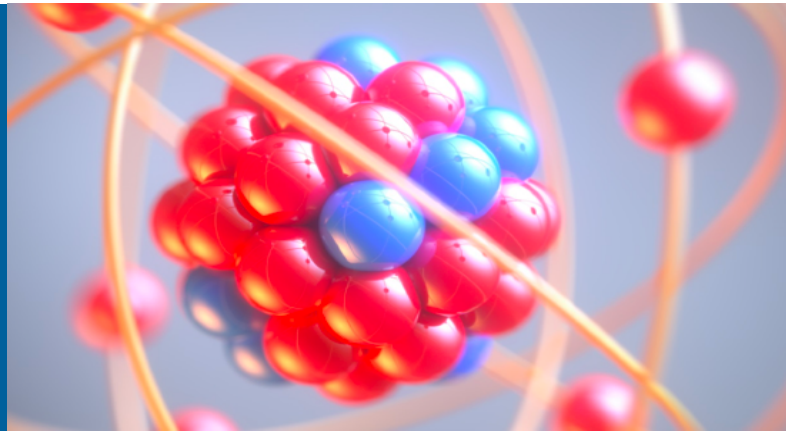
SOLVING THE QUANTUM MANY-BODY PROBLEM WITH NEURAL NETWORKS



ALESSANDRO LOVATO

TIFPA Trento Institute for Fundamental Physics and Applications

Argonne NATIONAL LABORATORY

Information and Statistics for Nuclear Experiment and Theory workshop (ISNET-9)

May 24, 2023

COLLABORATORS

C. Adams, **B. Fore, B. Hall**

G. Carleo, **G. Pescia**

J. Kim, M. Hjorth-Jensen

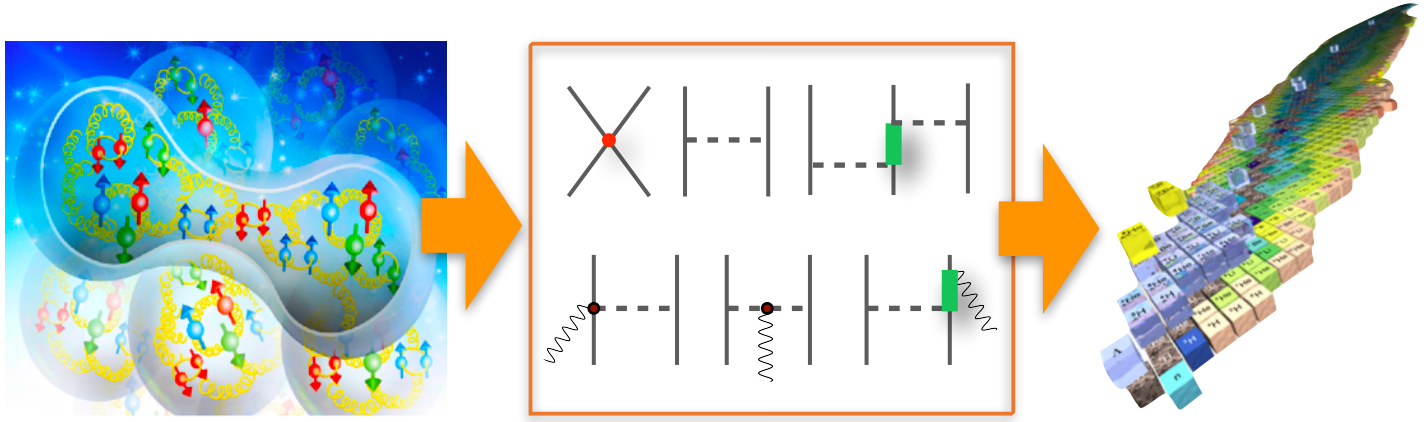
F. Pederiva, **M. Rigo**

N. Rocco

THE NUCLEAR MANY-BODY PROBLEM

In the low-energy regime, quark and gluons are confined within hadrons and the relevant degrees of freedoms are protons, neutrons, and pions

Effective field theories are the link between QCD and nuclear observables.

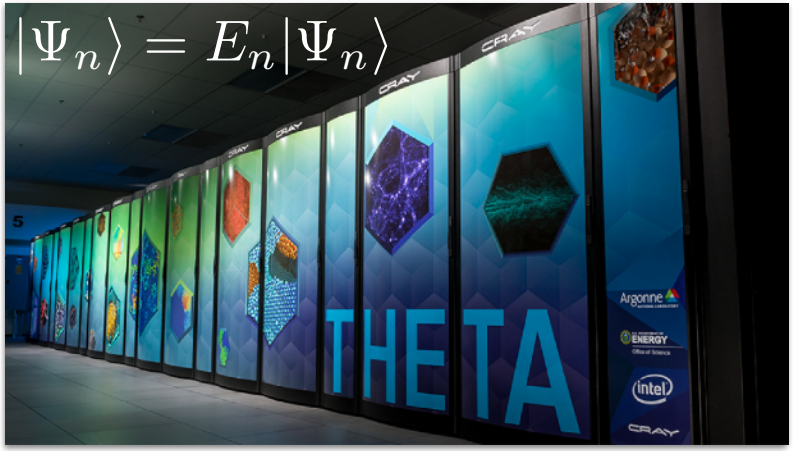


$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk}$$

$$J = \sum_i j_i + \sum_{i<j} j_{ij}$$

NUCLEAR MANY-BODY METHODS

- Hamiltonians and consistent currents are the main inputs to nuclear many-body methods
- These methods capitalize on high-performance computers to solve the Schrödinger equation with controlled approximation



- Nuclear many-body calculations are continually battling against the “curse of dimensionality,” the rapid growth with complexity of computational resources needed.

CONFIGURATION-INTERACTION

The exact ground-state wave function can be expressed as a sum of Slater determinants

$$\Psi_0(x_1,\ldots,x_A) = \sum_n c_n \Phi_n(x_1,\ldots,x_A)$$

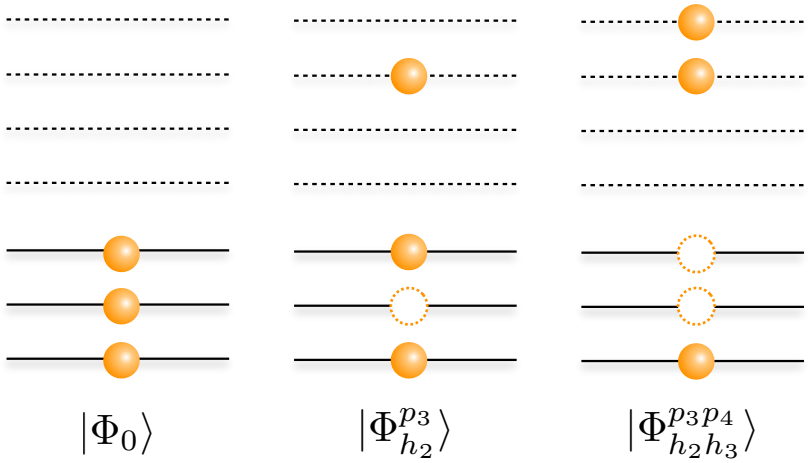
The occupation-number representation automatically encompasses fermion antisymmetry

$$|\Psi_0\rangle = \sum_{h_1\ldots,p_1\ldots} c^{p_1\cdots}_{h_1\ldots} |\Phi^{p_1\cdots}_{h_1\ldots}\rangle$$

$$|\Phi^{p_1\cdots}_{h_1\ldots}\rangle = a^\dagger_{p_1}\ldots a_{h_1}\ldots|\Phi_0\rangle$$

The dimensionality explodes quickly

$$\binom{N}{A} = \frac{N!}{(N-A)!A!}$$



CONTINUUM QUANTUM MONTE CARLO

The trial wave function can be expanded in the set of the Hamiltonian eigenstates

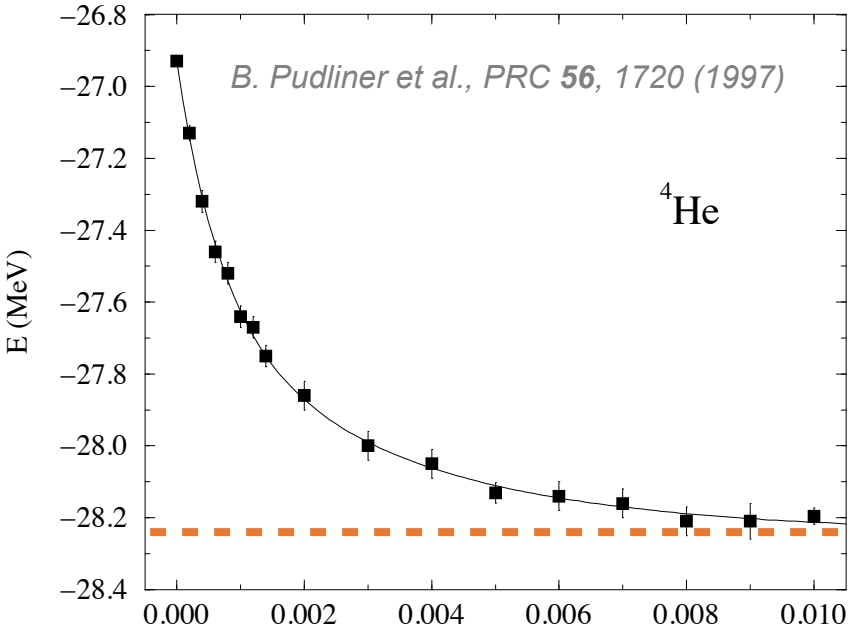
$$|\Psi_T\rangle = \sum_n c_n |\Psi_n\rangle \qquad H|\Psi_n\rangle = E_n|\Psi_n\rangle$$

