# Variational Monte Carlo Calculations of $A \leq 4$ Nuclei with an Artificial Neural-Network Correlator Ansatz

Corey Adams,[1,2] Giuseppe Carleo,[3] Alessandro Lovato,[1,4] and Noemi Rocco[5]
[1]Physics Division, Argonne National Laboratory, Argonne, Illinois 60439
[2]Leadership Computing Facility, Argonne National Laboratory, Argonne, Illinois 60439
[3]Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[4]INFN-TIFPA Trento Institute of Fundamental Physics and Applications, 38123 Trento, Italy
[5]Theoretical Physics Department, Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, USA

(Received 3 August 2020; revised 13 April 2021; accepted 25 May 2021; published 7 July 2021; corrected 12 July 2021)

The complexity of many-body quantum wave functions is a central aspect of several fields of physics and chemistry where nonperturbative interactions are prominent. Artificial neural networks (ANNs) have proven to be a flexible tool to approximate quantum many-body states in condensed matter and chemistry problems. In this work we introduce a neural-network quantum state ansatz to model the ground-state wave function of light nuclei, and approximately solve the nuclear many-body Schrödinger equation. Using efficient stochastic sampling and optimization schemes, our approach extends pioneering applications of ANNs in the field, which present exponentially scaling algorithmic complexity. We compute the binding energies and point-nucleon densities of $A \leq 4$ nuclei as emerging from a leading-order pionless effective field theory Hamiltonian. We successfully benchmark the ANN wave function against more conventional parametrizations based on two- and three-body Jastrow functions, and virtually exact Green's function Monte Carlo results.

DOI: 10.1103/PhysRevLett.127.022502

*Introduction.*—The last two decades have witnessed remarkable progress in our understanding of how the structure and dynamics of atomic nuclei emerge from the individual interaction among protons and neutrons. This progress has been primarily driven by the widespread use of nuclear-effective field theories to systematically construct realistic Hamiltonians [1–3], and the concurrent development of nuclear many-body techniques that solve the time-independent Schrödinger equation with controlled approximations [4–8]. The variational Monte Carlo (VMC) and Green's function Monte Carlo (GFMC) methods are ideally suited to tackle this problem and have been extensively applied to study properties of light nuclei [9]. Monte Carlo techniques also face important challenges. For example, the calculation of the spin-isospin dependent Jastrow correlations used in the VMC and GFMC scales exponentially with the number of nucleons, limiting the applicability of these methods to relatively small nuclear systems. Also, the auxiliary-field diffusion Monte Carlo [10] (AFDMC) samples the spin and isospin degrees of freedom to treat larger nuclei and infinite nucleonic matter [11,12], but it can only take as inputs somewhat simplified interactions [13]. In addition, the use of wave functions that scale polynomially with the number of nucleons exacerbates the AFDMC fermion sign problem for $A > 16$ nuclei. Therefore, extending VMC and GFMC calculations to medium-mass nuclei requires devising accurate wave functions that exhibit a polynomial scaling with A.

An alternative class of approaches being actively explored is based on machine learning (ML) techniques. These techniques typically rely on the ability of artificial neural networks (ANNs) to compactly represent complex high-dimensional functions, as already leveraged in several domains of physics [14]. For many-body quantum applications, neural-network-based variational representations have been introduced in Ref. [15], and have found applications as a tool to study ground state and dynamics of several interacting lattice quantum systems [16–24]. In a series of recent works [25–27] deep neural networks have been further developed to tackle *ab initio* chemistry problems within variational Monte Carlo, often resulting in accuracy improvements over existing variational approaches traditionally used to describe correlated molecules. While applications of ML approaches to the many-body problem in condensed matter, quantum chemistry, and quantum information have been proliferating in the past few years, the adoption in low-energy nuclear theory is still in its infancy [28,29]. Pioneering work in the field [30] has provided a proof-of-principle application of ANN to solve the Schrödinger equation of the deuteron. Extending the nonstochastic approach of Ref. [30] to larger nuclei, however, presents an intrinsically exponentially scaling challenge.

