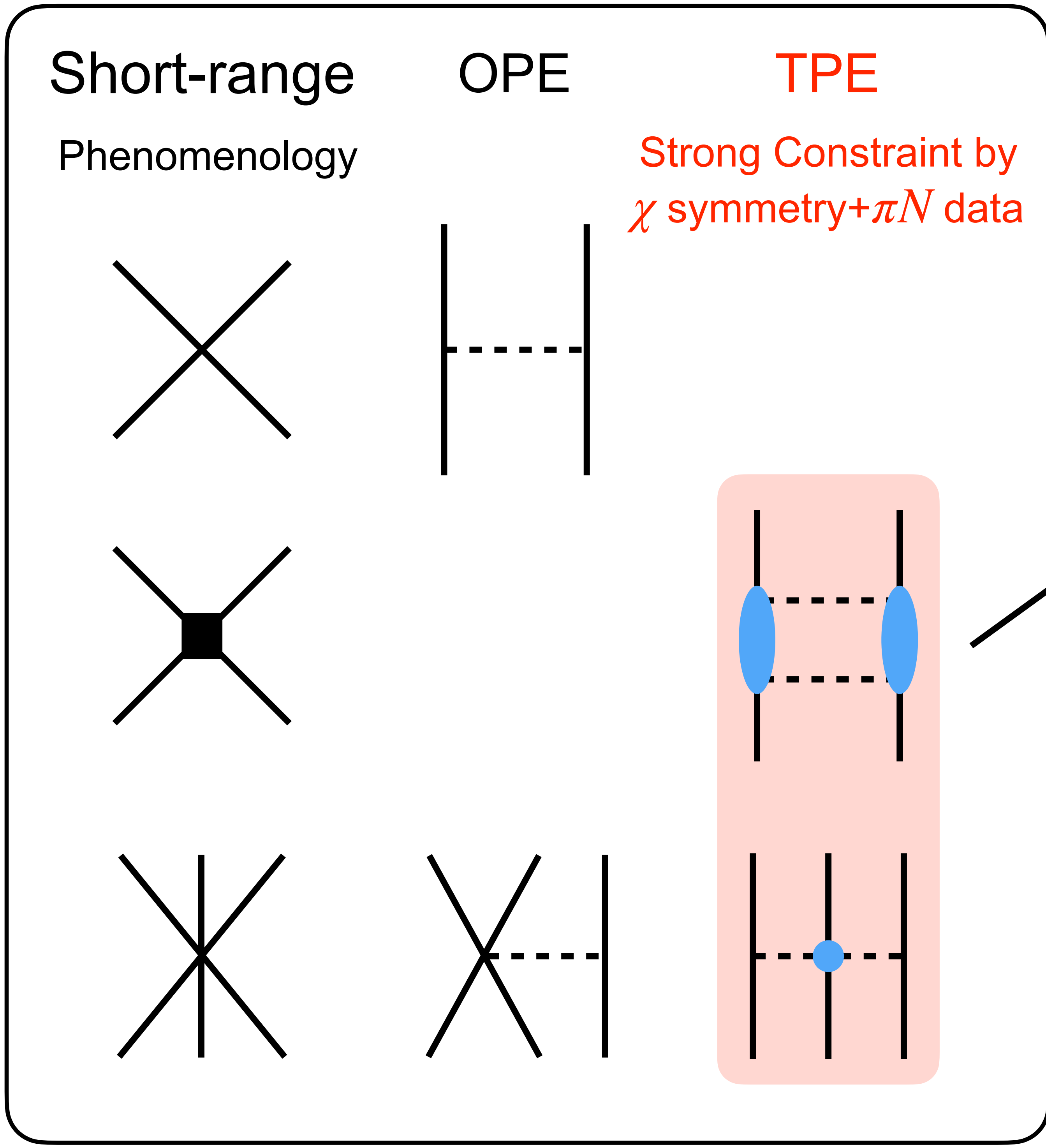
With new neural-network wave functions, VMC can for the first time provide virtually-exact solutions of the ground state energies with realistic NN+3N forces.

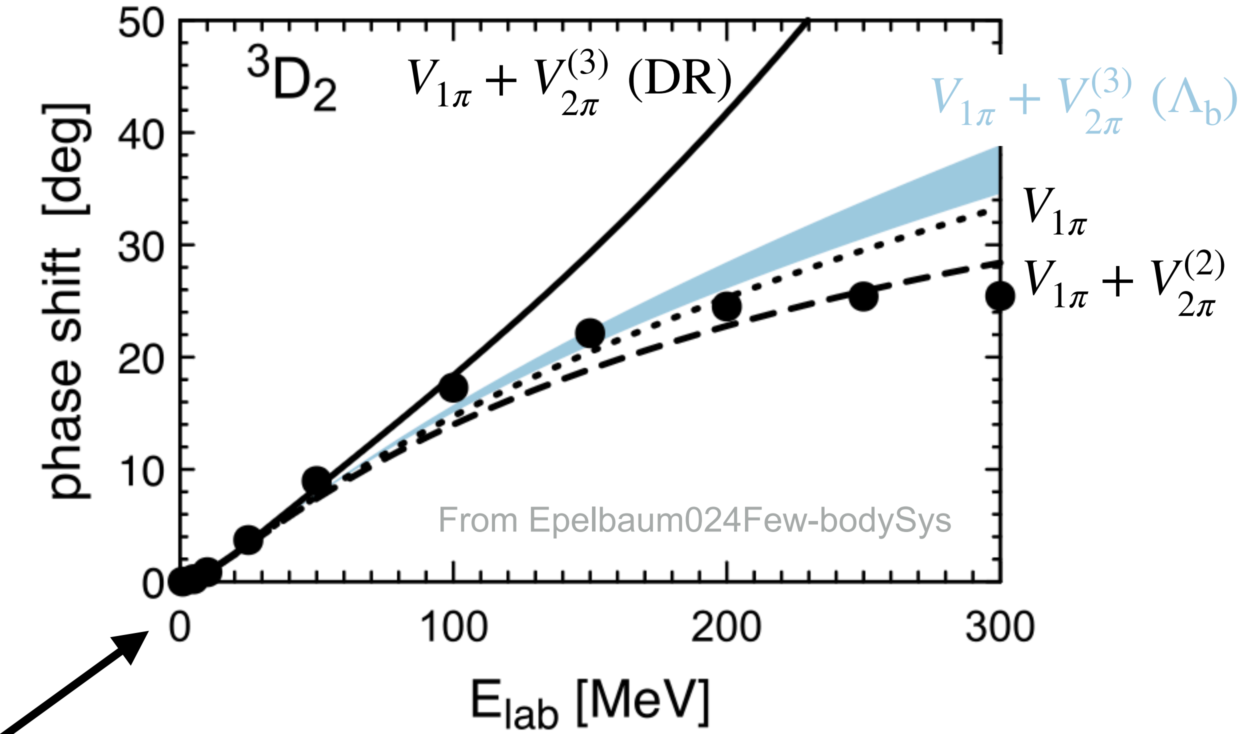
Probing long-range 3NF in peripheral $n\alpha$ scattering

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

Motivation

- Test the chiral EFT prediction of long-range behavior of nuclear force





Testing TPE NN force in peripheral NN scattering

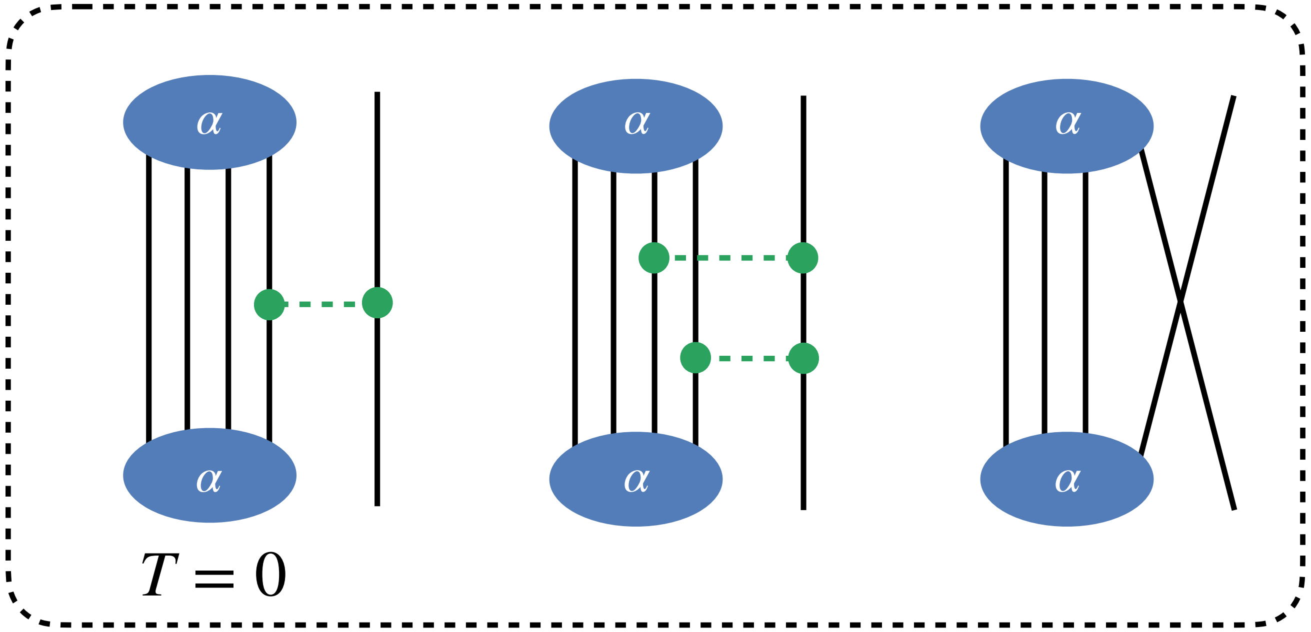
Kaiser, Brockmann, and Weise, NPA 625, 758 (1997)