GFMC relies on imaginary-time propagation

$$\lim_{\tau\to\infty} e^{-(H-E_0)\tau}|\Psi_T\rangle = c_0|\Psi_0\rangle$$

*J. Carlson Phys. Rev. C **36**, 2026 (1987)*

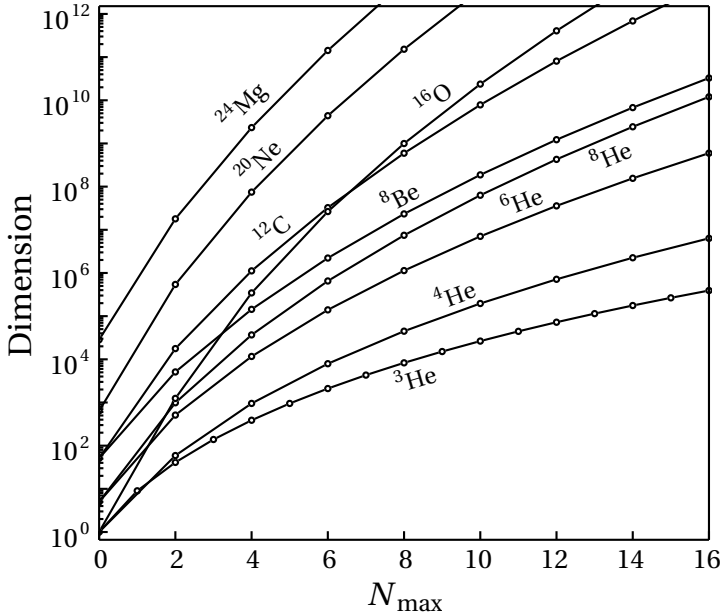
GFMC suffers from the fermion-sign problem, but it is “virtually exact” for light nuclear systems.



COURSE OF DIMENSIONALITY

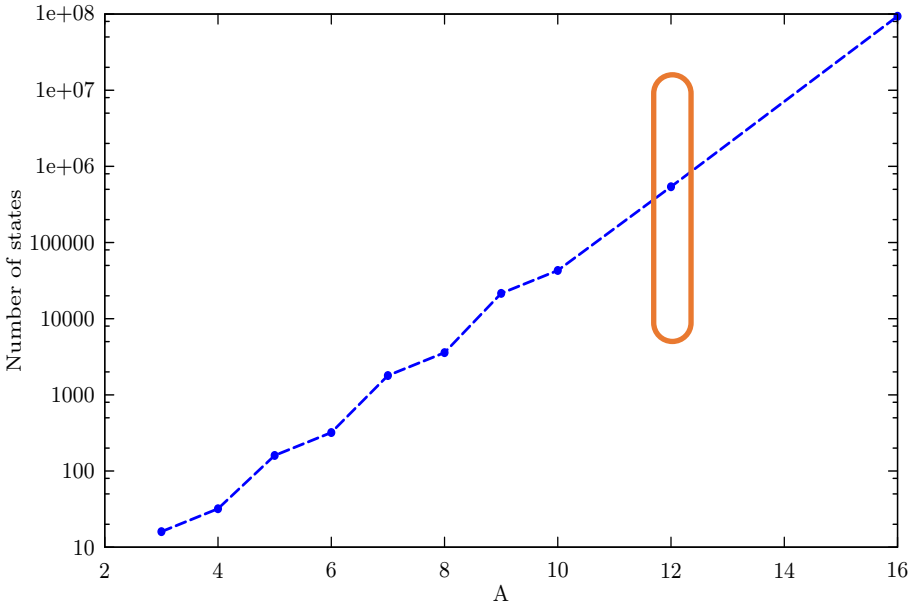
Configuration-Interaction

$$\Psi_0(x_1,\ldots,x_A) = \sum_n c_n \Phi_n(x_1,\ldots,x_A)$$



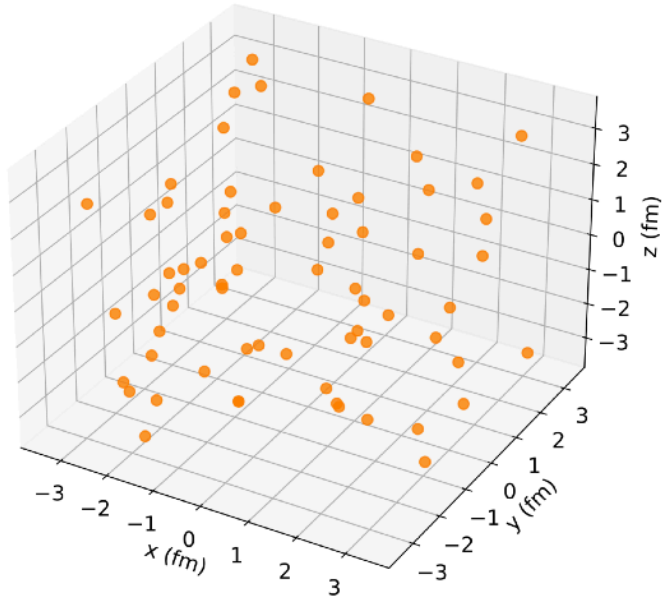
Green's function Monte Carlo

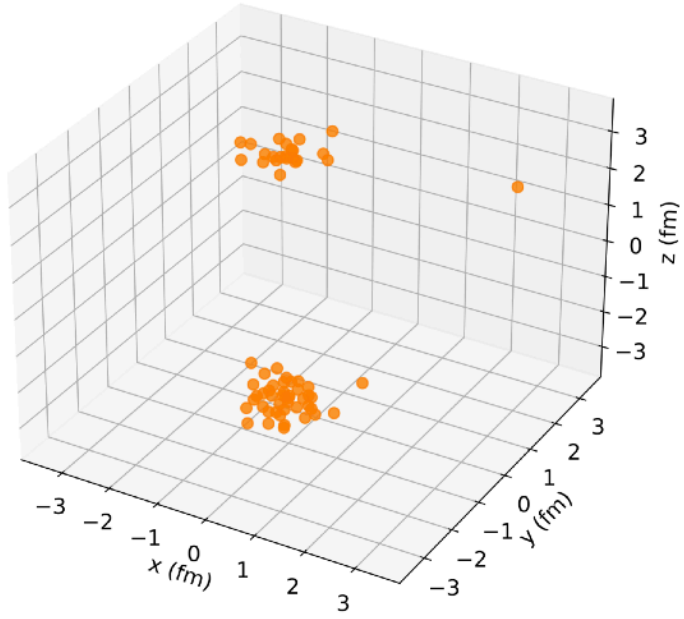
$$\lim_{\tau\to\infty} e^{-(H-E_0)\tau}\left|\Psi_T\right\rangle = c_0\left|\Psi_0\right\rangle$$



“ERRORS” IN MANY-BODY METHODS

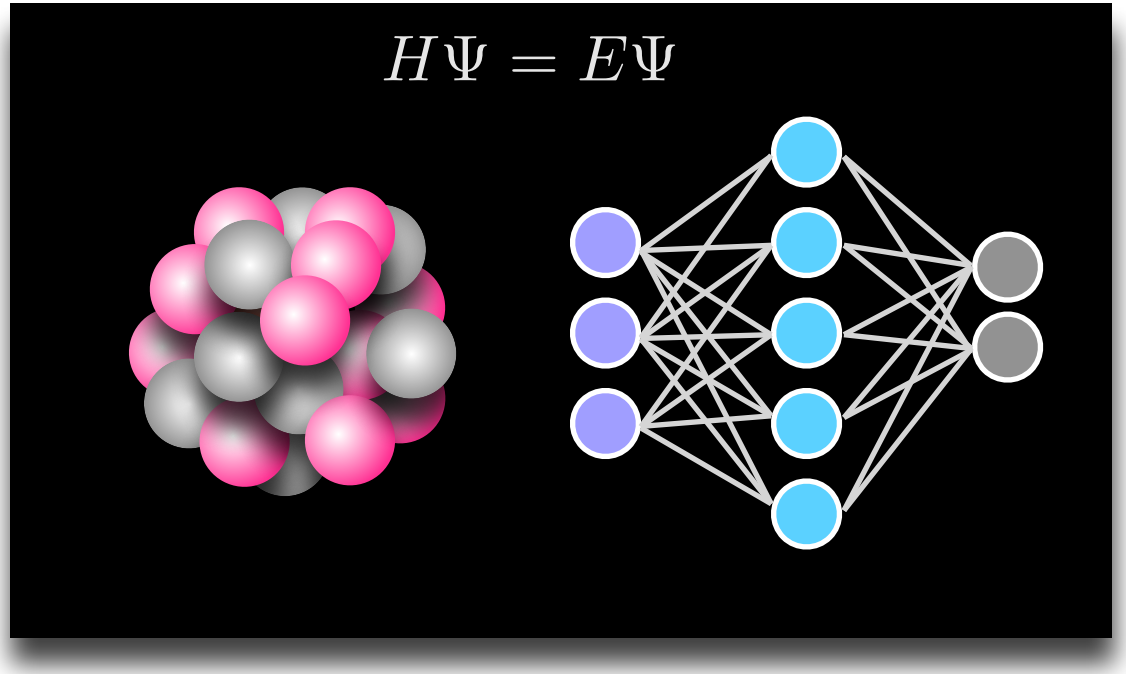
It is surprisingly easy to underestimate the errors we make in solving the many-body problem





*A. Lovato et al., PRC **105** (2022) 5, 055808*

NEURAL NETWORK QUANTUM STATES



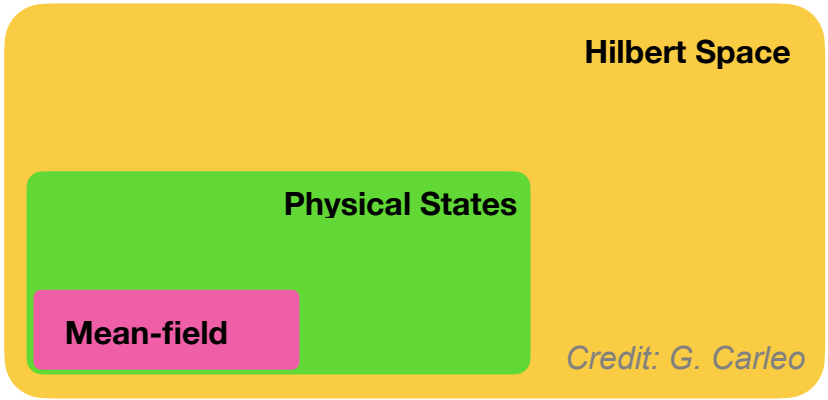
NEURAL-NETWORK QUANTUM STATES

Let's take a step back: spin problem  $H_{TIF} = -h \sum_i \sigma_i^x - \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z$

