

Emulators for scarce and noisy data: application to auxiliary field diffusion Monte Carlo for the deuteron

Rahul Somasundaram,^{1,2,*} Cassandra L. Armstrong,³
Pablo Giuliani,^{4,5} Kyle Godbey,⁴ Stefano Gandolfi,¹ and Ingo Tews¹

¹*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

²*Department of Physics, Syracuse University, Syracuse, NY 13244, USA*

³*Intelligence and Space Research Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

⁴*Facility for Rare Isotope Beams, Michigan State University, East Lansing, Michigan 48824, USA*

⁵*Department of Statistics and Probability, Michigan State University, East Lansing, Michigan 48824, USA*

(Dated: April 18, 2024)

The validation, verification, and uncertainty quantification of computationally expensive theoretical models of quantum many-body systems require the construction of fast and accurate emulators. In this work, we develop emulators for auxiliary field diffusion Monte Carlo (AFDMC), a powerful many-body method for nuclear systems. We introduce a reduced-basis method (RBM) emulator for AFDMC and study it in the simple case of the deuteron. Furthermore, we compare our RBM emulator with the recently proposed parametric matrix model (PMM) that combines elements of RBMs with machine learning. We contrast these two approaches with a traditional Gaussian Process emulator. All three emulators constructed here are based on a very limited set of 5 training points, as expected for realistic AFDMC calculations, but validated against $\mathcal{O}(10^3)$ exact solutions. We find that the PMM, with emulator errors of only $\approx 0.1\%$ and speed-up factors of $\approx 10^7$, outperforms the other two emulators when applied to AFDMC.

Introduction. In the next years, an explosion of new data from laboratory experiments, such as the Facility for Rare Isotope Beams (FRIB), and multi-messenger observations of neutron stars and their mergers [1–8] will provide exciting new information for nuclear physics. To robustly analyze the information provided during this data-rich era, reliable theoretical approaches with well-quantified uncertainties are key. These approaches can then be employed in statistical tools based on Bayesian inference [1, 9–12]. Quantum Monte Carlo (QMC) methods [13], such as auxiliary field diffusion Monte Carlo (AFDMC) [14], combined with interactions from chiral effective field theory (EFT) [15–19] are some of the most universal and reliable nuclear many-body approaches used in the community. QMC algorithms can be applied to both atomic nuclei and nuclear matter using the same input interactions, which enables us to straightforwardly connect nuclear experiments with astrophysical observations. QMC methods are also very accurate and precise and provide non-perturbative, virtually exact solutions to the Schrödinger equation [13, 20]. However, these benefits incur a large computational cost, of the order of several 100,000 CPU-h per typical simulation.

The Bayesian approaches necessary for the upcoming data-rich era typically require a large number of model evaluations across a broad parameter space, rendering their application to expensive numerical approaches, such as QMC, prohibitively expensive. Emulators, i.e., algorithms that mimic the behavior of a high-fidelity (HF) model at a fraction of its computational cost, have been proposed to circumvent this problem [21–27]. They can

broadly be classified into two categories: intrusive and non-intrusive [23]. Non-intrusive or data-driven emulators are usually trained only on the inputs and outputs of the HF model and are agnostic to the underlying physics. Common examples include Gaussian process (GP) regression and artificial neural networks [26–28]. Intrusive emulators, on the other hand, usually work with high-dimensional structures (such as wave functions), and respect certain physical aspects of the underlying equations or dynamics of the system. Examples are reduced order models, such as reduced basis methods (RBMs) [21, 24, 25, 29–31], or dynamic mode decomposition and SINDy [32]. While significant effort has been devoted to developing RBM-based emulators for some many-body techniques [23, 25, 33, 34], emulators for QMC methods are in their earlier stages [35, 36]. The primary obstacle to developing emulators for QMC methods is the inability to calculate inner products between eigenstates of different Hamiltonians - something that is crucial in the framework of RBMs for quantum systems. Recently Ref. [36] proposed the floating block method as a possible solution to this obstacle, and it was successfully applied to lattice Monte Carlo calculations of light nuclei.