In this work, we expand the domain of applicability of ANN-based representations of the wave function and compute ground-state properties of $A \leq 4$ nuclei as they

 

emerge from a leading-order pionless effective field theory (EFT) Hamiltonian, containing consistent two- and three-body potentials. Specifically, we develop a novel VMC algorithm based on an ANN representation of the spin-isospin dependent correlator that captures the vast majority of nuclear correlations and scales favorably with the number of nucleons. We benchmark our results against a more conventional parametrization of the variational wave function in terms of two- and three-body Jastrow functions, and virtually exact GFMC calculations.

*Hamiltonian.*—We employ nuclear Hamiltonians derived within pionless EFT, which is based on the tenet that the typical momentum of nucleons in nuclei is much smaller than the pion mass $m_\pi$ [3,31]. Under this assumption, largely justified for studying the structure and long-range properties of $A \leq 4$ nuclei, pion exchanges are unresolved contact interactions and nucleons are the only relevant degrees of freedom. The singularities of the contact terms are controlled introducing a Gaussian regulator that suppresses transferred momenta above the ultraviolet cutoff $\Lambda$. This regulator choice directly leads to a Gaussian radial dependence of the potential, which is local in coordinate [32,33]. The leading-order (LO) Hamiltonian reads

$$H_{\rm LO} = -\sum_i \frac{\vec{\nabla}_i^2}{2m_N} + \sum_{i<j}(C_1 + C_2 \vec{\sigma}_i \cdot \vec{\sigma}_j) e^{-r_{ij}^2 \Lambda^2/4} + D_0 \sum_{i<j<k} \sum_{\rm cyc} e^{-(r_{ik}^2 + r_{ij}^2)\Lambda^2/4}, \quad (1)$$

where $m_N$ is the mass of the nucleon, $\vec{\sigma}_i$ is the Pauli matrix acting on nucleon $i$, and $\sum_{\rm cyc}$ stands for the cyclic permutation of $i$, $j$, and $k$.

Following Ref. [34], the low-energy constants $C_1$ and $C_2$ are fit to the deuteron binding energy and to the neutron-neutron scattering length. In Eq. (1) we picked the operator basis 1 and $\vec{\sigma}_i \cdot \vec{\sigma}_j$, but this choice can be replaced by any other form equivalent under Fierz transformations in SU(2). Solving $A \geq 3$ nuclei with purely attractive two-nucleon potentials leads to the "Thomas collapse" [35], which can be avoided promoting a contact three-nucleon force to LO [36]. The values of the low-energy constants' adopted in this work can be found in Ref. [34]; since $C_1(\Lambda)$ is much larger than $C_2(\Lambda)$, the LO Hamiltonian has an approximate SU(4) symmetry.

*Variational wave function.*—A fundamental ingredient of the VMC method is the choice of a suitable variational wave function $\Psi_V$, whose parameters are found exploiting the variational principle

$$\frac{\langle \Psi_V | H | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} = E_V \geq E_0 \quad (2)$$

where $E_0$ is exact the ground-state energy: $H|\Psi_0\rangle = E_0|\Psi_0\rangle$. The metropolis Monte Carlo algorithm is used to evaluate the variational energy $E_V$ by sampling the spatial and spin-isospin coordinates. We introduce the following ANN representation of the variational wave function:

$$|\Psi_V^{\rm ANN}\rangle = e^{\mathcal{U}(\mathbf{r}_1,\ldots,\mathbf{r}_A)} \tanh[\mathcal{V}(\mathbf{s}_1, \mathbf{r}_1, \ldots, \mathbf{r}_A, \mathbf{s}_A)]|\Phi\rangle \quad (3)$$

where $\{\mathbf{r}_1, \ldots, \mathbf{r}_A\}$ and $\{\mathbf{s}_1, \ldots, \mathbf{s}_A\}$ denote the set of single-particle spatial three-dimensional coordinates and the z projection of the spin-isospin degrees of freedom $\mathbf{s}_i = \{s_i^z, t_i^z\}$, respectively. For the $s$-shell nuclei considered in this work, we take $|\Phi_{^2{\rm H}}\rangle = \mathcal{A}|\uparrow_p \uparrow_n\rangle$, $|\Phi_{^3{\rm He}}\rangle = \mathcal{A}|\uparrow_p \downarrow_p \uparrow_n\rangle$, and $|\Phi_{^4{\rm He}}\rangle = \mathcal{A}|\uparrow_p \downarrow_p \uparrow_n \downarrow_n\rangle$, with $\mathcal{A}$ being the antisymmetrization operator [37].