Finding the exact solution of this equation is, in principle, an **exponentially hard problem**

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

The majority of quantum states of physical interest have distinctive features and intrinsic structures



NEURAL-NETWORK QUANTUM STATES

$$\begin{cases} c_{\uparrow\uparrow\uparrow\ldots} \equiv \langle \uparrow\uparrow\uparrow \ldots | \Psi \rangle \equiv \Psi(\uparrow\uparrow\uparrow \ldots) \\ c_{\downarrow\uparrow\uparrow\ldots} \equiv \langle \downarrow\uparrow\uparrow \ldots | \Psi \rangle \equiv \Psi(\downarrow\uparrow\uparrow \ldots) \\ \\ c_{\downarrow\downarrow\downarrow\ldots} \equiv \langle \downarrow\downarrow\downarrow \ldots | \Psi \rangle \equiv \Psi(\downarrow\downarrow\downarrow \ldots) \end{cases} \quad \longleftrightarrow \quad c_S \equiv \langle S|\Psi\rangle \equiv \Psi(S)$$

Artificial neural networks (ANNs) can compactly represent complex high-dimensional functions;

$$\Psi(S) \simeq \langle S|\hat{\Psi}(\mathcal{W})\rangle \equiv \hat{\Psi}(S;\mathcal{W})$$

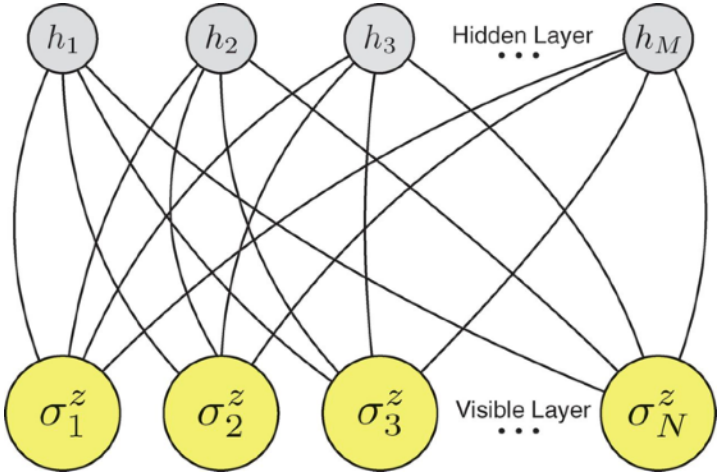
ANNs trained minimizing the energy, which is evaluated stochastically

$$E(\mathcal{W}) = \frac{\langle \hat{\Psi}(W)|H|\hat{\Psi}(W)\rangle}{\langle \hat{\Psi}(W)|\hat{\Psi}(W)\rangle} \simeq \sum_{S_n} \frac{\langle S_n|H|\hat{\Psi}(W)\rangle}{\langle S_n|\hat{\Psi}(W)\rangle} \qquad P(S_n) = |\langle S_n|\hat{\Psi}(W)\rangle|^2$$

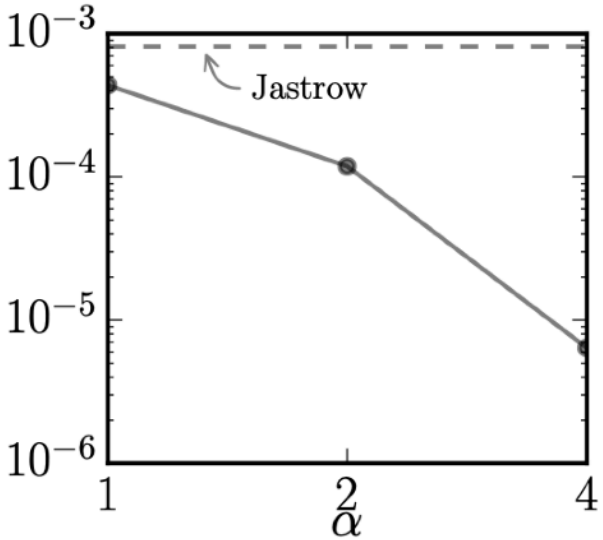
NEURAL-NETWORK QUANTUM STATES

Giuseppe Carleo and Mathias Troyer demonstrated that RMBs outperform traditional Jastrows

$$\hat{\Psi}(S;\mathcal{W}) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$



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

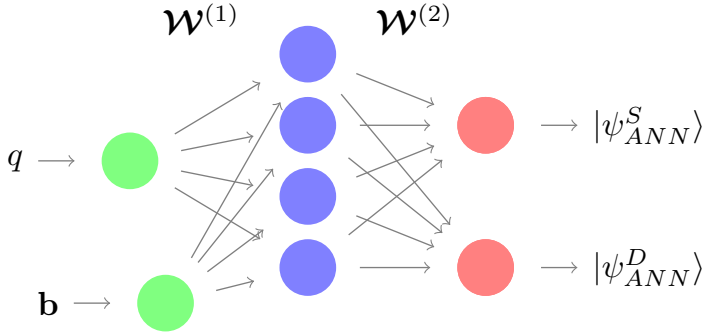


*G. Carleo et al. Science **355**, 602 (2017)*

MACHINE-LEARNING THE DEUTERON

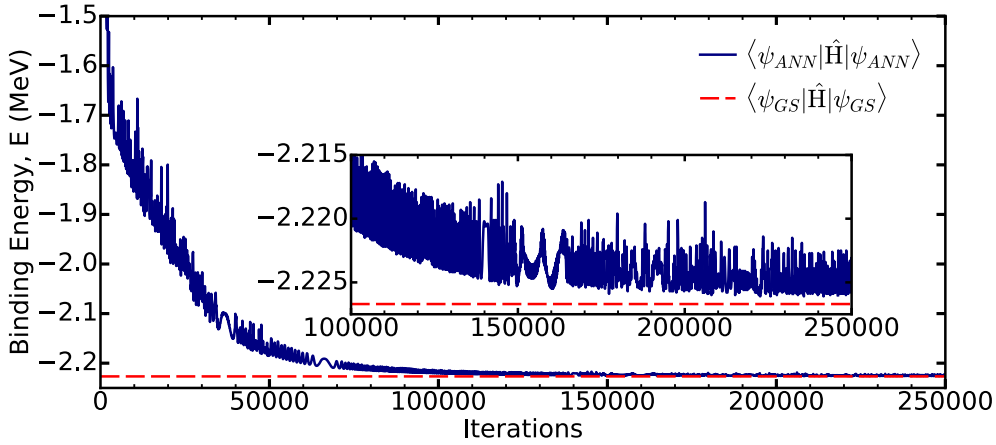
ANNs were recently applied to solve the deuteron in momentum space using the N3LO Entem-Machleidt chiral-EFT nucleon-nucleon force

*Keeble, Rios, PLB **809**, 135743 (2020)*



The parameters of the ANN are optimized minimizing the variational energy using RMSprop

$$E^{\mathcal{W}} = \frac{\left\langle \Psi^{\mathcal{W}}_{\mathrm{ANN}} \middle| \hat{H} \middle| \Psi^{\mathcal{W}}_{\mathrm{ANN}} \right\rangle}{\left\langle \Psi^{\mathcal{W}}_{\mathrm{ANN}} \middle| \Psi^{\mathcal{W}}_{\mathrm{ANN}} \right\rangle}$$

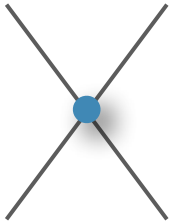


PIONLESS EFT HAMILTONIAN

We take as input a LO pionless-EFT Hamiltonian that we contributed developing

$$H_{LO} = -\sum_i \frac{\vec{\nabla}_i^2}{2m_N} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk}$$

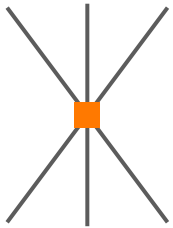
- NN potential: fit to np scattering lengths and effective radii and the deuteron binding energy



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

$$O_{ij}^{p=1,4} = (1, \tau_{ij}, \sigma_{ij}, \sigma_{ij}\tau_{ij})$$

- 3NF adjusted to reproduce the ^{3}H binding energy.



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

*R. Schiavilla, AL, PRC **103**, 054003(2021)*

NEURAL SLATER-JASTROW ANSATZ

The ANN variational state is a product of mean-field state modulated by a flexible correlator factor

$$\Psi_{SJ}(X) = e^{J(X)}\Phi(X)$$

NEURAL SLATER-JASTROW ANSATZ

The ANN variational state is a product of mean-field state modulated by a flexible correlator factor

$$\Psi_{SJ}(X) = e^{J(X)}\Phi(X)$$

VARIATIONAL THEORY OF NUCLEAR MATTER†

JOHN W. CLARK

McDonnell Center for the Space Sciences and Department of Physics, Washington University, St. Louis, Missouri 63130, U.S.A.

This elegant, sophisticated approach to the problem contrasts sharply with the following rather mundane variational treatment. The strong short-range repulsions may be dealt with simply enough with a trial wave function of the form

$$\Psi = \prod_{1 \leq i < j \leq A} f(r_{ij})\Phi_F, \qquad (1.4)$$

upper-bound property (Emery, 1958; Bell and Squires, 1961). The development of this approach was also hindered by a psychological obstacle—the embarrassing conceptual simplicity of the method, its lack of "snob appeal".