1. Can one find more direct probe for TPE, without OPE interfering?
2. How to probe TPE 3NF?

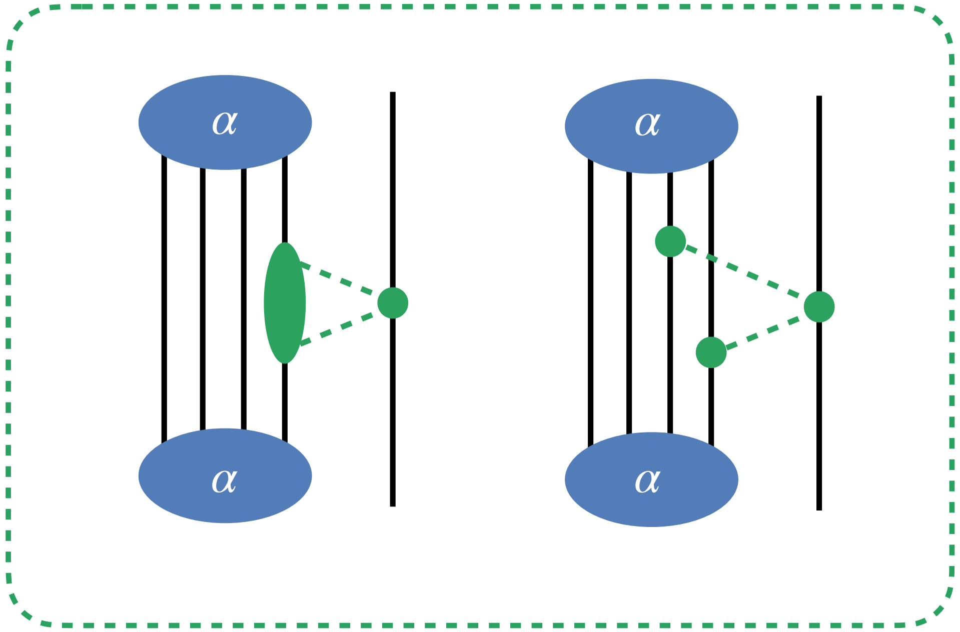
Peripheral $n\alpha$ scattering

- Peripheral $n\alpha$ scattering ($L \geq 2$) might be suitable!

Suggested in Higa and Robilotta, arXiv:nucl-th/9908062 (1999)



Suppressed



Allowed

- The existing studies have focused on S- and P-waves, where short-range mechanisms dominate, while no ab initio studies of peripheral $n\alpha$ scattering are available yet to probe long-range TPE three-nucleon forces.

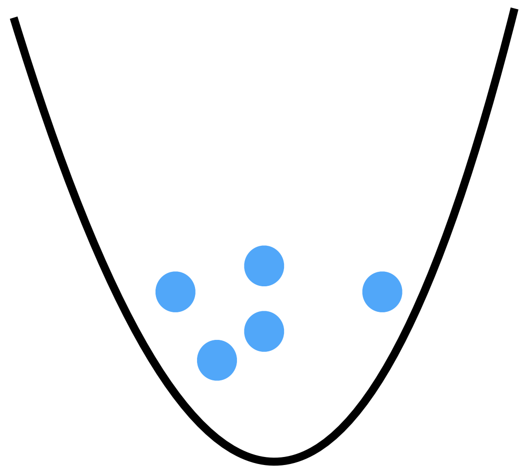
QMC: Carslon1987PRC, Lynn2016PRL, …
Faddeev-Yakubovsky: Lazauskas2018PRC, …
NCSM: Navrátil2016PS, Shirokov2018PRC, …
SVM: Bagnarol2023PLB, …

QMC + BERW formula for $n\alpha$ phase shifts

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

- $n\alpha$ phase shifts are extracted from the ^5He_l energy in a harmonic oscillator trap

$$k^{2l+1} \cot \delta_l^{n\alpha}(k) = (-1)^{l+1} \left(4\mu\omega\right)^{l+1/2} \frac{\Gamma((3+2l)/4 - \epsilon_l/(2\omega))}{\Gamma((1-2l)/4 - \epsilon_l/(2\omega))}$$



Busch et al., Found. Phys. 28, 549 (2008); Suzuki et al., PRA 80, 033601 (2009)

with $\epsilon_l = E(^5\mathrm{He}_l) - E_\alpha$. We focus on the $D_{5/2}$ wave (spin-orbit splitting between $D_{5/2}$ and $D_{3/2}$ at low energies are small).

- The trapped ^{5}He energy is calculated with neural-network VMC+DMC

YLY and Pengwei Zhao, Chinese Phys. Lett. 42, 052101 (2025); e-Print: 2405.04203

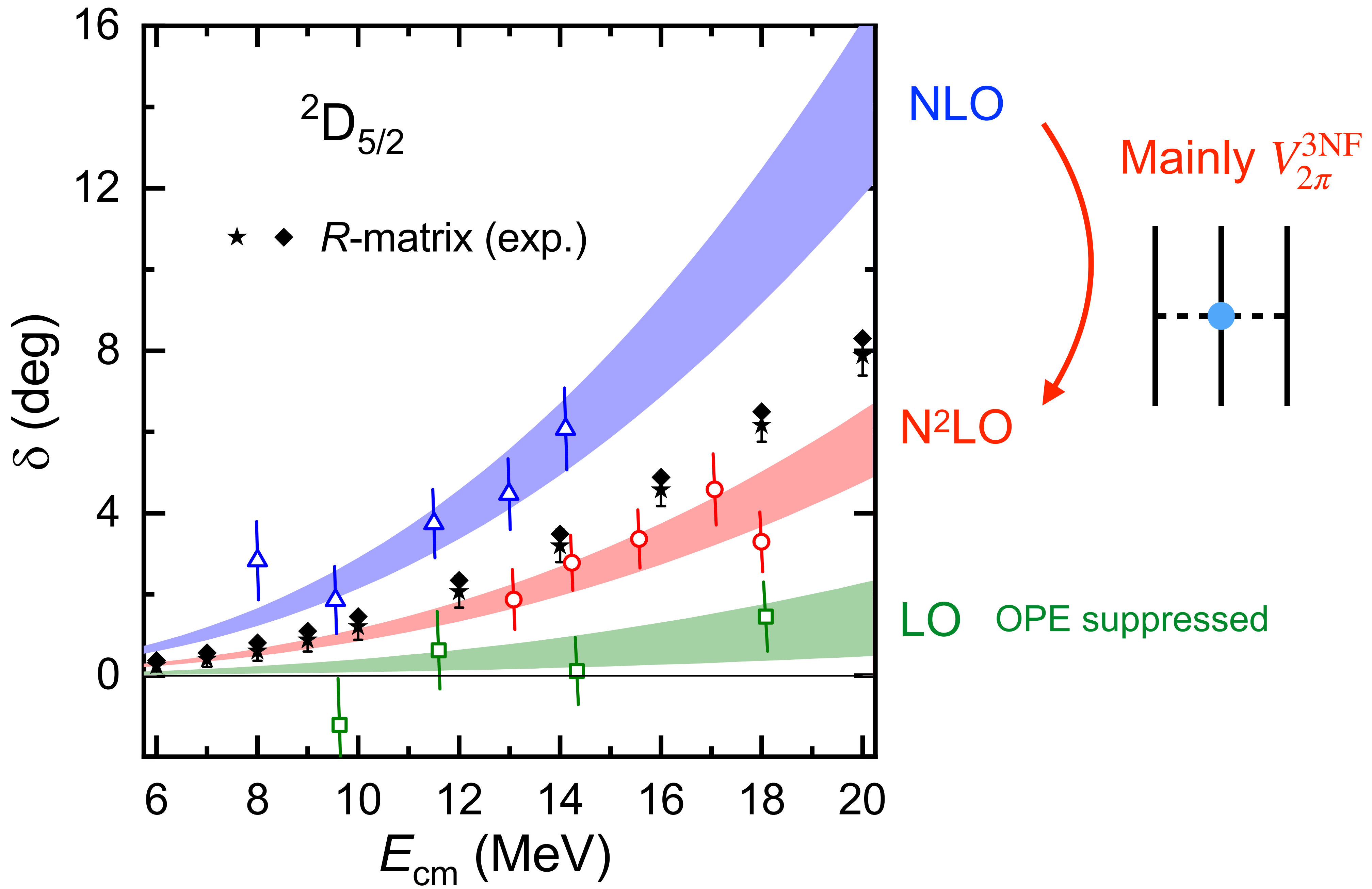
- The local N^{2}LO NN [Gezerlis2014PRC] +3N [Lynn2017PRC] forces with r-space cutoff $R = 1.2$ fm are employed.

- The BERW formula is benchmarked in NN scattering and the $n\alpha$ calculation is benchmarked with the existing P-wave calculations [Lynn2016PRL].