The real-valued correlating factors $\mathcal{U}(\mathbf{r}_1, \ldots, \mathbf{r}_A)$ and $\mathcal{V}(\mathbf{s}_1, \mathbf{r}_1, \ldots, \mathbf{r}_A, \mathbf{s}_A)$ are parametrized in terms of permutation-invariant ANNs, so that the total wave function is antisymmetric. To achieve this goal, we make use of the Deep Sets architecture [38,39], and map each of the single-particle inputs separately to a latent-space representation. We then apply a sum operation to destroy the ordering of the information and ensure permutation invariance

$$\mathcal{F}(\mathbf{x}_1, \ldots, \mathbf{x}_A) = \rho_{\mathcal{F}}\left(\sum_{\mathbf{x}_i} \phi_{\mathcal{F}}(\mathbf{x}_i)\right), \qquad \mathcal{F} = \mathcal{U}, \mathcal{V}. \quad (4)$$

Both $\phi_{\mathcal{U}}$ and $\rho_{\mathcal{U}}$ are represented by ANNs comprised of four fully connected layers with 32 nodes each, while $\phi_{\mathcal{V}}$ and $\rho_{\mathcal{V}}$ are made of two fully connected layers, again with 32 nodes, for total of 13 058 trainable parameters. The calculation of the kinetic energy requires using differentiable activation functions. We find that tanh and softplus [40] yield fully consistent results. The single-particle inputs are $\mathbf{x}_i \equiv \{\bar{\mathbf{r}}_i\}$ and $\mathbf{x}_i \equiv \{\bar{\mathbf{r}}_i, \mathbf{s}_i\}$ for $\mathcal{U}$ and $\mathcal{V}$, respectively, where we defined intrinsic spatial coordinates as $\bar{\mathbf{r}}_i = \mathbf{r}_i - \mathbf{R}_{\rm CM}$, with $\mathbf{R}_{\rm CM}$ being the center-of-mass coordinate. This procedure automatically removes spurious center-of-mass contributions from all observables [41]. Since the parameters of the network are randomly initialized, in the initial phases of the training, during the metropolis walk, the nucleons can drift away from $\mathbf{R}_{\rm CM}$. To control this behavior, a Gaussian function is added to confine the nucleons within a finite volume $\mathcal{U}(\mathbf{r}_1, \ldots, \mathbf{r}_A) \to \mathcal{U}(\mathbf{r}_1, \ldots, \mathbf{r}_A) - \alpha \sum_i \bar{\mathbf{r}}_i^2$ where we take $\alpha = 0.05$.

The choice of correcting a mean-field state $|\Phi\rangle$ with a flexible ANN correlator factor is similar in spirit to neural-network correlators introduced recently in condensed-matter [16,24] and chemistry applications [26], but it is more general as it encompasses spin-isospin dependent correlations. An appealing feature of the ANN ansatz is that it is more general than the more conventional product of two- and three-body spin-independent Jastrow functions

$$|\Psi_V^J\rangle = \prod_{i<j<k}\left(1-\sum_{\text{cyc}}u(r_{ij})u(r_{jk})\right)\prod_{i<j}f(r_{ij})|\Phi\rangle, \quad (5)$$

which is commonly used for nuclear Hamiltonians that do not contain tensor and spin-orbit terms [33,42].

Analogously to standard VMC calculations, as well as ML applications, the optimal set of weights and biases of the ANN is found minimizing a suitable cost function. Specifically, we exploit the variational principle of Eq. (2) and minimize the expectation value of the energy. The gradient components $G_i = \partial_i E(p)$ of the energy with respect to the variational parameters $p_i$ read

$$G_i = 2\left(\frac{\langle\partial_i\Psi_V|H|\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle} - E_V\frac{\langle\partial_i\Psi_V|\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle}\right) \quad (6)$$

and can be efficiently estimated through Monte Carlo sampling. While stochastic gradient descent can be readily used to compute parameters updates, for VMC applications it has been found that using a preconditioner based on the quantum Fisher information

$$S_{ij} = \frac{\langle\partial_i\Psi_V|\partial_j\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle} - \frac{\langle\partial_i\Psi_V|\Psi_V\rangle\langle\Psi_V|\partial_j\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle\langle\Psi_V|\Psi_V\rangle} \quad (7)$$