In this letter, we develop different new emulators for AFDMC. First, we develop an intrusive RBM-based emulator using the Petrov-Galerkin projection method [21, 22] which circumvents the need to compute overlaps between exact AFDMC eigenstates of different Hamiltonians while maintaining the fully intrusive nature of the emulator. Second, we implement the parametric matrix model (PMM) [37], a machine learning algorithm that combines elements of both intrusive and non-intrusive emulators. Last, we compare these emulation methods to a traditional, non-intrusive GP emulator. Be-

* rsomasun@syr.edu

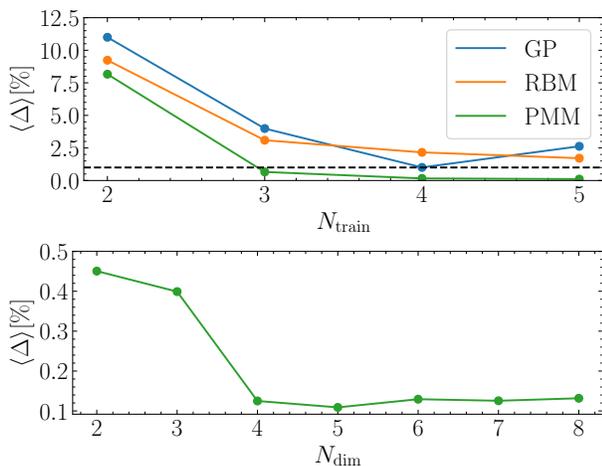


FIG. 1. Top: Averaged percentage error $\langle \Delta \rangle$ for the three emulators with respect to N_{train} . For the PMM, $N_{\text{dim}} = 5$. Bottom: $\langle \Delta \rangle$ of the PMM with respect to N_{dim} using 5 AFDMC deuteron calculations as training points. $\langle \Delta \rangle$ is computed by averaging over 1000 validation errors, each corresponding to a different validation sample.

cause AFDMC calculations are computationally expensive, when training such emulators we are limited by a very small set of training data, $N_{\text{train}} \approx 5$. Therefore, in this letter we require that our emulators achieve errors of a few percent when validated against $\mathcal{O}(10^3)$ exact solutions despite using a limited number of training data. We find that the intrusive RBM and PMM emulators far outperform the non-intrusive GP emulator, see Fig. 1. Furthermore, the PMM performs best, with an average emulation error of only $\approx 0.1\%$ but with a gain in speed of up to $\approx 10^7$ with respect to the AFDMC method.

Methods. We employ the local chiral EFT two-nucleon interactions of Refs. [15, 16, 19]. These interactions were calibrated to neutron-proton phase shifts in Ref. [19] using Bayesian inference, which results in posterior distributions over the low energy couplings (LECs). We can then write the nuclear Hamiltonian as $H(\vec{c})$, where \vec{c} is a set of control parameters, i.e., the LECs. We then employ these interactions in AFDMC, a continuum diffusion Monte Carlo (DMC) code [13, 14, 38]. Starting from a trial wave function for a specific system, AFDMC performs an evolution in imaginary time to project out the ground state of the system. All integrals appearing in this evolution are solved by means of Monte Carlo techniques. In addition, AFDMC achieves a better polynomial scaling with nucleon number than other DMC algorithms by linearizing spin-isospin states using a Hubbard-Stratonovich transformation [14]. However, as with all QMC algorithms, AFDMC results carry statistical noise. In this work, we develop emulators for the deuteron. Because the deuteron is the simplest atomic nucleus, we can obtain exact solutions in a reasonable time by solving the homogeneous part of the Lippmann-Schwinger (LS) equation [39]. Hence, it is easy to obtain validation