NEURAL SLATER-JASTROW ANSATZ

The ANN variational state is a product of mean-field state modulated by a flexible correlator factor

$$\Psi_{SJ}(X) = e^{J(X)}\Phi(X)$$

- The mean-field part is a Slater determinants of single-particle orbitals

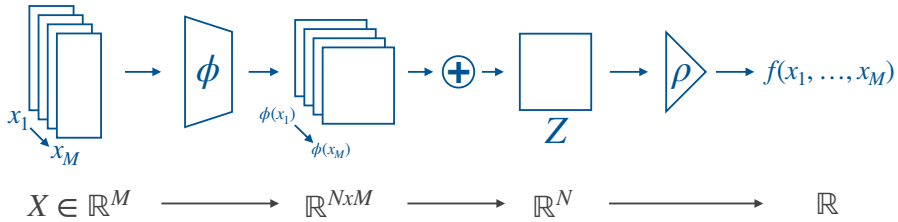
$$\det \begin{bmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_N) \end{bmatrix}$$

- Each orbital is a FFNN that takes as input

$$\bar{\mathbf{r}}_i = \mathbf{r}_i - \mathbf{R}_{CM}$$

- The Jastrow is a permutation-invariant function of the single-particle coordinates

$$J(X) = \rho_F\left[\sum_i \vec{\phi}_{\mathcal{F}}(\bar{\mathbf{r}}_i, \mathbf{s}_i)\right]$$



SAMPLING COORDINATES AND SPIN

The calculation of the observables involve integrating over 3A spatial and 2A spin-isospin variables

$$E_V = \frac{\langle \Psi_V | H | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} = \frac{\sum_S \int dR |\Psi_V(R,S)|^2 \frac{\langle RS|H|\Psi_V\rangle}{\langle RS|\Psi_V\rangle}}{\sum_S \int dR |\Psi_V(R,S)|^2} .$$

We evaluate it stochastically using the Metropolis-Hastings Markov Chain Mote Carlo algorithm

Spatial move → $P_R = \frac{|\Psi_V(R',S)|^2}{|\Psi_V(R,S)|^2}$ Spin-isospin move → $P_S = \frac{|\Psi_V(R,S')|^2}{|\Psi_V(R,S)|^2}$

The observables are estimated by taking averages over the sampled configurations

$$\frac{\langle \Psi_V | O | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} = \frac{1}{N_{\text{conf}}} \sum_{\{R,S\}} O_L(R,S) \longleftrightarrow P_V(R,S) = \frac{\Psi_V(R,S)|^2}{\sum_S \int dR |\Psi_V(R,S)|^2} .$$

STOCHASTIC RECONFIGURATION

The ANN is trained by performing an imaginary-time evolution in the variational manifold

$$(1 - H\delta\tau)|\Psi_V(\mathbf{p}_\tau)\rangle \simeq \Delta p^0 |\Psi_V(\mathbf{p}_\tau)\rangle + \sum_i \Delta p^i O^i |\Psi_V(\mathbf{p}_\tau)\rangle$$

During the optimization, then parameter are updated as

$$\mathbf{p}_{\tau+\delta\tau} = \mathbf{p}_\tau - \eta(S_\tau + \epsilon I)^{-1}\mathbf{g}_\tau$$

The gradient is supplemented by the quantum Fisher Information pre-conditioner

$$S_\tau^{ij} = \left\langle \frac{\partial \Psi_V(\mathbf{p}_\tau)}{\partial p_i} \middle| \frac{\partial \Psi_V(\mathbf{p}_\tau)}{\partial p_j} \right\rangle - \left\langle \frac{\partial \Psi_V(\mathbf{p}_\tau)}{\partial p_i} \right\rangle \left\langle \frac{\partial \Psi_V(\mathbf{p}_\tau)}{\partial p_j} \right\rangle$$

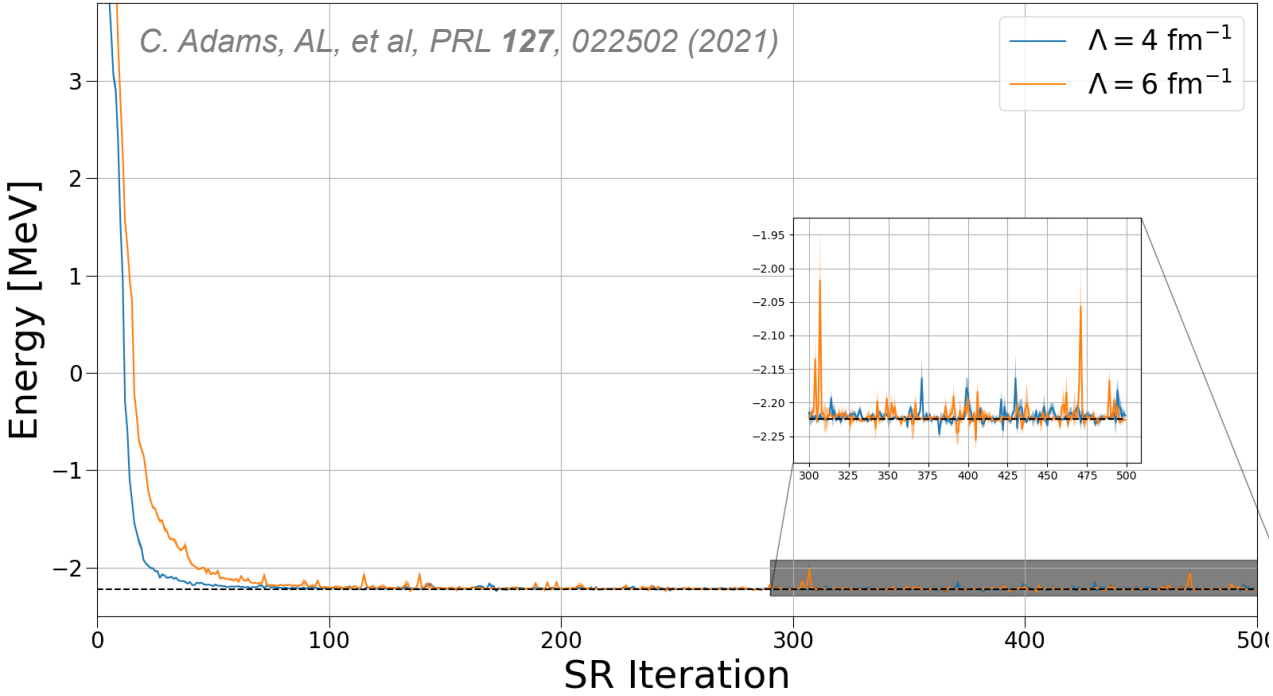
$$\gamma(\psi, \phi) = \arccos \sqrt{\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}}$$

*S. Sorella, Phys. Rev. B **64**, 024512 (2001)*

*J. Stokes, at al., Quantum **4**, 269 (2020).*

ADAPTIVE STOCHASTIC RECONFIGURATION

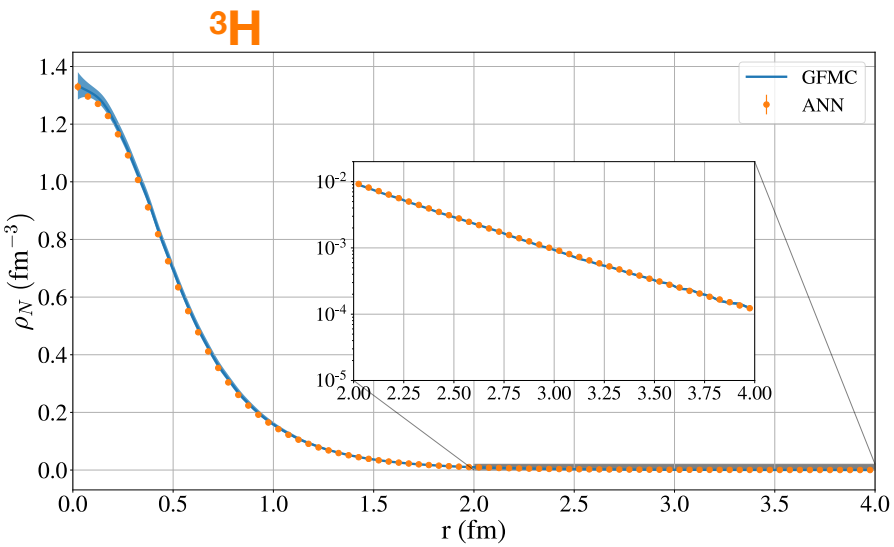
We use an adaptive learning rate with $10^{-7} < \eta < 10^{-2}$. It yields robust convergence patterns for all the nuclei and regulator choices that we have analyzed

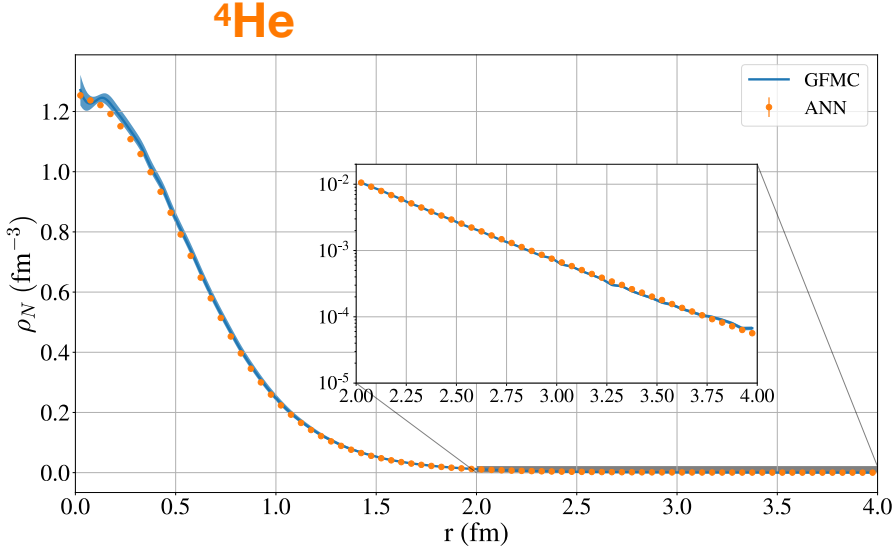


COMPARISON WITH QUANTUM MONTE CARLO

To further elucidate the quality of the ANN wave function we consider the point-nucleon density

$$\rho_N(r) = \frac{1}{4\pi r^2}\left\langle \Psi_V \middle| \sum_i \delta(r - |\mathbf{r}_i^{\text{int}}|) \middle| \Psi_V \right\rangle ,$$