Impact of leading TPE 3NF

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

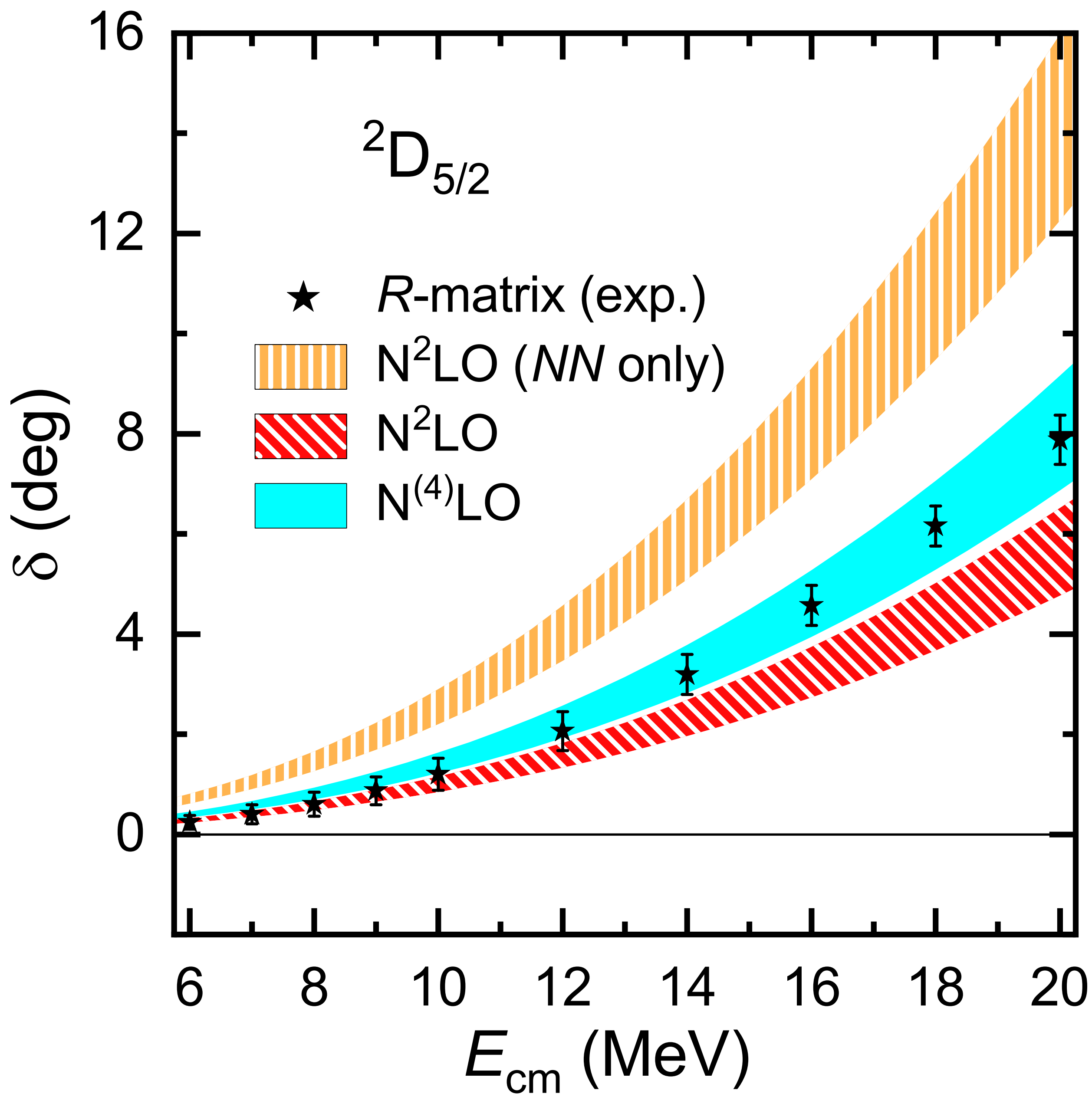
- Leading chiral TPE 3NF at N²LO provides a large repulsive contribution that improves the agreement with empirical D-wave $n\alpha$ phase shifts.



Impact of subleading corrections

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

- Peripheral $n\alpha$ scattering provides a sensitive probe to the long-range 3NF (governed by χ symmetry)



N²LO NN only

N²LO NN+3NF(effective c_i s)

$c_1 = -0.87,\ c_3 = -2.71,\ c_4 = 1.41$

N²LO NN+3NF

$c_1 = -0.81,\ c_3 = -3.40,\ c_4 = 3.40$

The N³LO and N⁴LO corrections effectively mimicked by reducing c_is

Krebs, Gasparyan, and Epelbaum PRC 85, 054006 (2012)

Resolving the Zemach radius puzzle in ^{6}Li

YLY, Evgeny Epelbaum, Chen Ji, Pengwei Zhao, in preparation

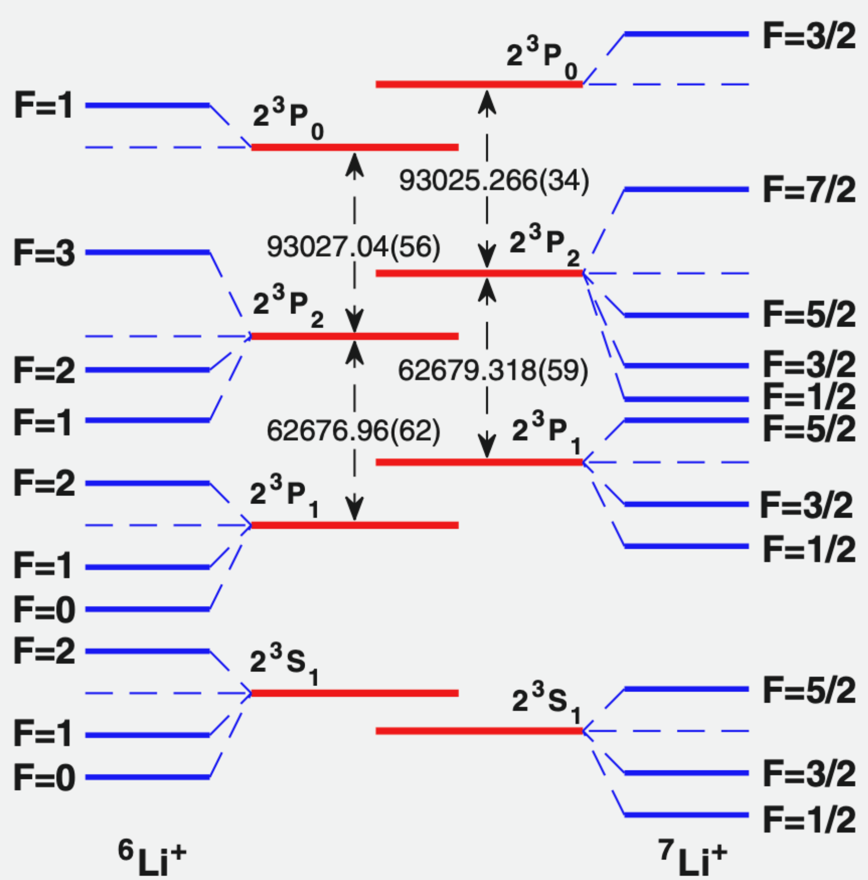
Zemach radius puzzle in ^{6}Li

- A large unexplained discrepancy between the two Zemach radius values of ^{6}Li obtained from the atomic hyperfine splitting and nuclear form factors.

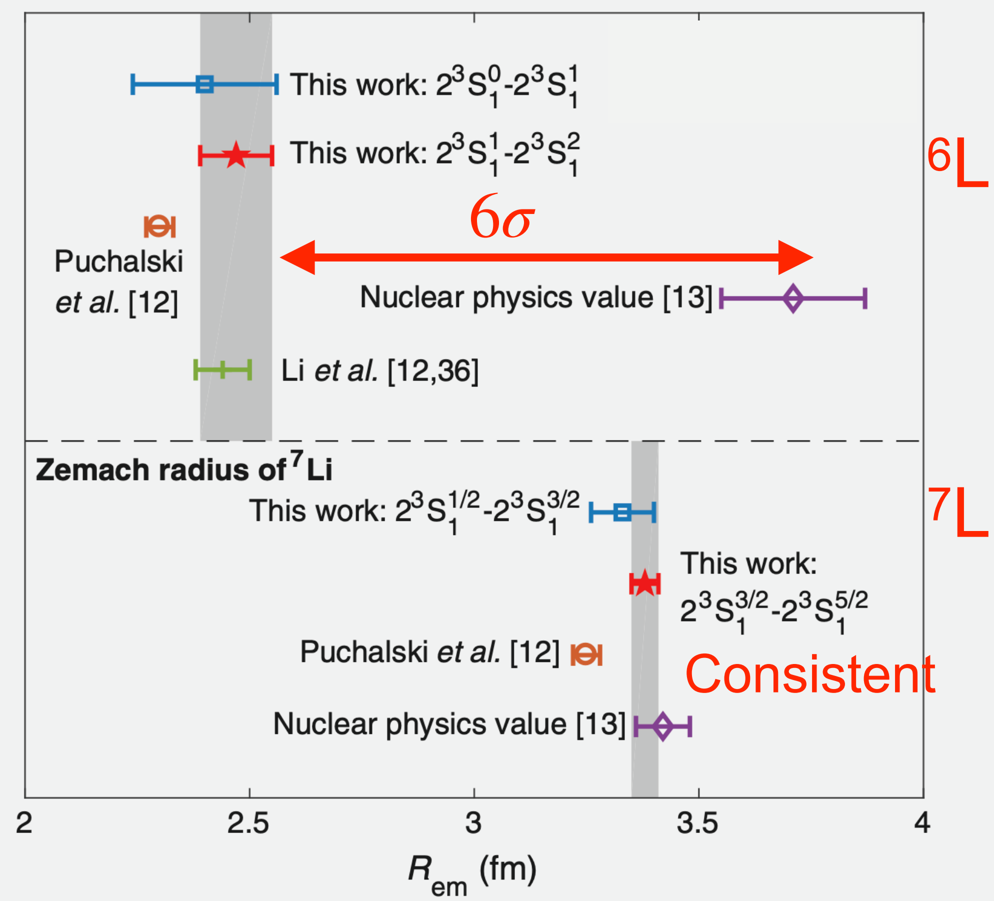
$$r_Z \simeq -\frac{1}{2Z\alpha m_e}\frac{E_{\rm expt}-E_{\rm QED}}{E_{\rm expt}}$$

$$r_Z = \int d^3r \int dr' \rho_E(\boldsymbol{r})\rho_M(\boldsymbol{r}')\,|\boldsymbol{r}-\boldsymbol{r}'|$$