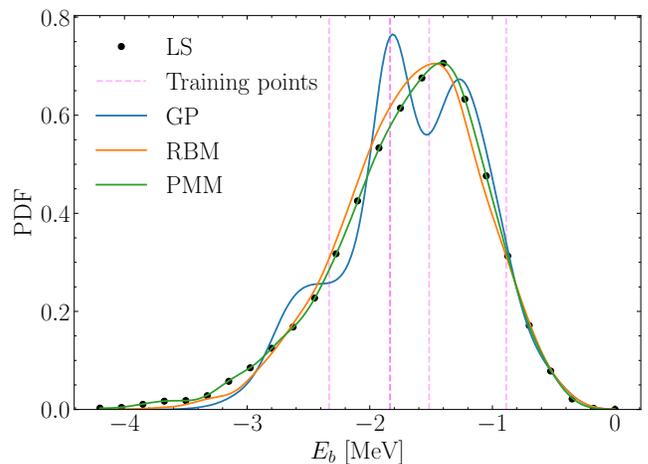


FIG. 2. Posterior distribution function (PDF) for the deuteron binding energy E_b calculated with the three emulators and the exact LS solver. The PDFs result from a Gaussian kernel density estimator calibrated to the predictions obtained from the emulators and the exact LS solver.

data. All training and validation samples are drawn from the posterior distributions of \vec{c} . We have checked that the deuteron energies calculated by the AFDMC and LS solvers agree to within 0.1%.

To construct an RBM, one typically obtains N_{train} HF solutions, i.e. ground states $\{|\psi\rangle\}_{j=1}^{N_{\text{train}}}$, corresponding to $\{\vec{c}\}_{j=1}^{N_{\text{train}}}$. Then, for \vec{c} not in the training set, one determines the ground state $|\psi\rangle$ within the subspace spanned by $\{|\psi\rangle\}_{j=1}^{N_{\text{train}}}$, i.e. we impose $|\psi\rangle \approx \sum_j a_j |\psi_j\rangle$. Under this ansatz, the Schrödinger equation becomes

$$\sum_j H|\psi_j\rangle a_j = E \sum_j |\psi_j\rangle a_j. \quad (1)$$

This equation is then projected onto a subspace spanned by “test” or “projecting” functions [21]. In several cases, these are chosen as $\langle\psi_i|$, thereby casting Eq. (1) as a generalized eigenvalue problem for the matrix $M_{ij} \equiv \langle\psi_i|H|\psi_j\rangle$ with the norm matrix $N_{ij} \equiv \langle\psi_i|\psi_j\rangle$. In QMC approaches, these overlaps are dominated by stochastic noise [13], see however Ref. [36] for a solution.

Here, we circumvent this problem using a novel application of the Petrov-Galerkin projection method [21, 30, 31] by choosing the projecting functions used to act on Eq. (1) not in the same subspace as that spanned by $\{|\psi\rangle\}_{j=1}^{N_{\text{train}}}$. Considering a different N_{train} dimensional subspace spanned by $\{|\phi\rangle\}_{i=1}^{N_{\text{train}}}$, the non-orthogonal projection results in a generalized eigenvalue problem for the matrix $\tilde{M}_{ij} \equiv \langle\phi_i|H|\psi_j\rangle$ with the corresponding norm matrix $\tilde{N}_{ij} \equiv \langle\phi_i|\psi_j\rangle$. Here, we choose the N_{train} trial wave functions that are used as initial conditions for the imaginary time evolution performed in AFDMC [13] as projecting functions. These trial wave functions are optimized using variational Monte Carlo [13] and have large overlap with the fully evolved AFDMC states $\{|\psi\rangle\}_{j=1}^{N_{\text{train}}}$.

Since the trial wave functions are analytic, all required overlaps and matrix elements can be readily computed in AFDMC. Upon solving this generalized eigenvalue problem for a non-Hermitian matrix, we discard the complex eigenvalues and then take the smallest real eigenvalue to be our physical ground state energy. In a few cases, the two smallest real eigenvalues were found to be very close to each other, less than 0.1 MeV apart. In these cases, we take the average of the two eigenvalues which does not significantly change our results.