C. Adams, AL, et al, PRL **127**, 022502 (2021)

COMPARISON WITH QUANTUM MONTE CARLO

- The ANN Slater Jastrow ansatz outperforms conventional Jastrow correlations

	Λ	VMC-ANN	VMC-JS	GFMC	GFMC$_c$
^{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}	$-8.26(1)$	$-7.80(1)$	$-8.38(2)$	$-7.82(1)$
	6 fm^{-1}	$-8.27(1)$	$-7.74(1)$	$-8.38(2)$	$-7.81(1)$
^{4}He	4 fm^{-1}	$-23.30(2)$	$-22.54(1)$	$-23.62(3)$	$-22.77(2)$
	6 fm^{-1}	$-24.47(3)$	$-23.44(2)$	$-25.06(3)$	$-24.10(2)$

- Remaining differences with the GFMC are due to deficiencies in the Slater-Jastrow ansatz

$$\Psi_{SJ}(X) = e^{J(X)}\Phi(X)$$

C. Adams, AL, et al, PRL **127**, 022502 (2021)

HIDDEN NUCLEONS

The “hidden fermion” approach was recently introduced to model fermion wave functions

$$\langle RS|\Psi_{HF}\rangle = \begin{pmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) & \phi_1(x_4) & \phi_1(y_1) & \phi_1(y_2) & \phi_1(y_3) & \phi_1(y_4) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(x_3) & \phi_2(x_4) & \phi_2(y_1) & \phi_2(y_2) & \phi_2(y_3) & \phi_2(y_4) \\ \phi_3(x_1) & \phi_3(x_2) & \phi_3(x_3) & \phi_3(x_4) & \phi_3(y_1) & \phi_3(y_2) & \phi_3(y_3) & \phi_3(y_4) \\ \phi_4(x_1) & \phi_4(x_2) & \phi_4(x_3) & \phi_4(x_4) & \phi_4(y_1) & \phi_4(y_2) & \phi_4(y_3) & \phi_4(y_4) \\ \chi_1(x_1) & \chi_1(x_2) & \chi_1(x_3) & \chi_1(x_4) & \chi_1(y_1) & \chi_1(y_2) & \chi_1(y_3) & \chi_1(y_4) \\ \chi_2(x_1) & \chi_2(x_2) & \chi_2(x_3) & \chi_2(x_4) & \chi_2(y_1) & \chi_2(y_2) & \chi_2(y_3) & \chi_2(y_4) \\ \chi_3(x_1) & \chi_3(x_2) & \chi_3(x_3) & \chi_3(x_4) & \chi_3(y_1) & \chi_3(y_2) & \chi_3(y_3) & \chi_3(y_4) \\ \chi_4(x_1) & \chi_4(x_2) & \chi_4(x_3) & \chi_4(x_4) & \chi_4(y_1) & \chi_4(y_2) & \chi_4(y_3) & \chi_4(y_4) \end{pmatrix}$$

Visible orbitals on visible coordinates

Visible orbitals on hidden coordinates

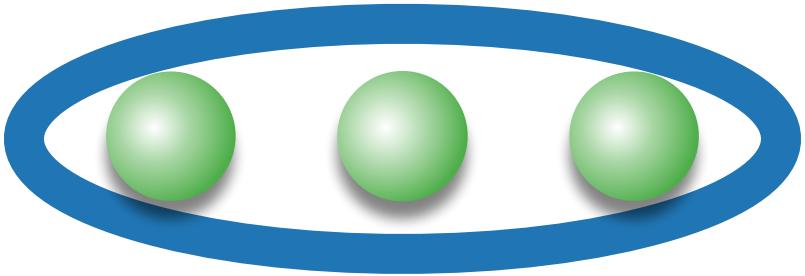
Hidden orbitals on visible coordinates

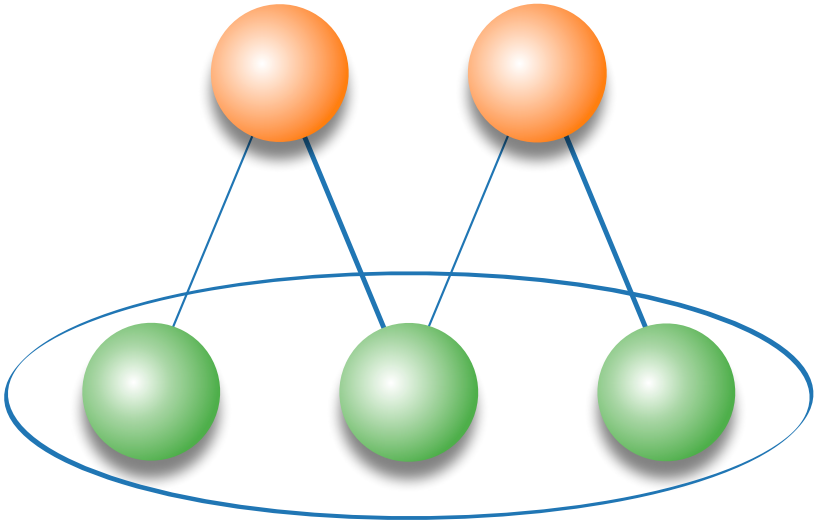
Hidden orbitals on hidden coordinates

*J. R. Moreno, et al., PNAS **119** (32)*

HIDDEN NUCLEONS

The “hidden fermion” approach was recently introduced to model fermion wave functions

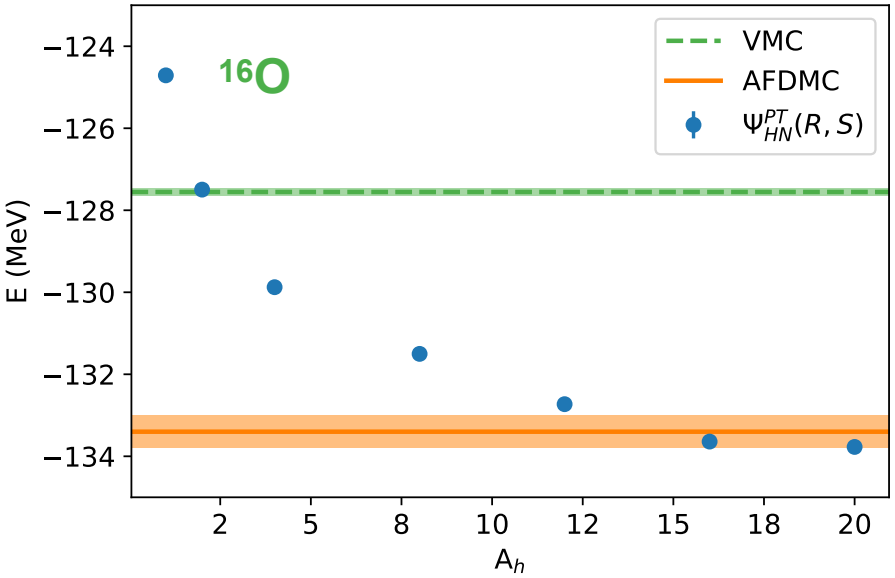


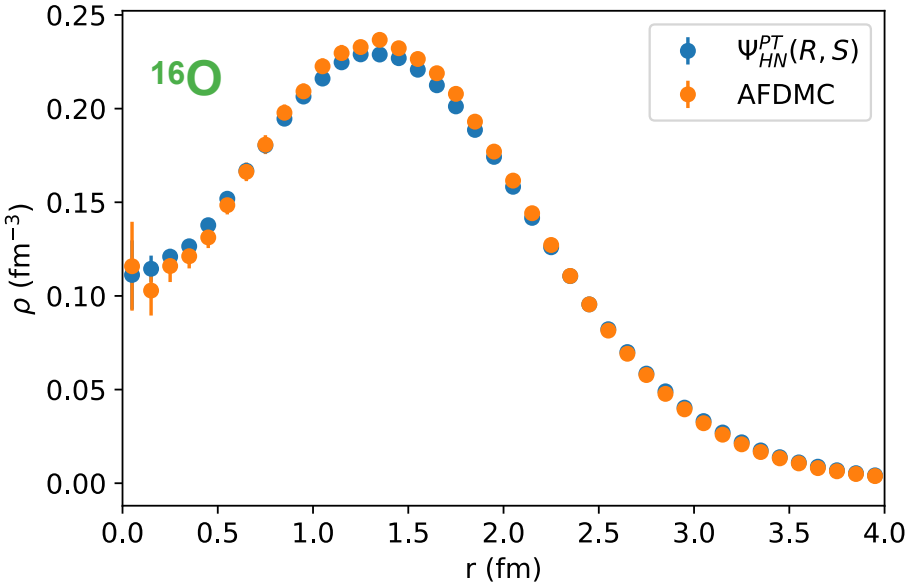


NUCLEAR PHYSICS APPLICATIONS

We extend the reach of neural quantum states to ^{16}O

In addition to its ground-state energy, we evaluate the point-nucleon density of ^{16}O with A_h=16