is significantly more efficient. During the optimization, then parameters at step $s$ are updated as $p^{s+1} = p^s - \eta(S+\Lambda)^{-1}G$, where $\eta$ is the learning rate and $\Lambda$ is a small positive diagonal matrix that is added to stabilize the method. This approach, known as the stochastic-reconfiguration (SR) algorithm [43,44] is equivalent to performing imaginary-time evolution in the variational manifold and it is in turn related to the natural gradient descent method [45] in unsupervised learning. Our computational techniques are based on the general ML framework TensorFlow [46], and it is scalable across more than 100 GPUs. We also maintain an additional developmental repository written in JAX [47] for fast prototyping of new features. More information about the architecture and performance of the software is available in the Supplemental Material [48].

Figure 1 displays the convergence pattern as a function of the optimization step of the $^2$H energy for the LO pionless EFT Hamiltonians with $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$. In the initial phase of the optimization, the softer cutoff exhibits a faster convergence than the stiffer one. However, the asymptotic value of the energy is reached after about 300 iterations for both values of the regulator. These results have been obtained using an adaptive learning rate in the range $10^{-7} \leq \eta \leq 10^{-2}$, which has proven to yield robust convergence patterns for all the nuclei and regulator choices that we have analyzed. The adaptive schedule of this *AdaptiveEta* algorithm is selected performing heuristic tests on the parameter change, similar to the ones introduced in Refs. [33,41] for regularizing the linear optimization method [58].

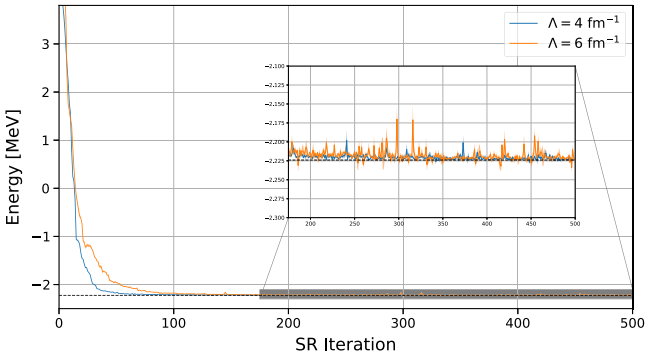


FIG. 1. Convergence pattern of the $^2$H variational energy for $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$ as a function of the number of optimization steps of the SR *AdaptiveEta* algorithm. The dashed line denotes the asymptotic value.

*Results and discussion.*—We analyze the accuracy of the ANN wave function ansatz by computing the ground-state energies of $^2$H, $^3$H, and $^4$He. In Table I we benchmark the ANN representation of $\Psi_T$ (VMC-ANN) against conventional VMC calculations carried out using a spline parametrization for the Jastrow functions [33] (VMC-JS), and virtually exact GFMC results.

The three methods provide fully compatible energies for $^2$H nucleus, within statistical errors, showing the flexibility of the ANN to accurately represent the ground-state wave function of the deuteron, consistent with the findings of Ref. [30]. Note that, since the LO pionless EFT Hamiltonian does not contain tensor or spin-orbit terms, the VMC-JS ansatz is exact. The perfect agreement with the experimental value is not surprising, as the potential has been fit to the deuteron binding energy using numerically exact few-body methods [32].

The VMC-ANN noticeably improves upon the VMC-JS energies of $^3$H, by $\simeq$0.5 MeV for both $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$. On the other hand, the GFMC results are $\simeq$0.1 MeV more bound than the VMC-ANN ones. This difference is due to spin-dependent correlations that are automatically generated by the GFMC imaginary-time propagation, but are not fully accounted for by the correlator ansatz of Eq. (3). To better quantify the spin-independent correlations entailed in the ANN, we have considered a simplified "ANN$_c$" ansatz $|\Psi_V^{\text{ANN}_c}\rangle = e^{\mathcal{U}(\mathbf{r}_1,\ldots,\mathbf{r}_A)}|\Phi\rangle$. In this case, the NN potential of Eq. (1) is equivalent to the $SU(4)$-symmetric interaction $\tilde{v}_c(r_{ij}) = v_c(r_{ij}) - v_\sigma(r_{ij})$. For $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$ ANN$_c$ yields $-7.85(2)$ and $-7.85(4)$ MeV, respectively. These numbers are in excellent agreement with the GFMC$_c$ calculations reported in Table I, which have also been carried out using $\tilde{v}_c(r_{ij})$.