In addition to this RBM emulator, we implement the PMM of Ref. [37] for AFDMC. Inspired by the reduced equations obtained from RBMs [40], we assume that the ground state energy of $H(\vec{c})$ can be well approximated by the lowest eigenvalue of a matrix given by

$$A(\vec{c}) = A_0 + \sum_i c_i A_i. \quad (2)$$

Here, the c_i are the LECs of the chiral Hamiltonian and we have used the fact that the LECs are affine. In contrast to the traditional RBM discussed above, we do not compute the matrix elements of A_i from AFDMC wave functions. Instead, we take a data-driven approach and infer the matrices by fitting the lowest eigenvalue of A to AFDMC results for the deuteron binding energy E_b for different \vec{c} . A global optimizer, such as the basin-hopping algorithm [41], helps find a suitable set of matrices that reproduce the desired dynamics. We impose that the A_i are real, symmetric matrices and that A_0 is diagonal [37]. The dimensionality of the matrices A_i , N_{dim} , is a hyperparameter of the emulator.

We compare our RBM and PMM emulators with a non-intrusive GP emulator. For the GP kernel, we use a linear combination of the Matérn and the dot-product kernels [28], since many of the other standard kernels (such as the radial basis function) performed poorly in comparison. The optimization of the kernel hyperparameters was performed using the python package SCIKIT-LEARN.

Results. We first discuss our results for leading-order (LO) chiral EFT interactions. At LO, we draw 5 samples from the posterior distribution on the single spectral LEC in the deuteron channel, C_{3S_1} , obtained in Ref. [19]. These 5 interactions are then used in AFDMC calculations of E_b as well as the various overlaps required for our RBM emulator. This set of 5 samples constitutes our training set. We draw a different set of 1000 samples from the same posterior distribution on C_{3S_1} and compute E_b for each interaction by solving the LS equation. These calculations are used for validation. For the training samples, the AFDMC and LS results agree on the sub-percent level.

In Fig 1, we show the performance of the three emulators as a function of N_{train} . The averaged percentage error is given as,

$$\langle \Delta \rangle [\%] = \frac{1}{N} \sum_i \left| \frac{E_{b,i}^{\text{pred}} - E_{b,i}^{\text{LS}}}{E_{b,i}^{\text{LS}}} \right| \times 100, \quad (3)$$

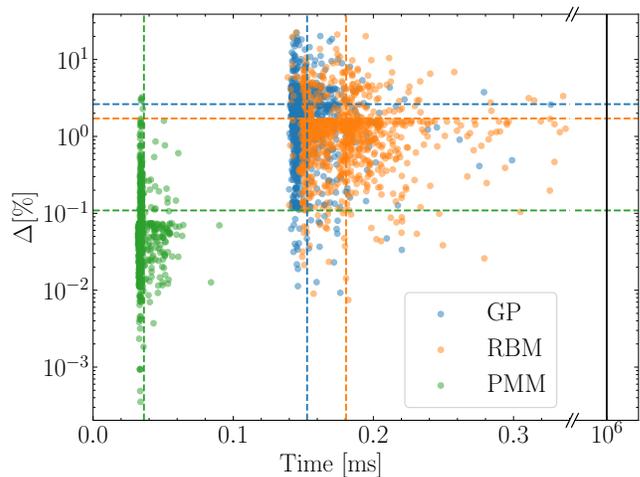


FIG. 3. Computational accuracy vs. time plot [42] for each model evaluation for our three emulators. Each dot corresponds to a validation sample. The dashed lines indicate the averages over all validation samples. The total CPU time required to perform a full AFDMC calculation is $\mathcal{O}(10^3)$ s, which is indicated by the solid black vertical line.

where $E_{b,i}^{\text{pred}}$ ($E_{b,i}^{\text{LS}}$) is the deuteron binding energy predicted by the emulator (LS solver) for sample i , and the sum is over the N validation samples. The PMM clearly performs better than the other two methods, achieving sub-percentage emulation errors for $N_{\text{train}} \geq 3$. We have checked that adding terms non-linear in c_i to Eq. (2) does not generally improve the accuracy of the PMM. We have also studied the performance of the PMM with respect to N_{dim} and found that it does not improve for $N_{\text{dim}} \gtrsim 5$. Therefore, we will use $N_{\text{dim}} = 5$ at leading order (LO) in the rest of this letter. The RBM emulator performs worse than the PMM. With $N_{\text{train}} = 5$, the RBM achieves an average percentage error of 1.8%, which is comparable to other RBMs reported in the literature [21, 36]. Finally, we find that the GP emulator generally performs worst.