Hyperfine splitting in Li ions



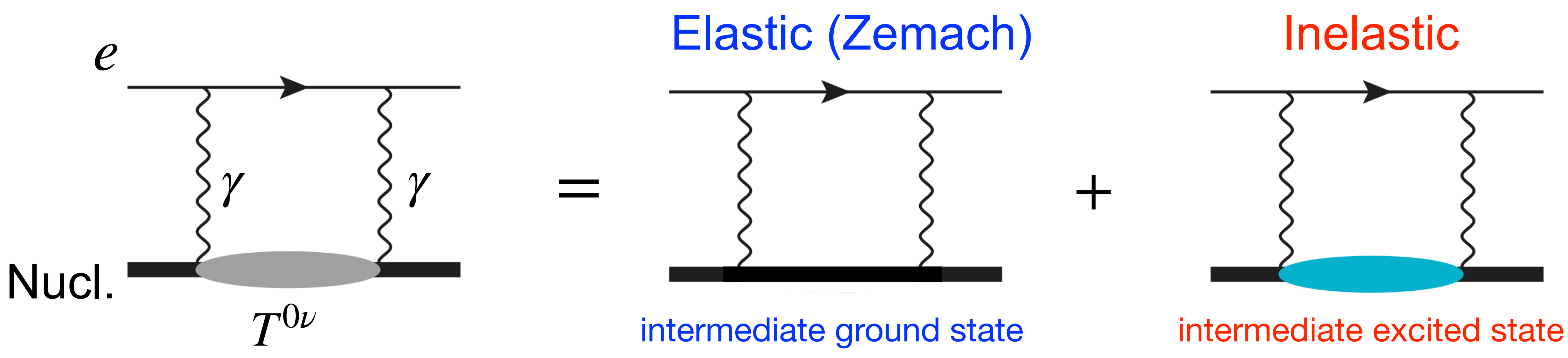
Extracted Zemach radius



Qi et al., PRL 125, 183002 (2020); Yerokhin et al., PRA 78, 012513 (2013)

Inelastic 2γ-exchange effects

- A probable explanation is the neglect of inelastic 2γ-exchange effects when extracting Zemach radius from hyperfine splitting.



- Theoretical calculations of inelastic contributions exist only for $A \leq 3$ nuclei

  ^{2,3}H and ^{3}He: Friar and Payne, PRC 72, 014002 (2005); ^{2}H: Ji, Zhang, and Platter, PRL 133, 042502 (2024)

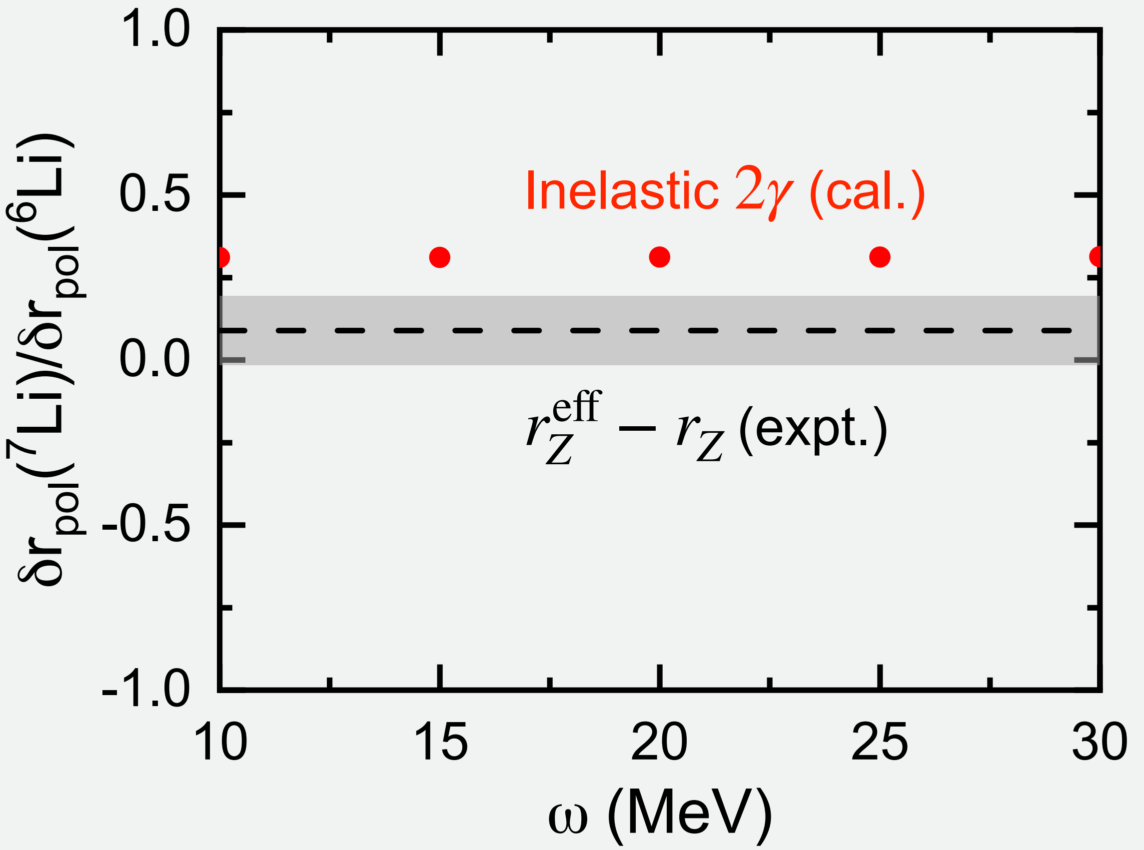
- We provide the first ab initio calculation of inelastic contributions of ^{6,7}Li using neural-network wave function, within closure approximation.

$$T^{0k}(q,-q) = \sum_N \frac{\langle 0|\rho(-\boldsymbol{q})|N\rangle\langle N|j^k(\boldsymbol{q})|0\rangle}{q_0 - \omega_N + \mathrm{i}\epsilon} + (\text{exch.}) \qquad \omega_N \to \overline{\omega}, \quad \sum_{N\neq 0} = 1 - |0\rangle\langle 0|$$

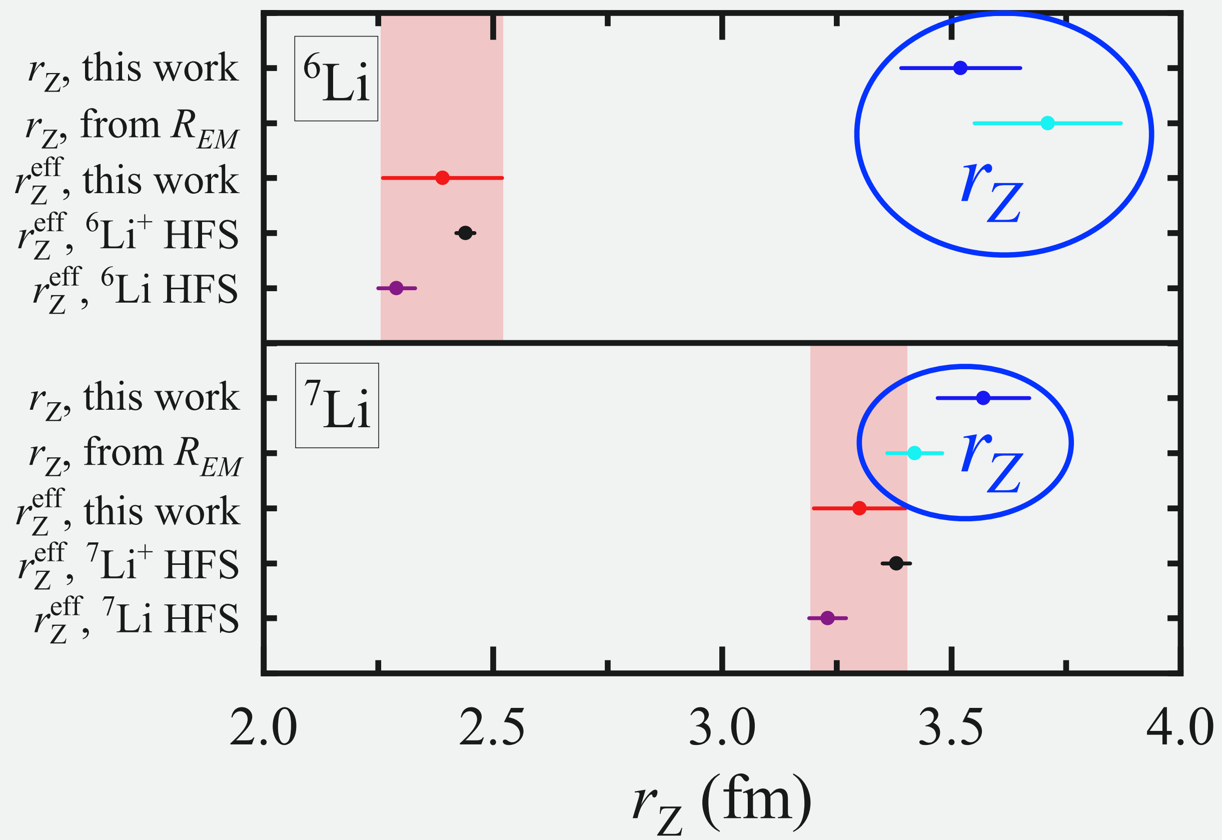
Resolving the Zemach radius puzzle in ^{6}Li