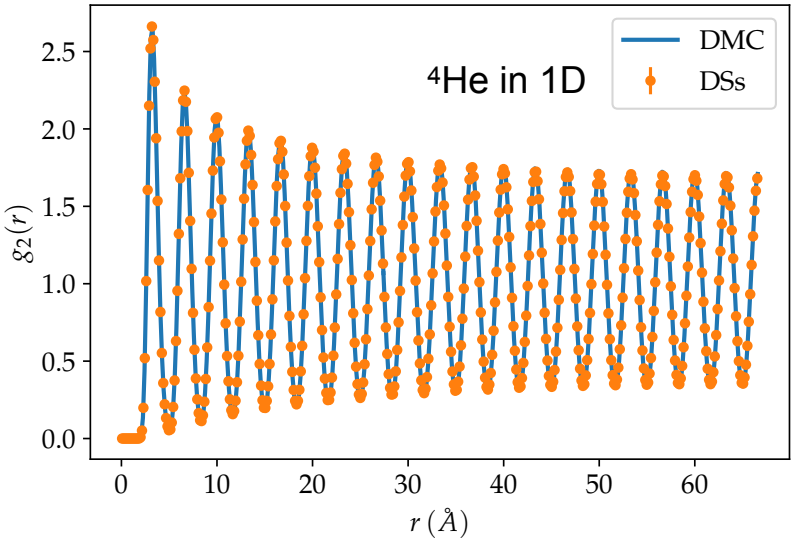
*AL, et al., Phys.Rev.Res. **4** (2022) 4, 043178*

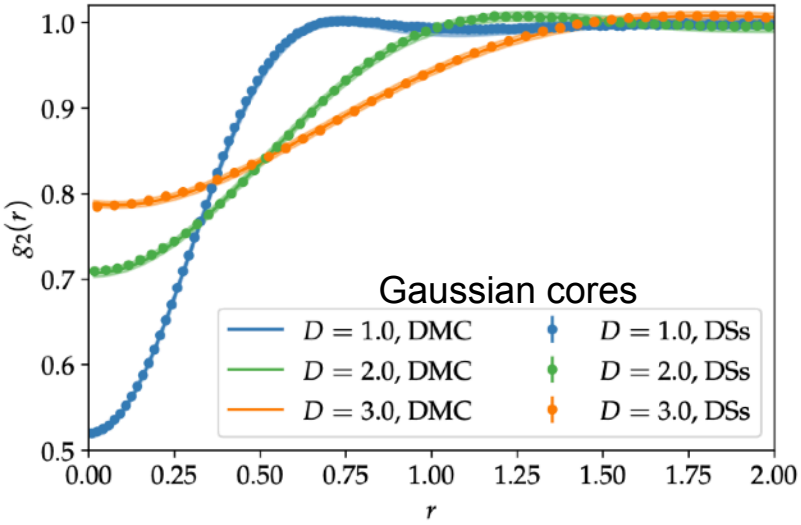
INFINITE PERIODIC SYSTEMS

- We extended our approach to periodic systems, such us liquid ^{4}He and soft (gaussian) spheres

  ➡ Periodic ANN by construction: $\mathbf{r}_i \longrightarrow \tilde{\mathbf{r}}_i = \left\{ \sin\left(\frac{2\pi}{L}\mathbf{r}_i\right), \cos\left(\frac{2\pi}{L}\mathbf{r}_i\right) \right\}$

  ➡ Permutation invariant Deep-Sets ANN for computing bosons: $\Psi_V(R) = e^{\mathcal{U}(R)}$





*G. Pescia, AL, et al., Phys. Rev .Res. **4** (2022) 023138*

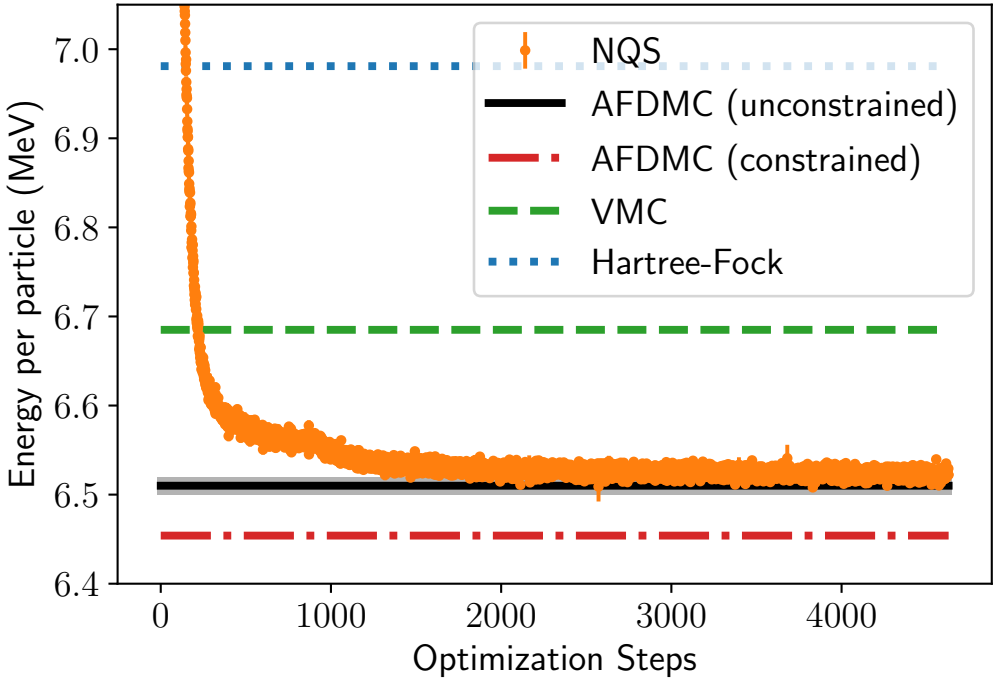
DILUTE NEUTRON MATTER

We have introduced a periodic hidden-nucleons ansatz to model low-density neutron matter

The NQS ansatz converges to the unconstrained AFDMC energy, using a fraction of the computing time

- NQS: 100 hours on NVIDIA-A100
- AFDMC: 1.2 million hours on Intel-KNL

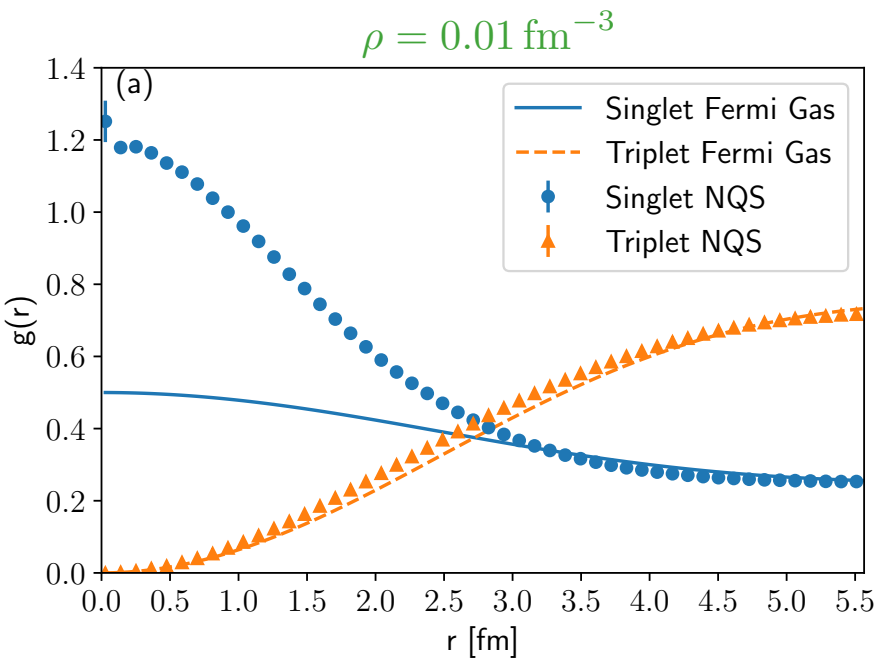
The hidden-nucleon ansatz captures the overwhelming majority of the correlation energy

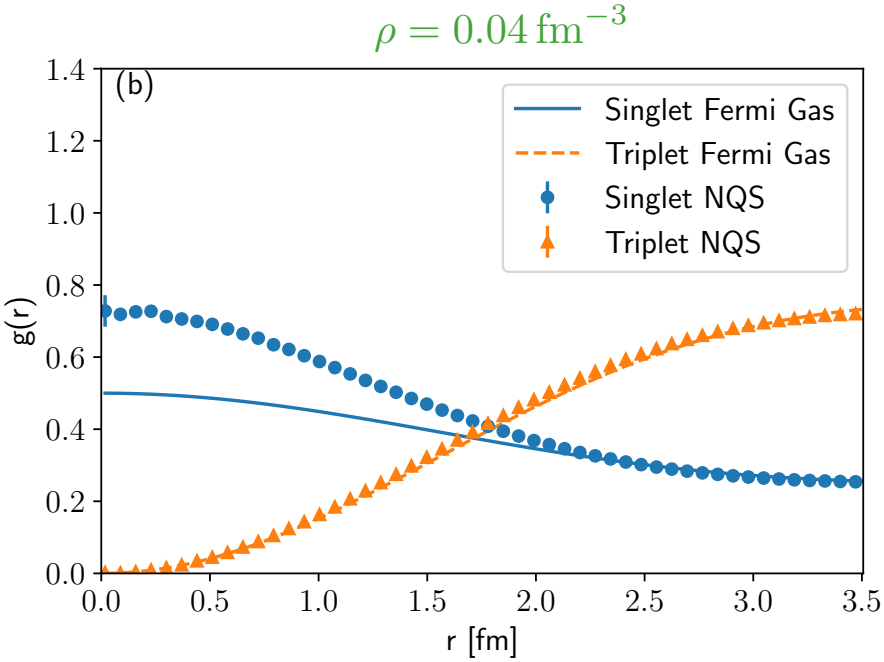


B. Fore, J. Kim, AL, arXiv:2212.04436 [nucl-th]

DILUTE NEUTRON MATTER

Low-density neutron matter is characterized by fascinating emergent quantum phenomena, such as the formation of Cooper pairs and the onset of superfluidity.