Next, we fix $N_{\text{train}} = 5$ and calculate the deuteron binding energy for the 1000 samples in our validation set. This results in emulated posterior distribution functions (PDF) on the binding energy shown in Fig. 2. The non-Gaussian PDF estimated from the exact LS solver is shown as reference. We see that the PDFs obtained from the PMM and the exact LS solver are virtually indistinguishable. The Kullback–Leibler (KL) divergence between the two distributions is 5×10^{-5} . On the other hand, the KL divergence of the PDF predicted by the RBM (GP) emulator with respect to the LS solution is 0.008 (0.03). We find that the RBM emulator performs comparably well and captures the general shape of the PDF which results only in small differences for the percentiles of the PDF. In contrast, the PDF obtained from the GP emulator shows significant deviations from the exact one. This demonstrates the benefit of building intrusive or hybrid emulators when the training data set

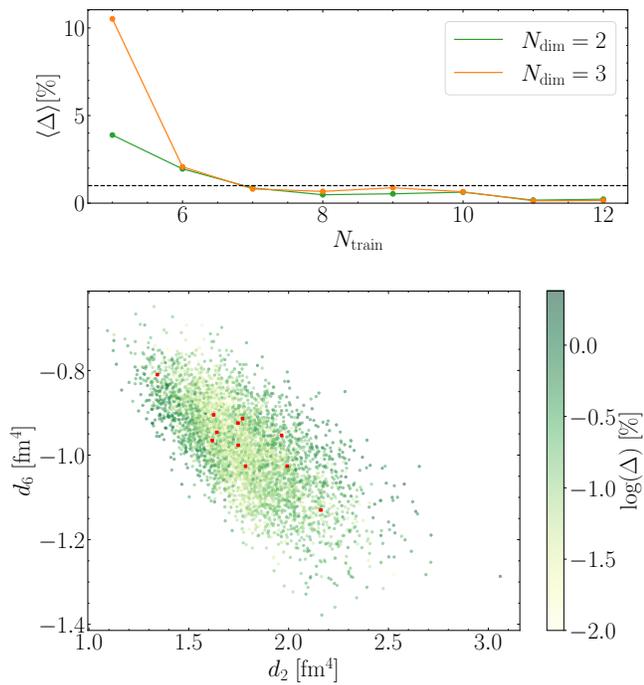


FIG. 4. Top: Behaviour of the PMM for local $N^2\text{LO}$ interactions with respect to the number of training samples. The average percentage error is calculated by validating against 5000 samples. Bottom: Relative error Δ of the PMM emulator for the 5000 validation samples with respect to the LECs d_2 and d_6 . The values of d_2 and d_6 for the training points are shown are red.

is limited. The 5 training points range from -2.3 MeV to -0.9 MeV and we find that the PMM accurately interpolates and extrapolates. The RBM performs well for interpolation but fails for extrapolation. Finally, the GP emulator fails on both counts due to the limited nature of the data set.

Fig. 3 depicts the percentage error for the emulators as a function of the computation time for each validation sample. We find that the PMM, in addition to being more accurate than the other two methods, is faster by almost an order of magnitude. The spread in the errors provides an estimate of the outliers present in the data. We find that the biggest outlier for the PMM has an error of $\approx 3\%$. On the other hand, both the RBM and the GP contain a significant number of outliers with errors as large as 20%. While expected for the GP, this is somewhat surprising for the RBM given its intrusive nature. Note that the emulators for QMC of Ref. [36] result in even bigger outliers. We found that the AFDMC calculations of the overlaps $\langle \phi_i | \psi_j \rangle$ are highly correlated throughout the imaginary-time evolution. As a consequence, our results carry larger stochastic noise. We will investigate this AFDMC noise for different choices of nuclear systems, trial wave functions, and nuclear Hamiltonians in the future.