YLY, Evgeny Epelbaum, Chen Ji, Pengwei Zhao, in preparation

- Inelastic contributions smaller in ^{6}Li than in ^{7}Li, consistent with the observation (insensitive to $\overline{\omega}$)



- With reasonable closure energies, the r_Z^{eff} from HFS and r_Z from form factors can be both reproduced

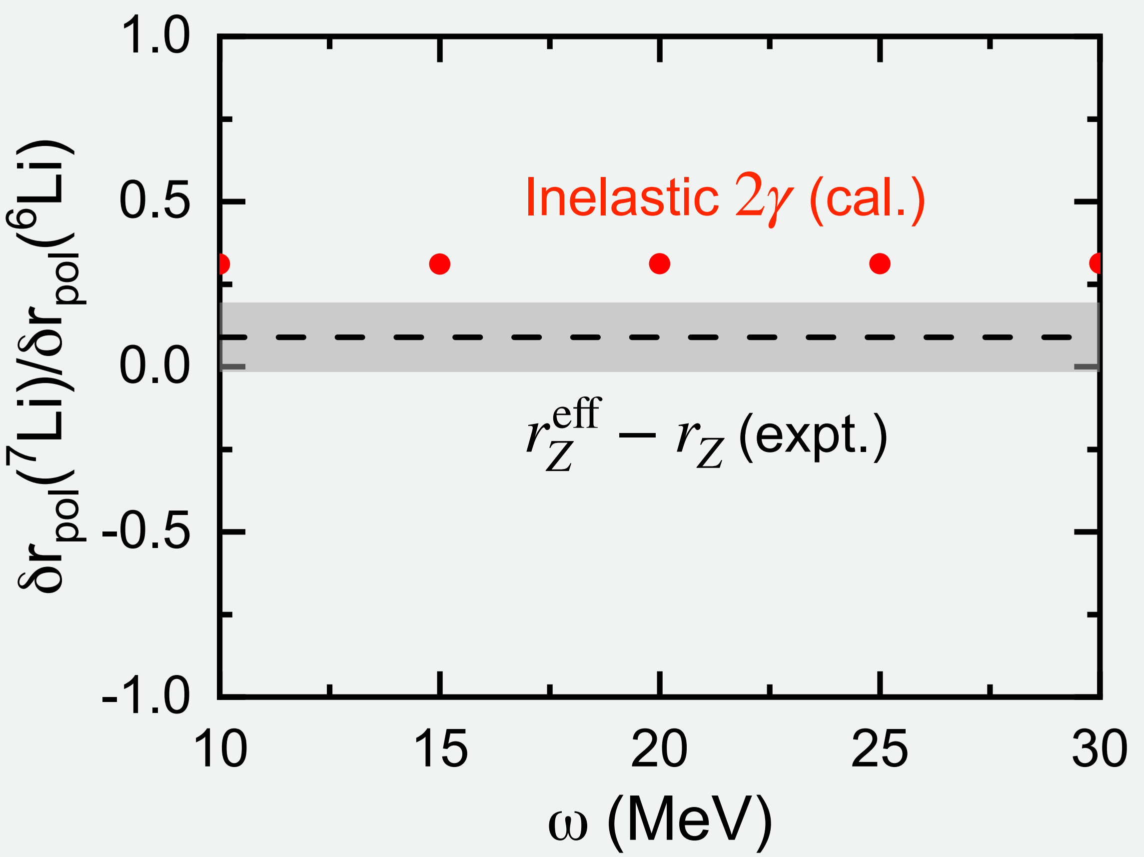


Strong evidence that inelastic 2γ-exchange effects are responsible for the observed discrepancy between r_Z^{eff} from hyperfine splitting (HFS) and r_Z from nuclear form factors.

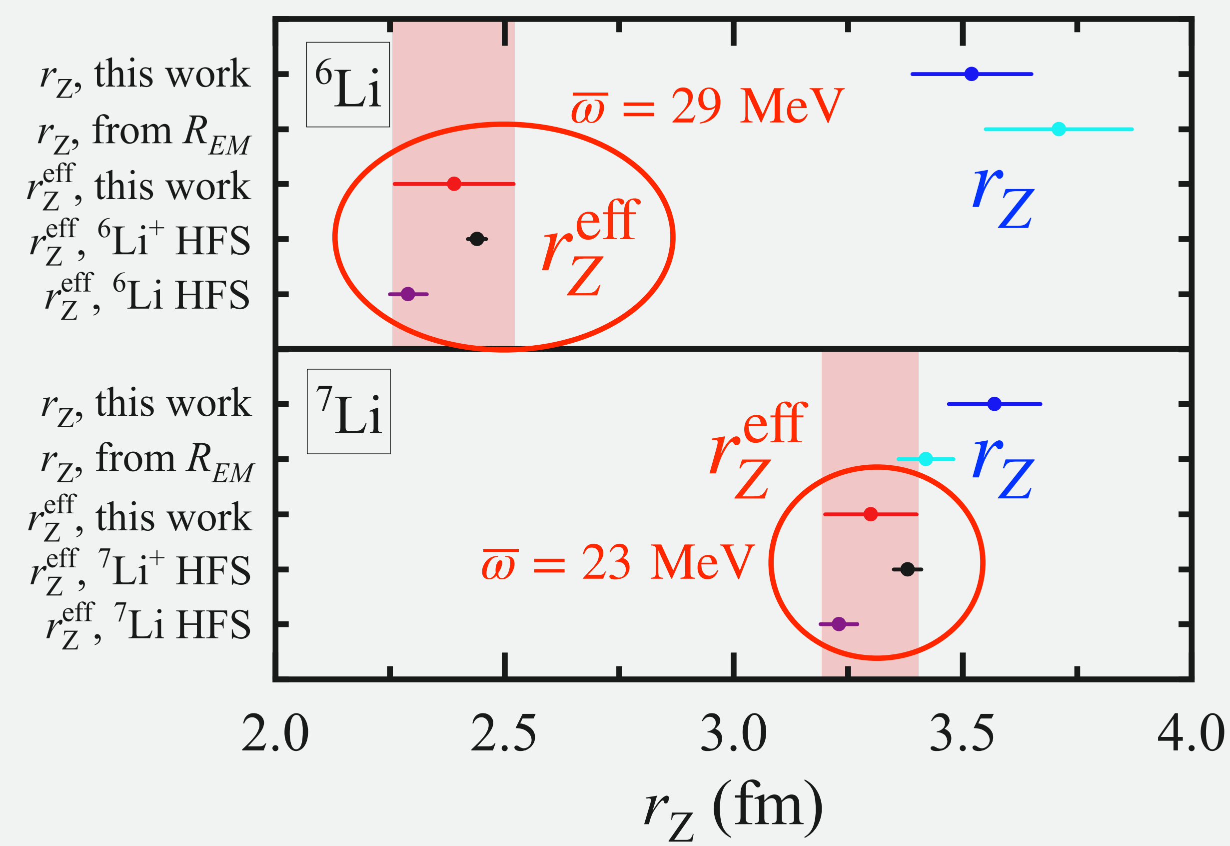
Resolving the Zemach radius puzzle in ^{6}Li

YLY, Evgeny Epelbaum, Chen Ji, Pengwei Zhao, in preparation

- Inelastic contributions smaller in ^{6}Li than in ^{7}Li, consistent with the observation (insensitive to $\overline{\omega}$)



- With reasonable closure energies, the $r_Z^{\rm eff}$ from HFS and r_Z from form factors can be both reproduced



Strong evidence that inelastic 2γ-exchange effects are responsible for the observed discrepancy between $r_Z^{\rm eff}$ from hyperfine splitting (HFS) and r_Z from nuclear form factors.

Summary and outlooks

Neural-network quantum Monte Carlo: A new accurate ab initio many-body method for studying nuclear and electroweak properties

- Achieve high accuracy of ground state variantionally

  ✓ $A \lesssim 16$ nuclei with pionless EFT

  ✓ $A \lesssim 7$ nuclei with chiral EFT

- Recent applications

  ✓ Probing long-range 3NF in peripheral $n\alpha$ scattering

  ✓ Solving the Zemach radius puzzle in ^{6}Li

➡ Towards heavier nuclei with more accurate nuclear interactions

➡ Excited states, resonances, real-time dynamics, …

Acknowledgement

- Special thanks to my supervisors and collaborators

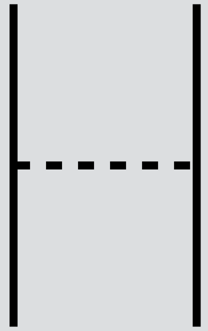
Jie Meng, Pengwei Zhao	Peking University
Evgeny Epelbaum, Lu Meng	Ruhr University Bochum
Chen Ji	Central China Normal University

Thank you for your attention!