A similar pattern emerges for $^4$He, with ANN wave functions outperforming the JS ones: the energy is improved by about 0.8 and 1.0 MeV for $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$, respectively. The small discrepancies

TABLE I. Ground-state energies in MeV of the $^2$H, $^3$H, and $^4$He for the LO pionless-EFT Hamiltonian for $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$. Numbers in parentheses indicate the statistical errors on the last digit.

| | $\Lambda$ | 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) |

with the GFMC are again due to missing spin-isospin dependent correlations in the ANN. In fact, the ANN$_c$ energies turn out to be −22.76(2) MeV and −24.05(5) for $\Lambda = 4$ fm$^{-1}$ and $\Lambda = 6$ fm$^{-1}$, which are fully compatible with the GFMC$_c$ results listed in Table I.

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

$$\rho_N(r) = \frac{1}{4\pi r^2}\langle \Psi_V | \sum_i \delta(r - |\mathbf{r}_i^{\rm int}|) | \Psi_V \rangle, \qquad (8)$$

which is of interest in a variety of experimental settings [59,60]. In the upper, medium, and lower panels of Fig. 2 we display $\rho_N(r)$ of $^2$H, $^3$H, and $^4$He as obtained from VMC-ANN and GFMC calculations that use as input the LO pionless-EFT Hamiltonian with $\Lambda = 4$ fm$^{-1}$. There is an excellent agreement between the two methods, which further corroborates the representative power of the ANN ansatz for the wave functions of $A \leq 4$ nuclei. The VMC-ANN and GFMC densities overlap both at short distances and in the slowly decaying asymptotic exponential tails, highlighted in the insets of Fig. 2. It should be emphasized that the ANN learns how to compensate for the original Gaussian confining function and reproduce the correct exponential falls off of the nuclear wave function, which is notoriously delicate to obtain within nuclear methods that rely on harmonic-oscillator basis expansions [61,62].

*Conclusions.*—In this work we have carried out proof-of-principle calculations that demonstrate the capability of ANNs to represent the variational state of $A \leq 4$ nuclei encompassing the vast majority of nuclear correlations and scale favorably with the number of nucleons. Exploiting the Deep Sets architecture, we have devised permutation-invariant, spin-isospin dependent correlators whose computational cost scales polynomially with the number of nucleons. Using the stochastic-reconfiguration algorithm, we solve the Schrödinger equation of a LO pionless-EFT Hamiltonian that contains two- and three-nucleon potentials characterized by highly nonperturbative, spin-dependent, short-range components. The spin-isospin dependent ANN variational wave function outperforms the routinely employed two- and three-body Jastrow parametrization of the correlation function. The small remaining differences with the exact GFMC result will likely be solved once spin-dependent backflow correlations are introduced in the Slater determinant, as in Ref. [25–27], paving the way