So far, we have considered only LO interactions that

contain only one spectral LEC in the deuteron channel. We now study a larger parameter space by employing interactions at $N^2\text{LO}$. At this order, four independent spectral LECs contribute to the deuteron. As we found the PMM to perform best at LO, we focus on the PMM for $N^2\text{LO}$ and draw up to 15 samples from the posteriors on the LECs calculated in Ref. [19] for training. Since the PMM requires only the energy of the deuteron for training, we generate the binding energies using the LS solver. For validation, we now use 5000 samples drawn from the same posterior. We show our results for PMMs with $N_{\text{dim}} = 2$ and $N_{\text{dim}} = 3$ in Fig. 4 (top). We find that both PMMs achieve sub-percentage accuracies for $N_{\text{train}} \geq 7$, even in a 4 dimensional parameter space. For the PMM with $N_{\text{dim}} = 2$ and $N_{\text{train}} = 11$, the average error is 0.2% with the largest outlier having an error of 2.4%. Similar to the results of Fig. 2, the PDF on the deuteron binding energy calculated with this PMM has a vanishingly small KL divergence of 3×10^{-4} with respect to the PDF evaluated with the LS solver. Finally, we see that the PMM with $N_{\text{dim}} = 3$ performs slightly worse than the one with $N_{\text{dim}} = 2$ for $N_{\text{train}} < 10$. We conclude that in a 4–5 dimensional parameter space with a limited training data set ($N_{\text{train}} \approx 10$), PMMs with small values for the hyperparameter N_{dim} are sufficient.

For the PMM with $N_{\text{dim}} = 2$ and $N_{\text{train}} = 11$, we further study the relative errors within the parameter space spanned by the 4 relevant spectral LECs. In Fig. 4 (bottom), we show the relative error Δ of the PMM for all of our validation samples. We limit ourselves to a projection onto the plane spanned by the tensor coupling d_2 and the spin-orbit coupling d_6 but other LEC choices lead to similar results. We find that the PMM both interpolates and extrapolates well even when the number of unknowns (14) is larger than the number of training points (11). The PMM is still very effective for parameter values far from those of the training points and we see some regions of particularly good performance outside of the range of training points. Note that the same point in $d_2 - d_6$ space can have different Δ due to different values of the other LECs.

Conclusion. We have developed three emulators for AFDMC calculations of the deuteron: a Petrov-Galerkin RBM-based emulator, a PMM, and a GP emulator. We found the RBM and PMM to be generally superior to a traditional GP regression, both performing well despite being trained on a very limited data set. We expect that the RBM performance can be further optimized in the future. This is an important goal because these intrusive emulators give access to the complete wave functions of the many-body system, which allows us to easily compute other matrix elements of interest. On the other hand, we have demonstrated that PMMs are already a good choice for the purposes of emulating the ground state energy of these systems. In addition to their performance, the PMM’s almost straightforward set-up establishes them as a valuable new emulation tool for a broad class of systems. We believe that

the calculations and tools developed in this work will enable novel applications of chiral interactions and QMC methods, such as their implementation in data analyses pipelines used to interpret multi-messenger neutron star observations.

We thank J. Carlson, B. Reed, and R. Weiss for insightful discussions. R.S. acknowledges support from the Nuclear Physics from Multi-Messenger Mergers (NP3M) Focused Research Hub which is funded by the National Science Foundation under Grant Number 21-16686, and by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20220541ECR. C.L.A. was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20230315ER. I.T. and S.G. were supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract No. DE-AC52-

06NA25396 and by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) NUCLEI program. I.T. was also supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project numbers 20220541ECR and 20230315ER. P.G. and K.G. were supported by the National Science Foundation CSSI program under award No. OAC-2004601 (BAND Collaboration). Computational resources have been provided by the Los Alamos National Laboratory Institutional Computing Program, which is supported by the U.S. Department of Energy National Nuclear Security Administration under Contract No. 89233218CNA000001, and by the National Energy Research Scientific Computing Center (NERSC), which is supported by the U.S. Department of Energy, Office of Science, under contract No. DE-AC02-05CH11231.