Appendix

Chiral symmetry in nuclear force

Chiral symmetry + πN data = predictions for the large-distance behavior of the nuclear forces

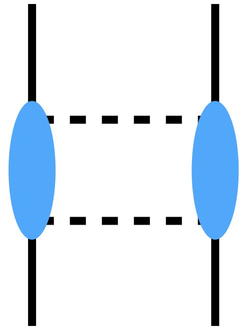


$$\mathscr{L}_{pv} = -\frac{g}{2M}\overline{N}\gamma_5\gamma^\mu\tau N\cdot\partial_\mu\pi$$

~~$\mathscr{L}_{ps} = -g\overline{N}i\gamma_5\tau N\cdot\pi$~~

$\Big\} \Rightarrow$ Same 1π exchange (on-shell)

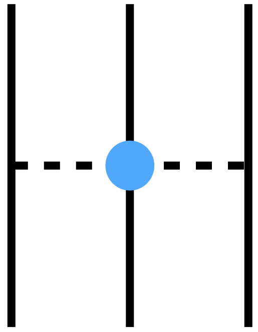
i.e., ***NOT*** contrained by χ symmetry



2π exchange strongly constrained by χ symmetry

$\mathscr{L}_{pv}$ vs $\mathscr{L}_{ps}$ matters; $\pi\pi$, $\pi\pi N$ interactions enter (fixed in πN systems)

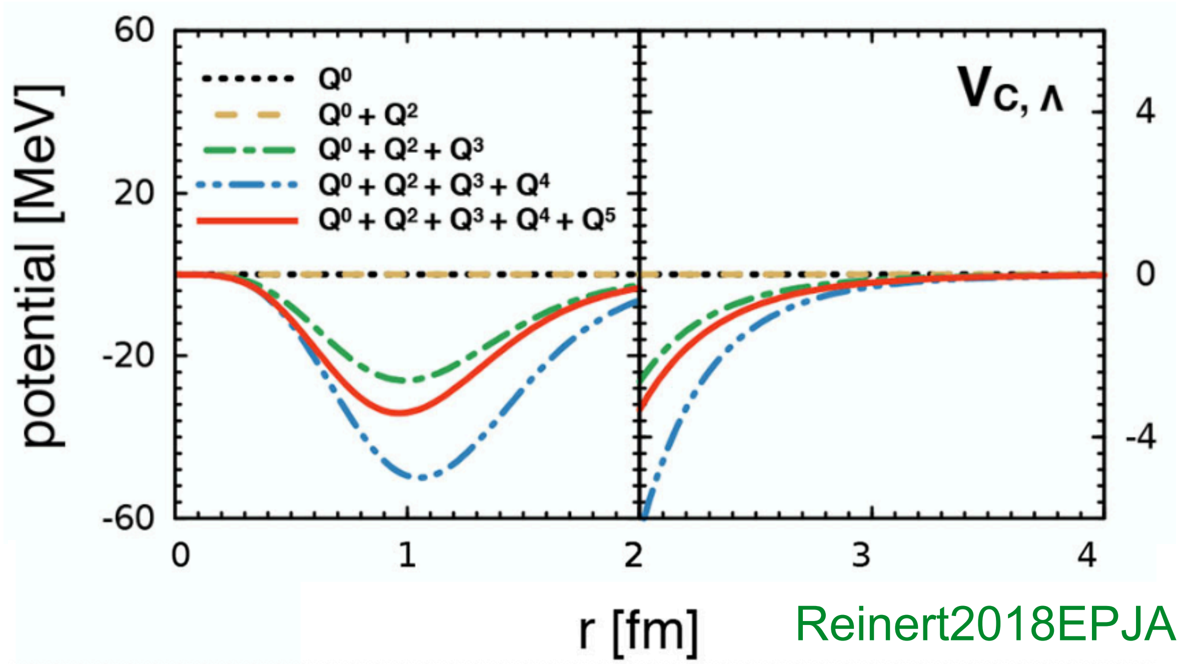
predictions for long-range NN and 3N force



e.g., a N²LO contribution

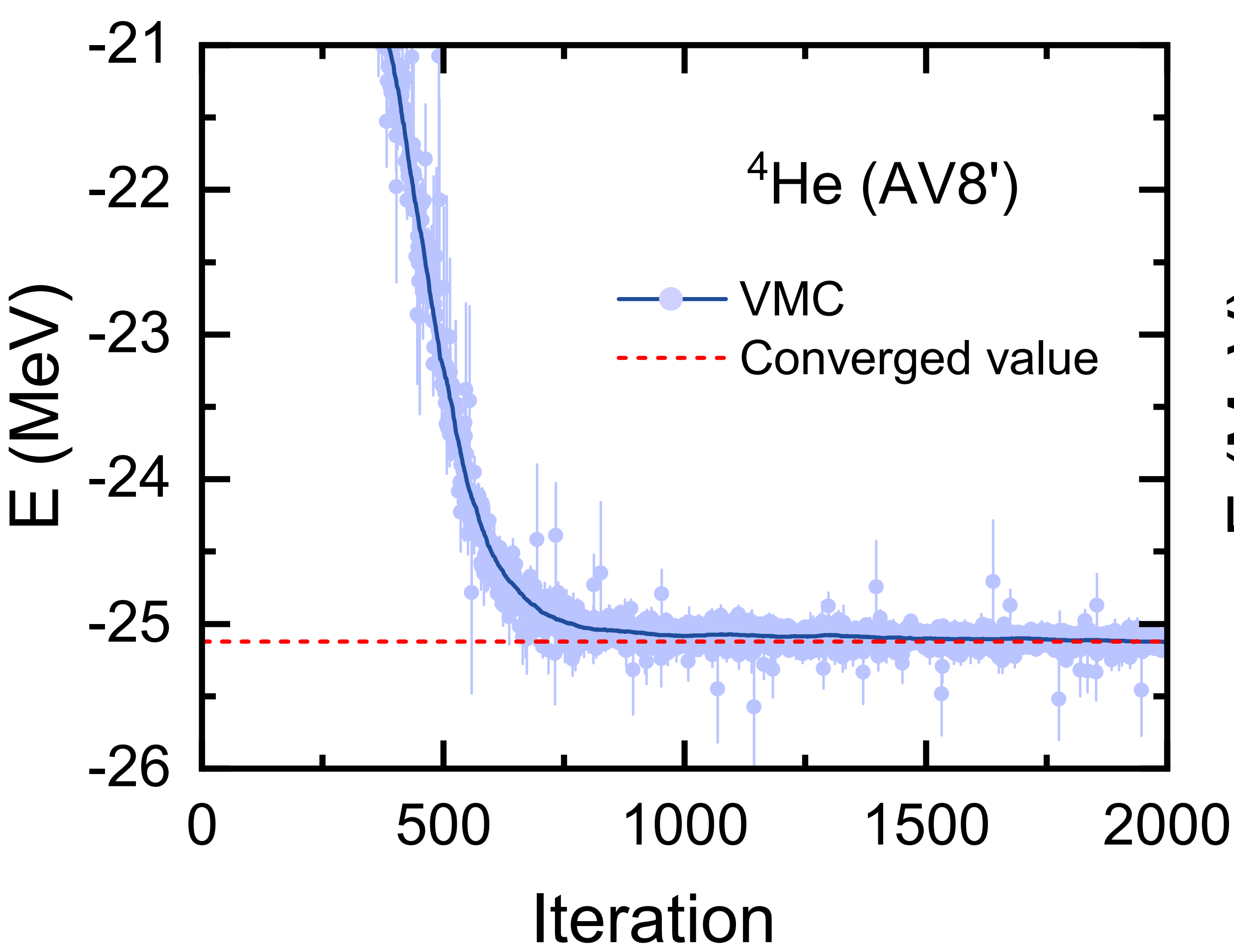
$$V_C^{(3)}(r) = c_1\frac{3m_\pi^6 g_A^2}{16\pi^2 f_\pi^4}\frac{e^{-2x}}{x^4}(1+x^2)$$