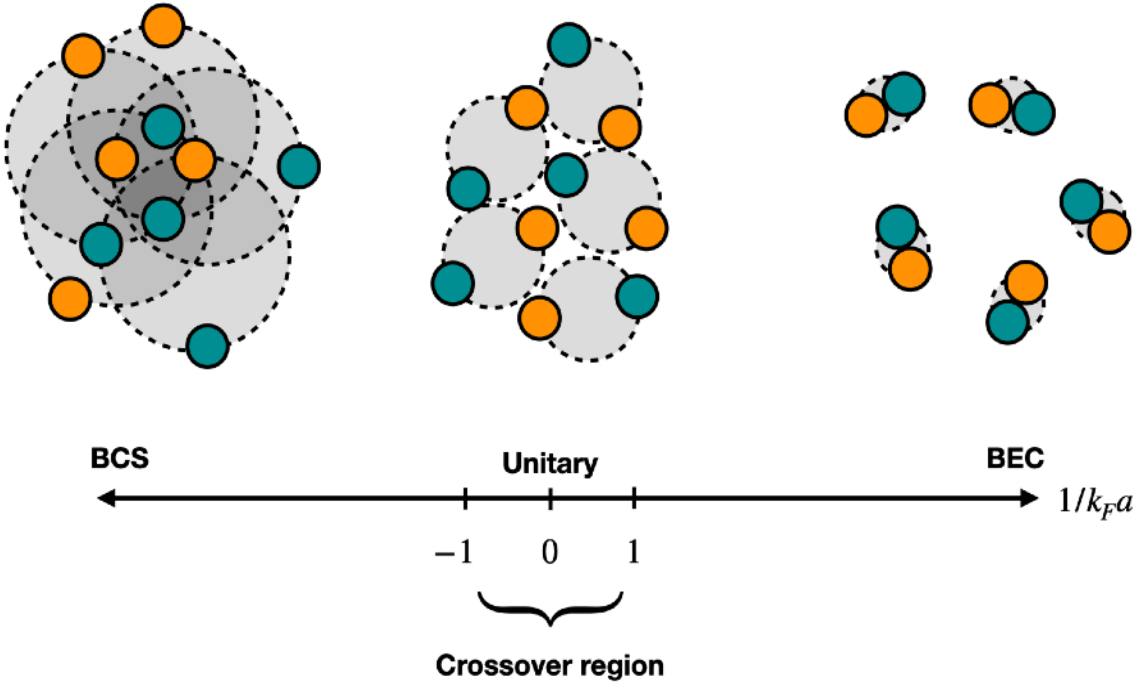


B. Fore, J. Kim, AL, arXiv:2212.04436 [nucl-th]

J. Kim, B. Fore, AL, arXiv:2305.08831

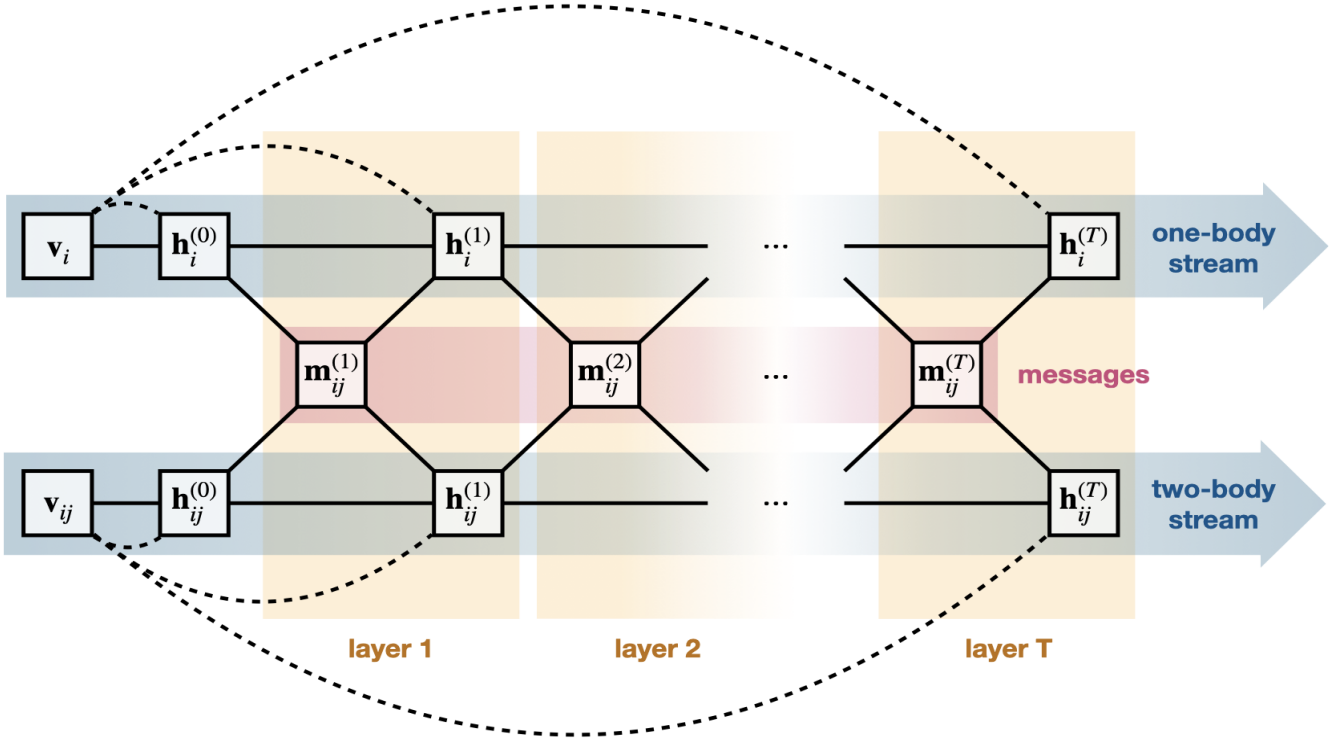
COLD FERMI GASES

Periodic-NQS can accurately model the BCS-BEC crossover in two-components cold Fermi gases



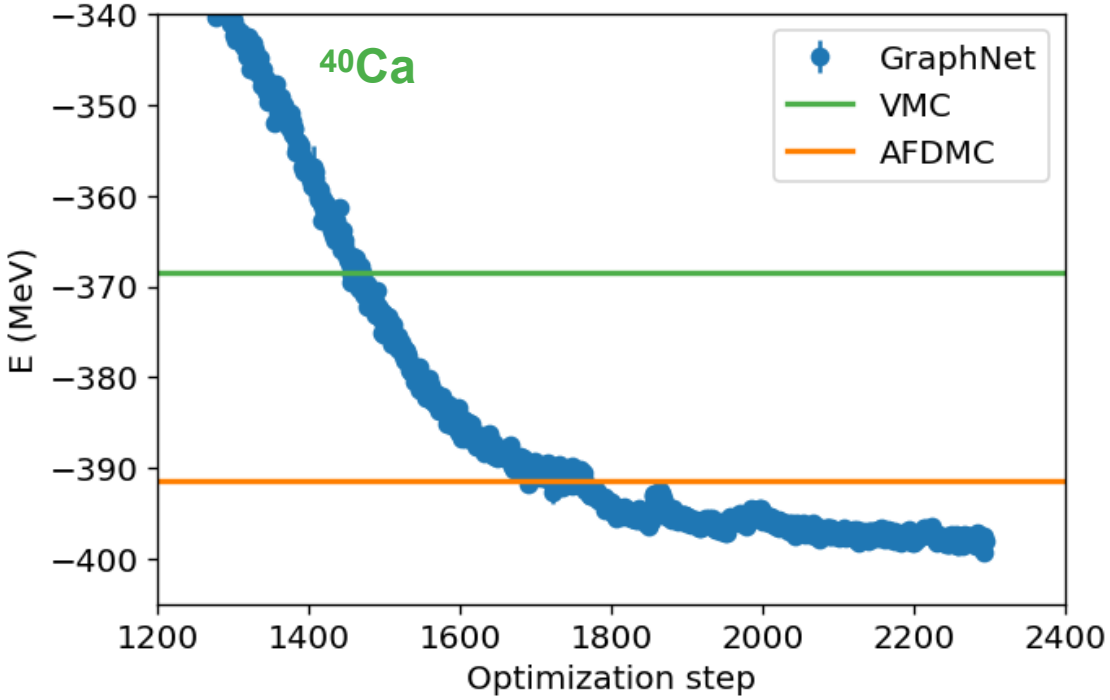
COLD FERMI GASES

Develop a permutation-invariant message-passing neural network to iteratively build correlations



BACK TO NUCLEI, WITH MPNN

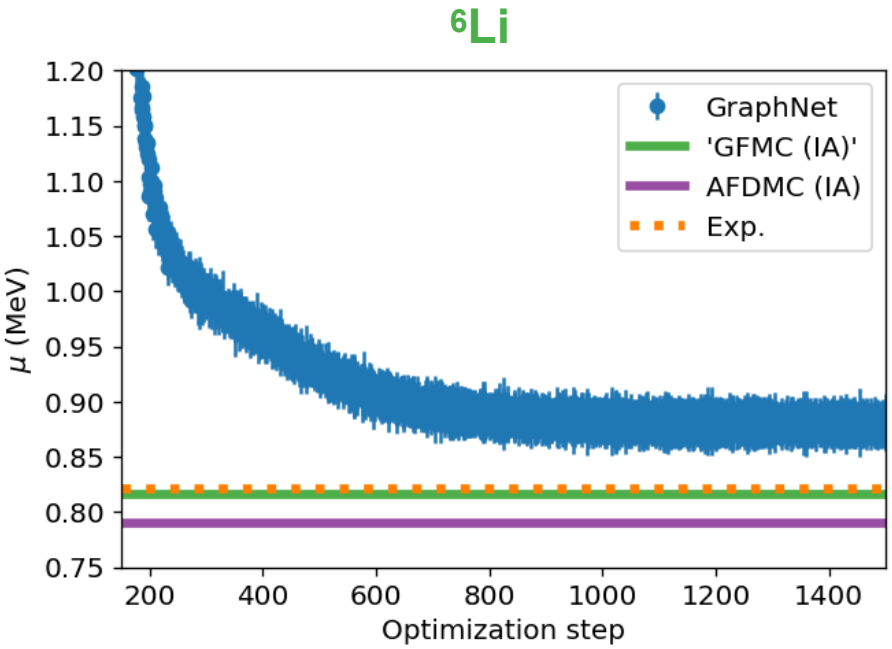
Even with just one hidden-nucleon we do better than AFDMC for medium-mass nuclei