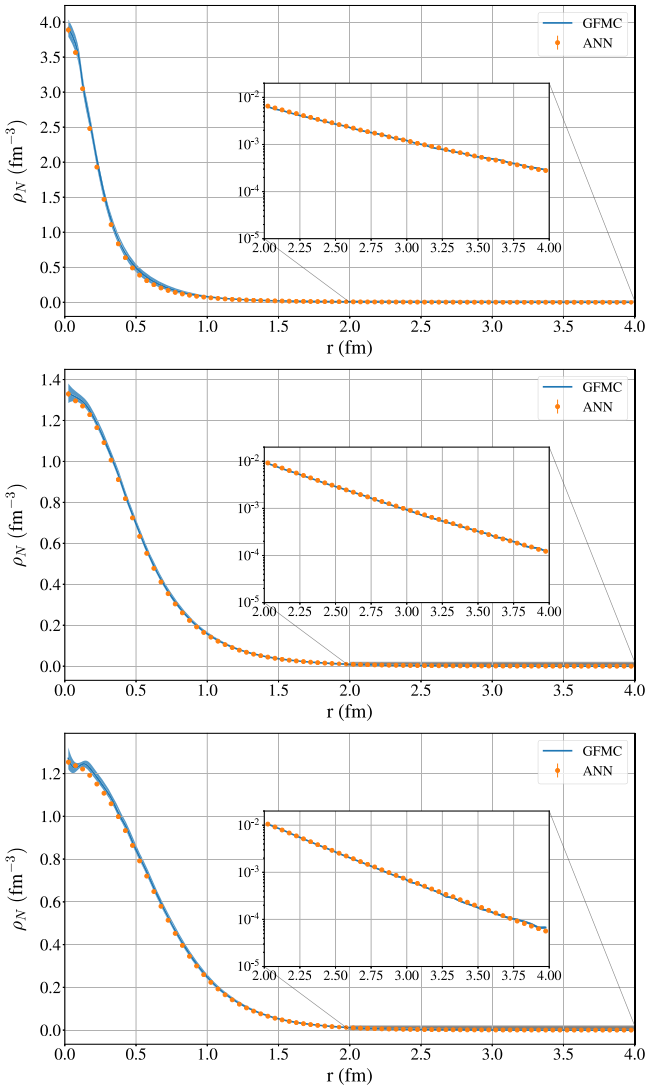


FIG. 2. Point-nucleon densities of $^2$H (upper panel), $^3$H (middle panel), and $^4$He (lower panel) for the LO pionless-EFT Hamiltonian with $\Lambda = 4$ fm$^{-1}$. The solid points and the shaded area represent the VMC-ANN and GFMC results, respectively.

for performing accurate quantum Monte Carlo studies of medium-mass nuclei.

The single-particle densities obtained with ANN wave functions are also in excellent agreement with GFMC results, both at short distances and in the slowly decaying exponential tails, which are notoriously difficult to reproduce.

The present research is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contracts DE-AC02-06CH11357, by the NUCLEI SciDAC program (A. L. and N. R.) and by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics (N. R.). This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract No. DE-AC02-06CH11357. The calculations were performed using resources of the Laboratory Computing Resource Center of Argonne National Laboratory. We acknowledge discussions with Markus Holzmann, Dean Lee, James Stokes, and James Vary. Numerical calculations were made possible also through a CINECA-INFN agreement, providing access to resources on MARCONI at CINECA.