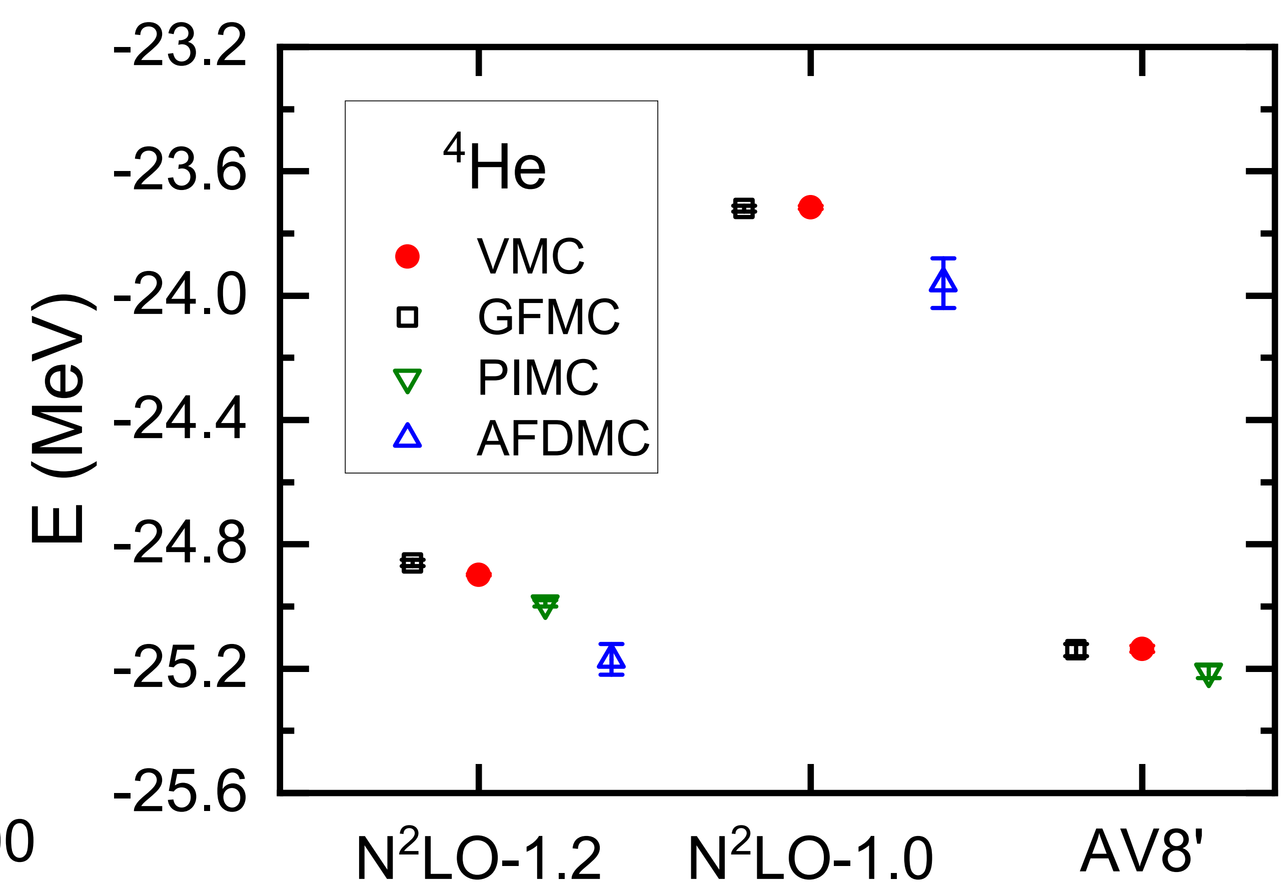
$$x = m_\pi r$$



Reinert2018EPJA

^{4}He calculated with several NN forces





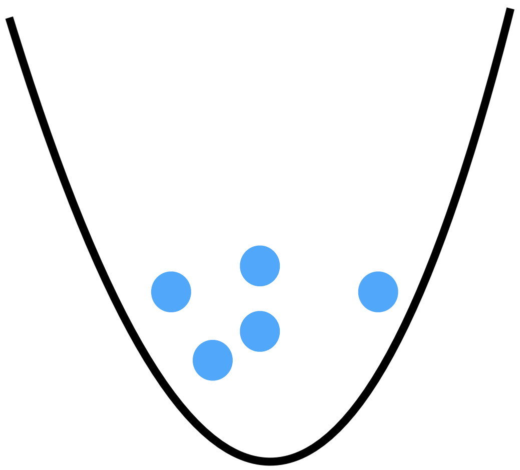
QMC + BERW formula for $n\alpha$ phase shifts

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

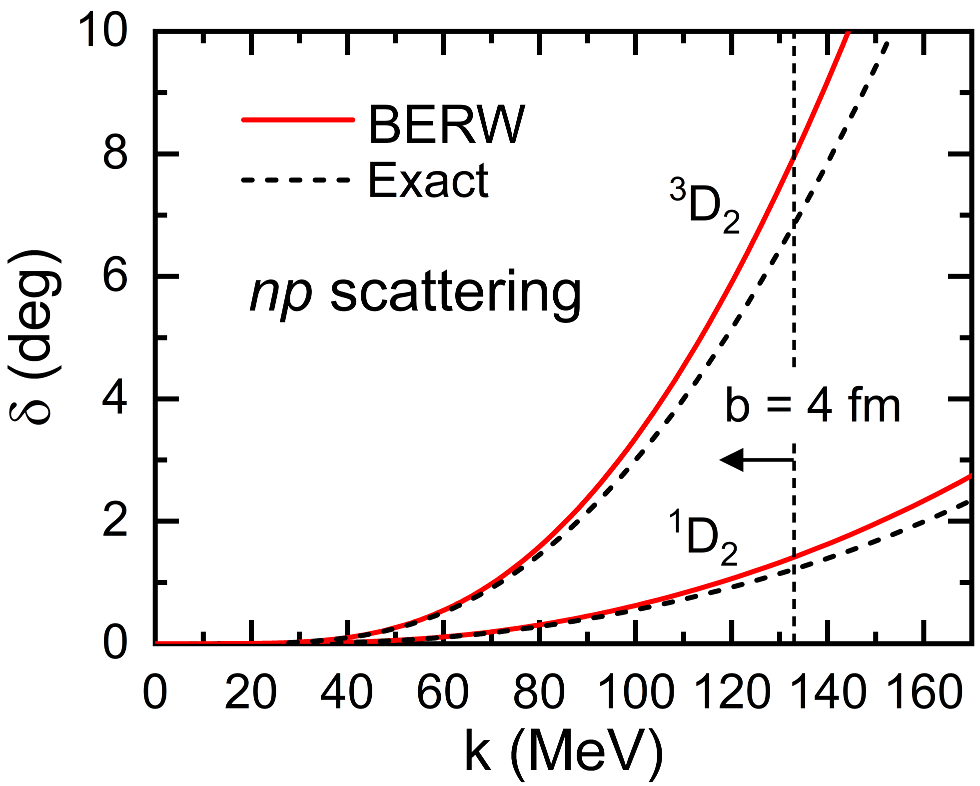
- $n\alpha$ phase shifts are extracted from the $^5\mathrm{He}_l$ energy in a harmonic oscillator trap

$$k^{2l+1}\cot\delta_l^{n\alpha}(k) = (-1)^{l+1}\left(4\mu\omega\right)^{l+1/2}\frac{\Gamma((3+2l)/4-\epsilon_l/(2\omega))}{\Gamma((1-2l)/4-\epsilon_l/(2\omega))}$$

Busch et al., Found. Phys. 28, 549 (2008); Suzuki et al., PRA 80, 033601 (2009)



with $\epsilon_l = E(^5\mathrm{He}_l) - E_\alpha$. We focus on the $D_{5/2}$ wave (spin-orbit splitting between $D_{5/2}$ and $D_{3/2}$ at low energies are small).



- The oscillator lengths $b > 4$ fm are used.

  much larger than interaction range $m_\pi b \geq 3$

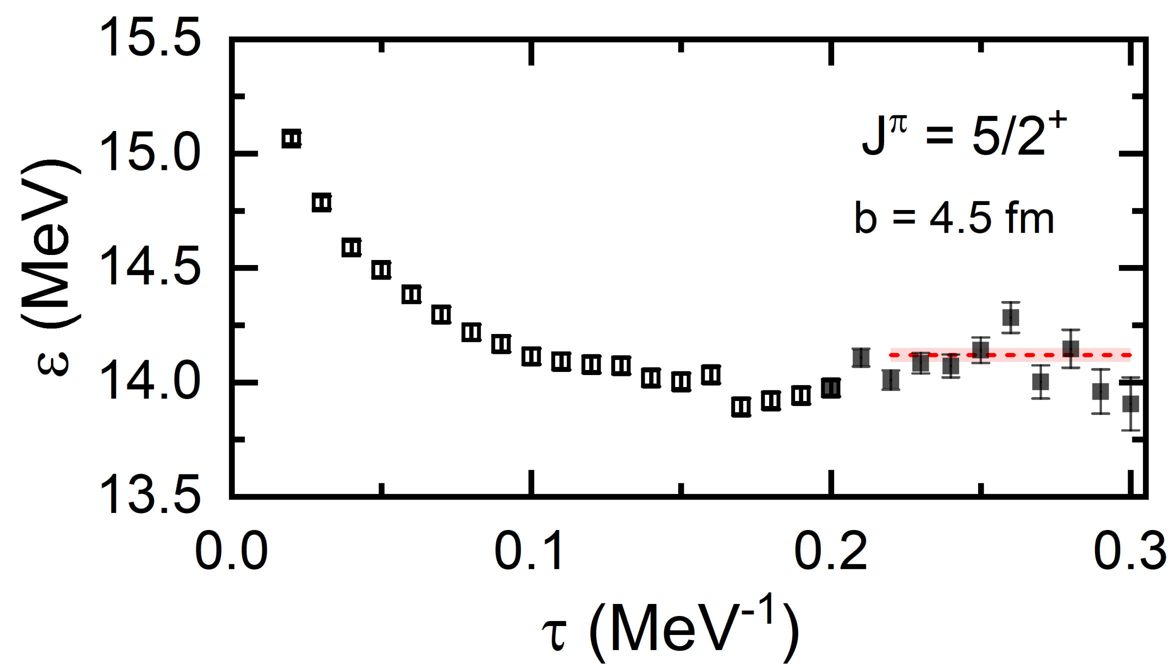
- Benchmarked in D-wave NN scattering.

  systematic error within $\sim$10% at low energies.