-
- [1] H. Koehn *et al.*, (2024), [arXiv:2402.04172 \[astro-ph.HE\]](#).
- [2] H. T. Cromartie *et al.* (NANOGrav), *Nature Astron.* **4**, 72 (2019), [arXiv:1904.06759 \[astro-ph.HE\]](#).
- [3] B. P. Abbott *et al.* (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAAS-TRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT), *Astrophys. J. Lett.* **848**, L12 (2017), [arXiv:1710.05833 \[astro-ph.HE\]](#).
- [4] B. P. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. Lett.* **119**, 161101 (2017), [arXiv:1710.05832 \[gr-qc\]](#).
- [5] T. E. Riley *et al.*, *Astrophys. J. Lett.* **887**, L21 (2019), [arXiv:1912.05702 \[astro-ph.HE\]](#).
- [6] T. E. Riley *et al.*, *Astrophys. J. Lett.* **918**, L27 (2021), [arXiv:2105.06980 \[astro-ph.HE\]](#).
- [7] M. C. Miller *et al.*, *Astrophys. J. Lett.* **887**, L24 (2019), [arXiv:1912.05705 \[astro-ph.HE\]](#).
- [8] M. C. Miller *et al.*, *Astrophys. J. Lett.* **918**, L28 (2021), [arXiv:2105.06979 \[astro-ph.HE\]](#).
- [9] P. T. H. Pang *et al.*, *Nature Commun.* **14**, 8352 (2023), [arXiv:2205.08513 \[astro-ph.HE\]](#).
- [10] G. Raaijmakers *et al.*, *Astrophys. J. Lett.* **887**, L22 (2019), [arXiv:1912.05703 \[astro-ph.HE\]](#).
- [11] J. N. Farr, Z. Meisel, and A. W. Steiner, (2021), [arXiv:2111.11536 \[nucl-th\]](#).
- [12] D. R. Phillips, R. J. Furnstahl, U. Heinz, T. Maiti, W. Nazarewicz, F. M. Nunes, M. Plumlee, M. T. Prato, S. Pratt, F. G. Viens, and S. M. Wild, *J. Phys. G Nucl. Part. Phys.* **48**, 072001 (2021).
- [13] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, *Rev. Mod. Phys.* **87**, 1067 (2015), [arXiv:1412.3081 \[nucl-th\]](#).
- [14] K. E. Schmidt and S. Fantoni, *Phys. Lett. B* **446**, 99 (1999).
- [15] A. Gezerlis, I. Tews, E. Epelbaum, S. Gandolfi, K. Hebeler, A. Nogga, and A. Schwenk, *Phys. Rev. Lett.* **111**, 032501 (2013), [arXiv:1303.6243 \[nucl-th\]](#).
- [16] A. Gezerlis, I. Tews, E. Epelbaum, M. Freunek, S. Gandolfi, K. Hebeler, A. Nogga, and A. Schwenk, *Phys. Rev. C* **90**, 054323 (2014), [arXiv:1406.0454 \[nucl-th\]](#).
- [17] I. Tews, S. Gandolfi, A. Gezerlis, and A. Schwenk, *Phys. Rev. C* **93**, 024305 (2016), [arXiv:1507.05561 \[nucl-th\]](#).
- [18] J. E. Lynn, I. Tews, J. Carlson, S. Gandolfi, A. Gezerlis, K. E. Schmidt, and A. Schwenk, *Phys. Rev. Lett.* **116**, 062501 (2016), [arXiv:1509.03470 \[nucl-th\]](#).
- [19] R. Somasundaram, J. E. Lynn, L. Huth, A. Schwenk, and I. Tews, *Phys. Rev. C* **109**, 034005 (2024), [arXiv:2306.13579 \[nucl-th\]](#).
- [20] D. Lonardoni, J. Carlson, S. Gandolfi, J. E. Lynn, K. E. Schmidt, A. Schwenk, and X. Wang, *Phys. Rev. Lett.* **120**, 122502 (2018), [arXiv:1709.09143 \[nucl-th\]](#).
- [21] E. Bonilla, P. Giuliani, K. Godbey, and D. Lee, *Phys. Rev. C* **106**, 054322 (2022), [arXiv:2203.05284 \[nucl-th\]](#).
- [22] J. A. Melendez, C. Drischler, R. J. Furnstahl, A. J. Garcia, and X. Zhang, *J. Phys. G* **49**, 102001 (2022), [arXiv:2203.05528 \[nucl-th\]](#).
- [23] T. Duguet, A. Ekström, R. J. Furnstahl, S. König, and D. Lee, (2023), [arXiv:2310.19419 \[nucl-th\]](#).
- [24] D. Frame, R. He, I. Ipsen, D. Lee, D. Lee, and E. Rrapaj, *Phys. Rev. Lett.* **121**, 032501 (2018), [arXiv:1711.07090 \[nucl-th\]](#).
- [25] S. König, A. Ekström, K. Hebeler, D. Lee, and A. Schwenk, *Phys. Lett. B* **810**, 135814 (2020), [arXiv:1909.08446 \[nucl-th\]](#).
- [26] D. Lay, E. Flynn, S. A. Giuliani, W. Nazarewicz,