[1] E. Epelbaum, H.-W. Hammer, and U.-G. Meissner, Modern theory of nuclear forces, Rev. Mod. Phys. **81**, 1773 (2009).
[2] R. Machleidt and D. R. Entem, Chiral effective field theory and nuclear forces, Phys. Rep. **503**, 1 (2011).
[3] H.-W. Hammer, S. König, and U. van Kolck, Nuclear effective field theory: Status and perspectives, Rev. Mod. Phys. **92**, 025004 (2020).
[4] B. R. Barrett, P. Navratil, and J. P. Vary, Ab initio no core shell model, Prog. Part. Nucl. Phys. **69**, 131 (2013).
[5] G. Hagen, T. Papenbrock, M. Hjorth-Jensen, and D. J. Dean, Coupled-cluster computations of atomic nuclei, Rep. Prog. Phys. **77**, 096302 (2014).
[6] H. Hergert, S. K. Bogner, T. D. Morris, A. Schwenk, and K. Tsukiyama, The in-medium similarity renormalization group: A novel ab initio method for nuclei, Phys. Rep. **621**, 165 (2016).
[7] A. Carbone, A. Cipollone, C. Barbieri, A. Rios, and A. Polls, Self-consistent Green's functions formalism with three-body interactions, Phys. Rev. C **88**, 054326 (2013).
[8] E. Epelbaum, H. Krebs, D. Lee, and U.-G. Meissner, Ab Initio Calculation of the Hoyle State, Phys. Rev. Lett. **106**, 192501 (2011).
[9] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, Quantum Monte Carlo methods for nuclear physics, Rev. Mod. Phys. **87**, 1067 (2015).
[10] K. E. Schmidt and S. Fantoni, A quantum Monte Carlo method for nucleon systems, Phys. Lett. B **446**, 99 (1999).
[11] M. Piarulli, I. Bombaci, D. Logoteta, A. Lovato, and R. B. Wiringa, Benchmark calculations of pure neutron matter with realistic nucleon-nucleon interactions, Phys. Rev. C **101**, 045801 (2020).
[12] D. Lonardoni, I. Tews, S. Gandolfi, and J. Carlson, Nuclear and neutron-star matter from local chiral interactions, Phys. Rev. Research **2**, 022033(R) (2020).
[13] S. Gandolfi, D. Lonardoni, A. Lovato, and M. Piarulli, Atomic nuclei from quantum Monte Carlo calculations with chiral EFT interactions, Front. Phys. **8**, 117 (2020).
[14] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).
[15] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).
[16] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).
[17] H. Saito, Method to solve quantum few-body problems with artificial neural networks, J. Phys. Soc. Jpn. **87**, 074002 (2018).
[18] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Symmetries and Many-Body Excitations with Neural-Network Quantum States, Phys. Rev. Lett. **121**, 167204 (2018).
[19] Y. Nomura, Machine learning quantum states—extensions to fermion-boson coupled systems and excited-state calculations, J. Phys. Soc. Jpn. **89**, 054706 (2020).
[20] N. Yoshioka and R. Hamazaki, Constructing neural stationary states for open quantum many-body systems, Phys. Rev. B **99**, 214306 (2019).
[21] A. Nagy and V. Savona, Variational Quantum Monte Carlo Method with a Neural-Network Ansatz for Open Quantum Systems, Phys. Rev. Lett. **122**, 250501 (2019).
[22] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Variational Neural-Network Ansatz for Steady States in Open Quantum Systems, Phys. Rev. Lett. **122**, 250503 (2019).
[23] M. J. Hartmann and G. Carleo, Neural-Network Approach to Dissipative Quantum Many-Body Dynamics, Phys. Rev. Lett. **122**, 250502 (2019).
[24] F. Ferrari, F. Becca, and J. Carrasquilla, Neural Gutzwiller-projected variational wave functions, Phys. Rev. B **100**, 125131 (2019).
[25] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, Ab initio solution of the many-electron Schrödinger equation with deep neural networks, Phys. Rev. Research **2**, 033429 (2020).
[26] J. Hermann, Z. Schätzle, and F. Noé, Deep-neural-network solution of the electronic Schrödinger equation, Nat. Chem. **12**, 891 (2020).
[27] K. Choo, A. Mezzacapo, and G. Carleo, Fermionic neural-network states for ab initio electronic structure, Nat. Commun. **11**, 2368 (2020).
[28] G. A. Negoita *et al.*, Deep learning: Extrapolation tool for ab initio nuclear theory, Phys. Rev. C **99**, 054308 (2019).
[29] W. G. Jiang, G. Hagen, and T. Papenbrock, Extrapolation of nuclear structure observables with artificial neural networks, Phys. Rev. C **100**, 054326 (2019).