QMC + BERW formula for $n\alpha$ phase shifts

YLY, Evgeny Epelbaum, Jie Meng, Lu Meng, and Pengwei Zhao, e-Print: 2502.9961

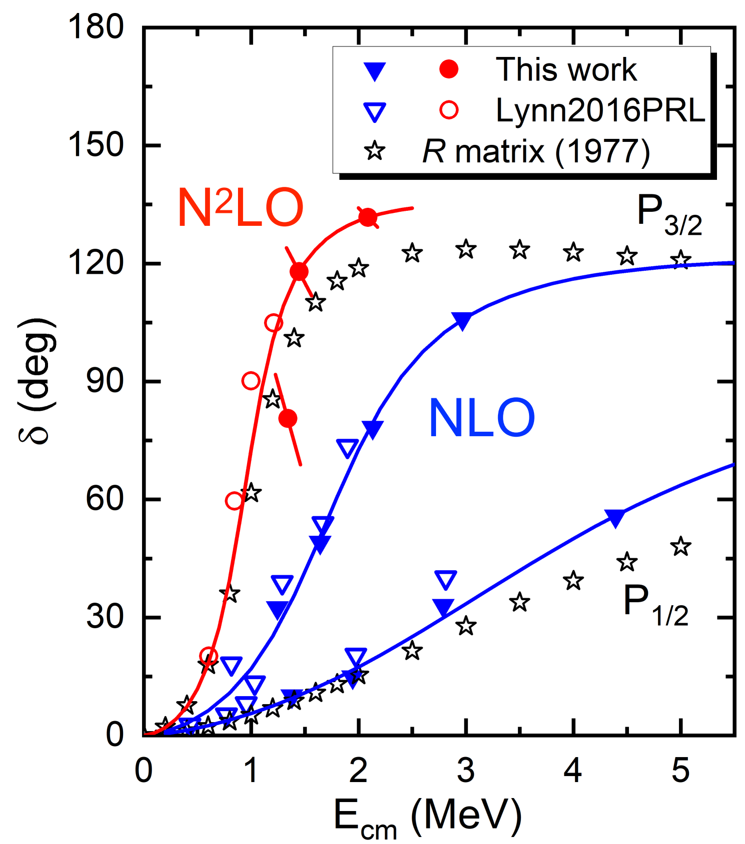
- The trapped ^{5}He energy is calculated with neural-network VMC+GFMC



Here, the trial function includes mainly two-body correlations, so GFMC is necessary.

- The local N²LO NN [Gezerlis2014PRC] +3N [Lynn2017PRC] forces with the softest r space cutoff $R = 1.2$ fm are employed. (due to the sign problem; harder cutoff can be made possible by the new neural-network wave function with many-body correlations)

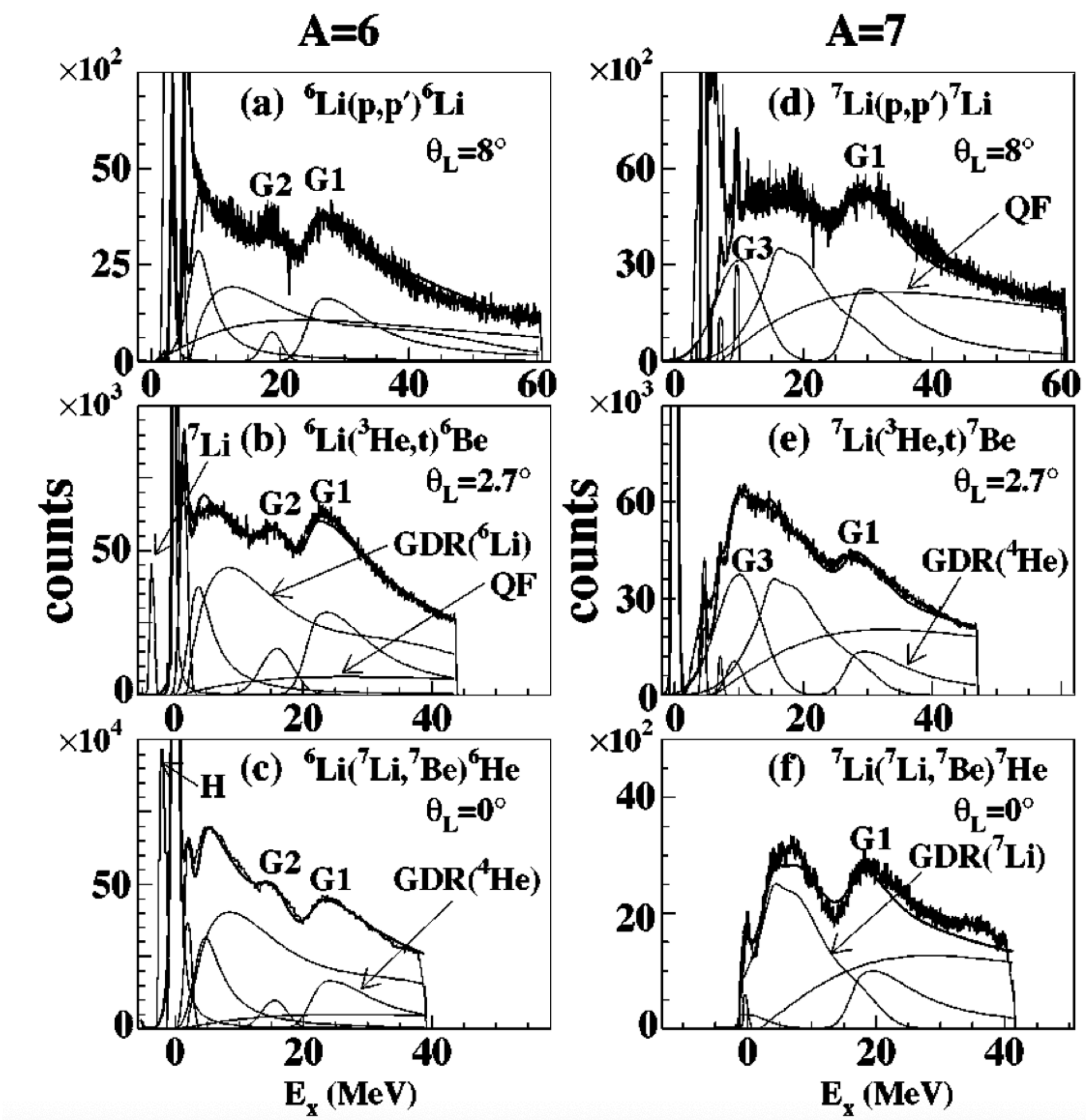
- $n\alpha$ caluclations for P-wave benchmarked against [Lynn2016PRL]

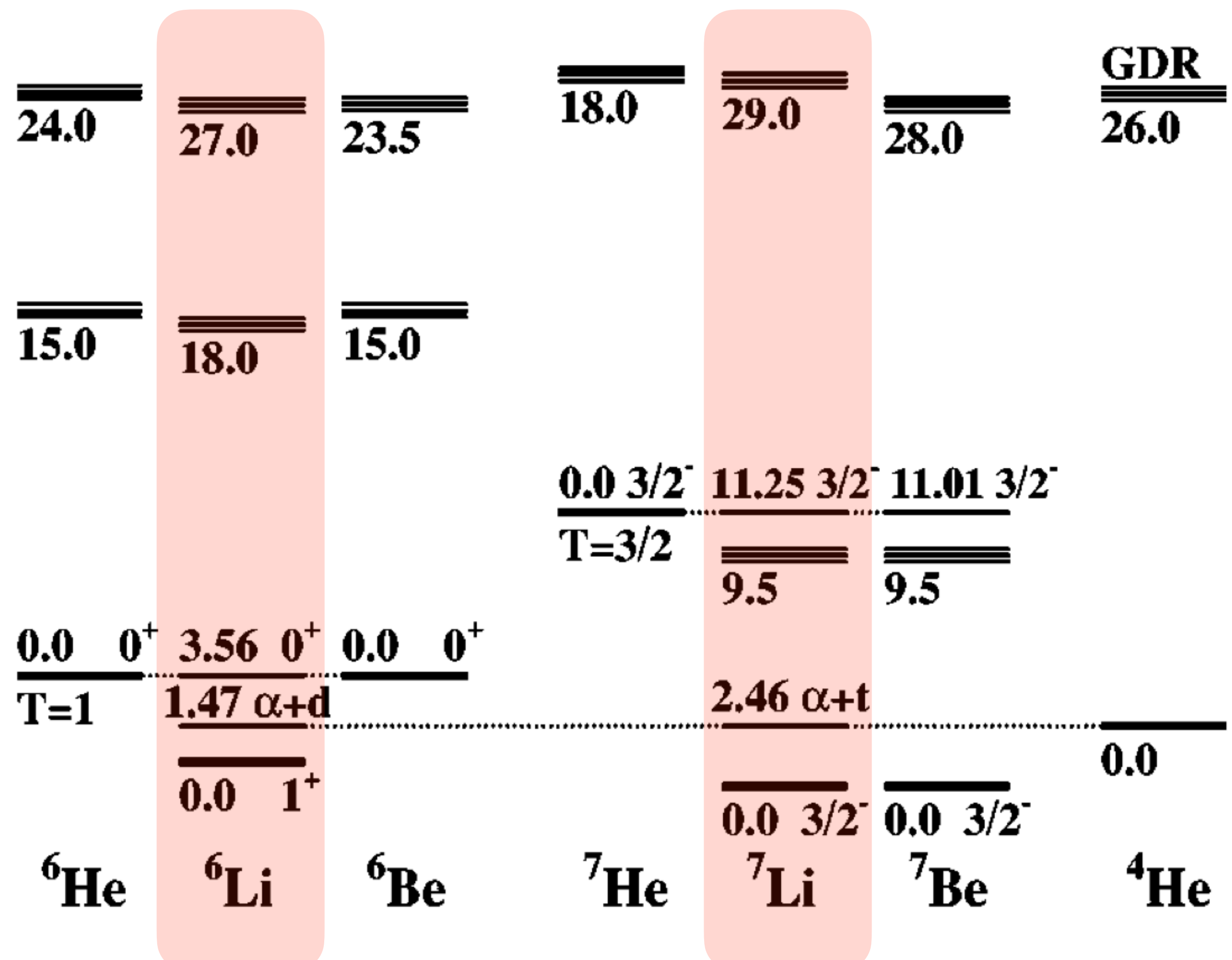


Excited spectrum in ^{6,7}Li

Excitations of the cluster in *A*=6 and 7 nuclei

T. Yamagata et al., PHYSICAL REVIEW C **69**, 044313 (2004)





Unit: MeV