- and L. Neufcourt, *Phys. Rev. C* **109**, 044305 (2024), [arXiv:2310.01608 \[nucl-th\]](#).
- [27] Ö. Sürer, F. M. Nunes, M. Plumlee, and S. M. Wild, *Physical Review C* **106**, 024607 (2022).
- [28] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, 2005).
- [29] J. S. Hesthaven, G. Rozza, B. Stamm, *et al.*, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Vol. 590 (Springer, 2016).
- [30] J. N. Reddy, *An introduction to the finite element method*, 3rd ed. (McGraw-Hill Higher Education, New York, NY, 2006).
- [31] C. A. Fletcher, *Computational galerkin methods* (Springer, 1984).
- [32] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control* (Cambridge University Press, 2019).
- [33] W. G. Jiang, C. Forssén, T. Djärv, and G. Hagen, (2022), [arXiv:2212.13203 \[nucl-th\]](#).
- [34] W. G. Jiang, C. Forssén, T. Djärv, and G. Hagen, (2022), [arXiv:2212.13216 \[nucl-th\]](#).
- [35] D. K. Frame, *Ab initio simulations of light nuclear systems using eigenvector continuation and auxiliary field Monte Carlo*, Ph.D. thesis, Michigan State U., Michigan State U. (2019), [arXiv:1905.02782 \[nucl-th\]](#).
- [36] A. Sarkar, D. Lee, and U.-G. Meißner, *Phys. Rev. Lett.* **131**, 242503 (2023), [arXiv:2306.11439 \[nucl-th\]](#).
- [37] P. Cook, D. Jammooa, M. Hjorth-Jensen, D. D. Lee, and D. Lee, (2024), [arXiv:2401.11694 \[cs.LG\]](#).
- [38] J. E. Lynn, I. Tews, S. Gandolfi, and A. Lovato, *Ann. Rev. Nucl. Part. Sci.* **69**, 279 (2019), [arXiv:1901.04868 \[nucl-th\]](#).
- [39] E. Epelbaum, W. Glockle, and U.-G. Meissner, *Nucl. Phys. A* **747**, 362 (2005), [arXiv:nucl-th/0405048](#).
- [40] K. Godbey, P. Giuliani, E. Bonilla, E. Flynn, D. Odell, K. Beyer, D. Lay, D. Figueroa, R. Garg, and M. Campbell, “Dimensionality reduction in nuclear physics,” https://dr.ascsn.net/blackbox_galerkin/Introduction.html.
- [41] D. J. Wales and J. P. Doye, *The Journal of Physical Chemistry A* **101**, 5111 (1997).
- [42] D. Odell, P. Giuliani, K. Beyer, M. Catacora-Rios, M. Y.-H. Chan, E. Bonilla, R. J. Furnstahl, K. Godbey, and F. M. Nunes, *Phys. Rev. C* **109**, 044612 (2024).