[30] J. W. T. Keeble and A. Rios, Machine learning the deuteron, Phys. Lett. B **809**, 135743 (2020).
[31] P. F. Bedaque and U. van Kolck, Effective field theory for few nucleon systems, Annu. Rev. Nucl. Part. Sci. **52**, 339 (2002).
[32] N. Barnea, L. Contessi, D. Gazit, F. Pederiva, and U. van Kolck, Effective Field Theory for Lattice Nuclei, Phys. Rev. Lett. **114**, 052501 (2015).
[33] L. Contessi, A. Lovato, F. Pederiva, A. Roggero, J. Kirscher, and U. van Kolck, Ground-state properties of $^4$He and $^{16}$O extrapolated from lattice QCD with pionless EFT, Phys. Lett. B **772**, 839 (2017).
[34] J. Kirscher, N. Barnea, D. Gazit, F. Pederiva, and U. van Kolck, Spectra and scattering of light lattice nuclei from effective field theory, Phys. Rev. C **92**, 054002 (2015).
[35] C.-J. Yang, Do we know how to count powers in pionless and pionful effective field theory? Eur. Phys. J. A **56**, 96 (2020).
[36] P. F. Bedaque, H. W. Hammer, and U. van Kolck, Renormalization of the Three-Body System with Short Range Interactions, Phys. Rev. Lett. **82**, 463 (1999).
[37] J. Lomnitz-Adler, V. R. Pandharipande, and R. A. Smith, Monte Carlo calculations of triton and 4 He nuclei with the Reid potential, Nucl. Phys. **A361**, 399 (1981).
[38] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, and A. Smola, Deep sets, arXiv:1703.06114.
[39] E. Wagstaff, F. B. Fuchs, M. Engelcke, I. Posner, and M. Osborne, On the limitations of representing functions on sets, arXiv:1901.09006.
[40] C. Dugas, Y. Bengio, F. Bélisle, C. Nadeau, and R. Garcia, Incorporating second-order functional knowledge for better option pricing, in *Advances in Neural Information Processing Systems 13*, edited by T. K. Leen, T. G. Dietterich, and V. Tresp (MIT Press, Cambridge, MA, 2001), pp. 472–478.
[41] P. Massella, F. Barranco, D. Lonardoni, A. Lovato, F. Pederiva, and E. Vigezzi, Exact restoration of Galilei invariance in density functional calculations with quantum Monte Carlo, J. Phys. G **47**, 035105 (2020).
[42] R. Schiavilla, L. Girlanda, A. Gnech, A. Kievsky, A. Lovato, L. E. Marcucci, M. Piarulli, and M. Viviani, Two- and three-nucleon contact interactions and ground-state energies of light- and medium-mass nuclei, Phys. Rev. C **103**, 054003 (2021).
[43] S. Sorella, Green Function Monte Carlo with Stochastic Reconfiguration, Phys. Rev. Lett. **80**, 4558 (1998).
[44] S. Sorella, Wave function optimization in the variational Monte Carlo method, Phys. Rev. B **71**, 241103(R) (2005).
[45] S.-i. Amari, Natural gradient works efficiently in learning, Neural Comput. **10**, 251 (1998).
[46] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, (2015), software available from https://www.tensorflow.org.
[47] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, JAX: Composable transformations of Python + NumPy programs (2018), http://github.com/google/jax.
[48] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.127.022502, which includes Refs. [49–57], for details the computational algorithm, including the architecture of the network, the optimization procedure, and the scaling performance of the code.
[49] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, PointNet: Deep learning on point sets for 3D classification and segmentation, arXiv:1612.00593.
[50] O. Vinyals, S. Bengio, and M. Kudlur, Order matters: Sequence to sequence for sets, arXiv:1511.06391.
[51] J. Lee, Y. Lee, J. Kim, A. R. Kosiorek, S. Choi, and Y. W. Teh, Set transformer: A framework for attention-based permutation-invariant neural networks, arXiv:1810.00825.
[52] K. T. Schütt, H. E. Sauceda, P. J. Kindermans, A. Tkatchenko, and K. R. Müller, SchNet: A deep learning architecture for molecules and materials, J. Chem. Phys. **148**, 241722 (2018).
[53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087 (1953).
[54] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika **57**, 97 (1970).
[55] N. Qian, On the momentum term in gradient descent learning algorithms, Neural Netw. **12**, 145 (1999).
[56] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.
[57] T. Tieleman and G. Hinton, Lecture 6.5—RmsProp: Divide the gradient by a running average of its recent magnitude, *COURSERA: Neural Networks for Machine Learning* (2012), https://www.cs.toronto.edu/~hinton/coursera/lecture6/lec6.pdf.
[58] J. Toulouse and C. J. Umrigar, Optimization of quantum Monte Carlo wave functions by energy minimization, J. Chem. Phys. **126**, 084102 (2007).
[59] A. Lovato, S. Gandolfi, R. Butler, J. Carlson, E. Lusk, S. C. Pieper, and R. Schiavilla, Charge Form Factor and Sum Rules of Electromagnetic Response Functions in $^{12}$C, Phys. Rev. Lett. **111**, 092501 (2013).
[60] R. Weiss, A. Schmidt, G. A. Miller, and N. Barnea, Short-range correlations and the charge density, Phys. Lett. B **790**, 484 (2019).
[61] J. P. Vary, P. Maris, P. J. Fasano, and M. A. Caprio, Perspectives on nuclear structure and scattering with the *abinitio* no-core shell model, J. Phys. Soc. Jpn. Conf. Proc. **23**, 012001 (2018).
[62] M. Sharaf, R. McCarty, R. A. M. Basili, and J. P. Vary, Comparing sinc and harmonic oscillator basis for bound states of a Gaussian interaction, arXiv:1912.07155.

*Correction:* The affiliation indicator for the fourth author was presented incorrectly and has been fixed.