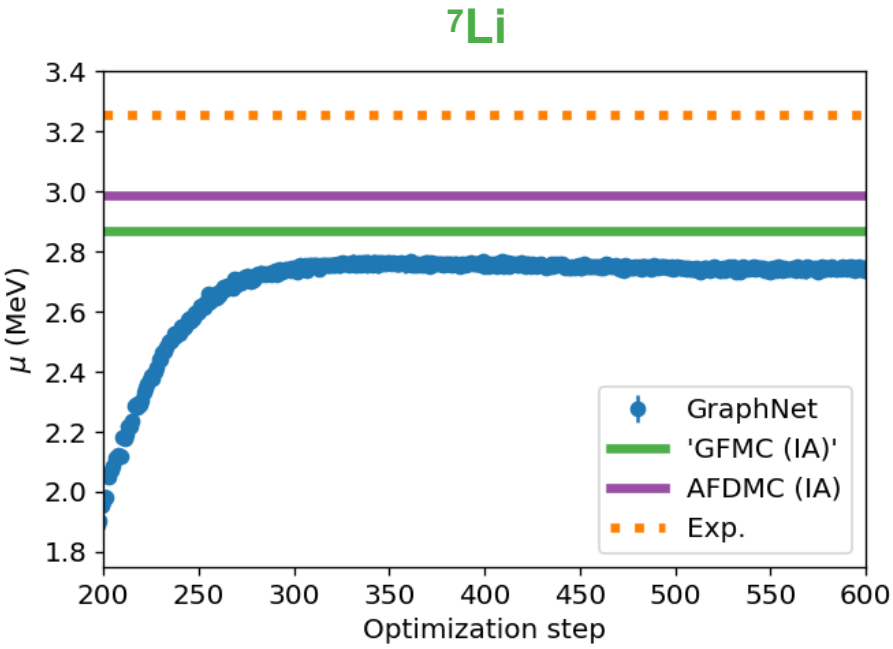


AL, B. Fore, et al., in preparation

BACK TO NUCLEI, WITH MPNN

In addition to energies and single-particle densities, we can compute electroweak properties





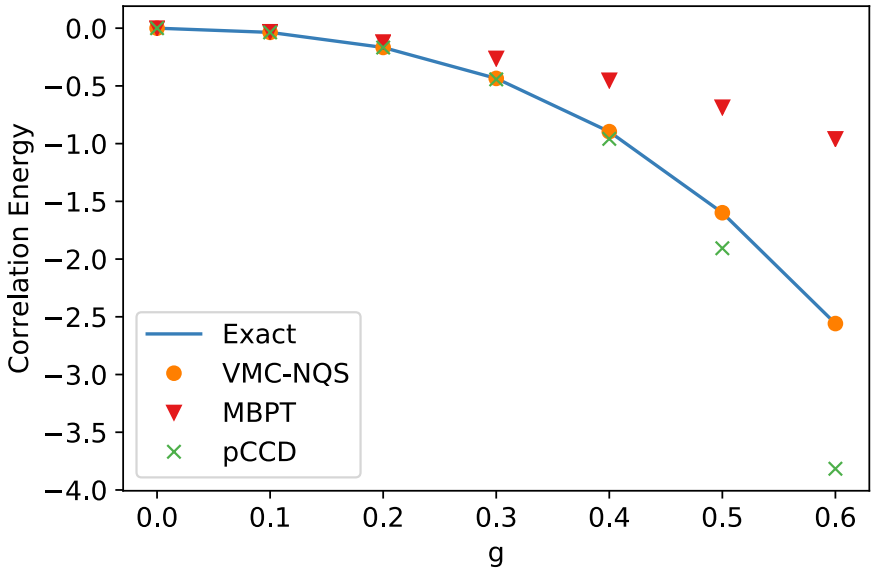
A. Gnech, AL, in preparation

NUCLEAR PAIRING MODEL

We are exploring the solution of the many-body problem in the occupation-number formalism. As a prototypal example, we consider a simplified pairing model

$$H = \sum_{p,\sigma} d_p a^\dagger_{p\sigma} a_{p\sigma} + \sum_{pq} g_{pq} a^\dagger_{p\uparrow} a^\dagger_{p\downarrow} a_{q\downarrow} a_{q\uparrow} \quad \Longleftrightarrow \quad d_p = p \;\; ; \;\; g_{pq} = g$$

- The Jordan-Wigner transformations map the Hamiltonian into the spin basis;
- We solve it using ANN ansatz and compare with exact diagonalization;
- The algorithm scales as the number of single-particle states squared;



*M. Rigo, B. Hall, AL, Phys. Rev.E **107** (2023) 2, 025310*

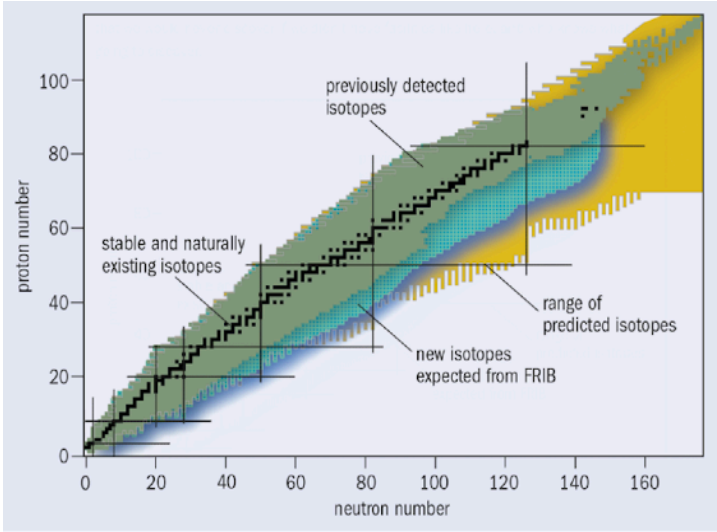
CONCLUSIONS

Neural network quantum states are extending the reach of conventional QMC methods

- Favorable scaling with the number of fermions;
- Universal and accurate approximations for fermion wave functions;
- Suitable for confined and periodic systems;
- Scalable to leadership-class hybrid CPU/GPU computers

PERSPECTIVES

- NQS calculations of medium-mass stable and exotic nuclei relevant to FRIB and ATLAS



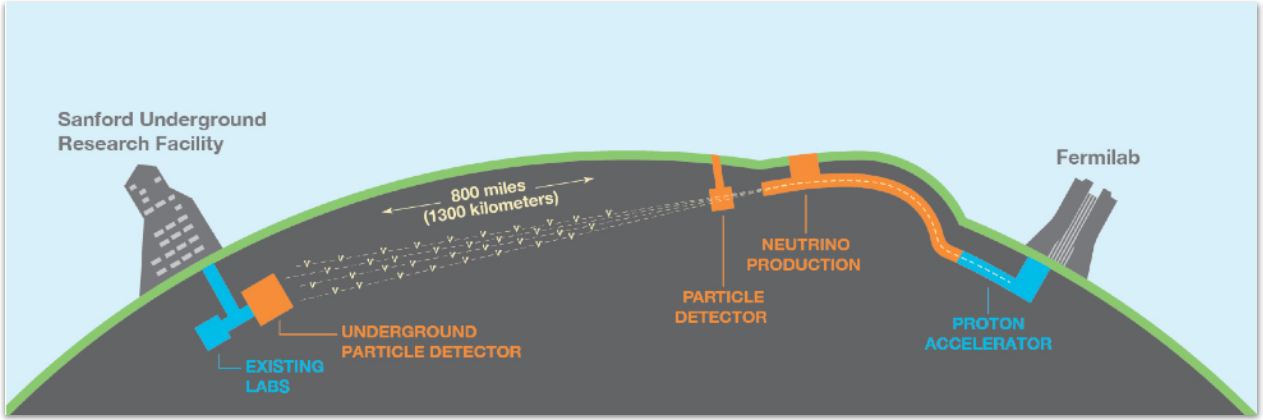
- High-precision electroweak transitions, including magnetic moments and beta-decay rates;
- Compute low-density isospin-asymmetric nucleonic matter: the flexibility of NQS will allow us to see self-emerging clustering in the low-density region;

PERSPECTIVES

- Access “real-time” dynamics: the prototypal exponentially-hard problem in many-body theory

$$\mathcal{D}\left(|\Psi(\mathbf{p}_{t+\delta t})\rangle, e^{-iHt}|\Psi(\mathbf{p}_t)\rangle\right)^2 = \arccos\left(\sqrt{\frac{\langle\Psi(\mathbf{p}_{t+\delta t})|e^{-iHt}|\Psi(\mathbf{p}_t)\rangle\langle\Psi(\mathbf{p}_t)|e^{iHt}|\Psi(\mathbf{p}_{t+\delta t})\rangle}{\langle\Psi(\mathbf{p}_{t+\delta t})|\Psi(\mathbf{p}_{t+\delta t})\rangle\langle\Psi(\mathbf{p}_t)|\Psi(\mathbf{p}_{\tau+\delta t})\rangle}}\right)^2$$

- Relevant for: fusion, lepton-nucleus scattering, and collective neutrino oscillation;



CONGRATULATIONS, JOHN!