

Generative Modeling of Nucleon-Nucleon Interactions

Pengsheng Wen^{1,2,*}, Jeremy W. Holt^{1,2,†}, and Maggie Li^{3,4,5,‡}

¹*Cyclotron Institute, Texas A&M University, College Station, Texas 77843, USA*

²*Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843, USA*

³*Cornell University, Ithaca, New York 14853, USA*

⁴*Department of Astronomy, Cornell University, Ithaca, New York 14853, USA*

⁵*Cahill Center for Astrophysics, California Institute of Technology,*

MC 249-17, 1200 E California Boulevard, Pasadena, California 91125, USA

(Received 11 July 2023; revised 27 June 2024; accepted 18 November 2024; published 20 December 2024)

Developing high-precision models of the nuclear force and propagating the associated uncertainties in quantum many-body calculations of nuclei and nuclear matter remain key challenges for *ab initio* nuclear theory. In this Letter, we demonstrate that generative machine learning models can construct novel instances of the nucleon-nucleon interaction when trained on existing potentials from the literature. In particular, we train the generative model on nucleon-nucleon potentials derived at second and third order in chiral effective field theory and at three different choices of the resolution scale. We then show that the model can be used to generate samples of the nucleon-nucleon potential drawn from a continuous distribution in the resolution scale parameter space. The generated potentials are shown to produce high-quality nucleon-nucleon scattering phase shifts. This work provides an important step toward a comprehensive estimation of theoretical uncertainties in nuclear many-body calculations that arise from the arbitrary choice of nuclear interaction and resolution scale.

DOI: 10.1103/PhysRevLett.133.252501

Introduction—High-precision models of the nuclear interaction are essential not only for explaining the structure and dynamics of atomic nuclei [1,2] but also for describing the properties of hot and dense matter in extreme astrophysical environments, such as neutron stars, core-collapse supernovae, and neutron star mergers [3,4]. In this Letter, we show that modern machine learning generative models have the ability to learn salient features of the nucleon-nucleon (NN) interaction, reconstruct distributions of potentials, and create novel and physically reasonable instances of the nuclear potential. Generative models have become a popular research topic in machine learning due to their ability to generate new data based on existing data distributions, which has many practical applications to image and speech synthesis [5–8]. Here, we apply generative models as a tool for uncertainty quantification in nuclear many-body calculations, which depend on the resolution scale of the employed nuclear force. This work is related to other recent investigations across diverse fields that have explored generative models in the context of the renormalization group and uncertainty quantification [9–12].

Since the nuclear potential arises from quantum chromodynamics as the low-energy effective interaction among composite nucleons, there is an inherent uncertainty due to

the choice of resolution scale at which nuclear dynamics is resolved [13,14]. This resolution scale is typically parametrized in the nuclear potential through a momentum-space regulating function parameter Λ that demarcates the separation between low-energy and high-energy (unresolved) physics. In principle, physical observables should be independent of the resolution scale, but in practice, the results of nuclear many-body calculations exhibit a moderate residual uncertainty due to this choice [15–17]. In recent years, chiral effective field theory (EFT) [18–20] has emerged as a suitable tool to systematically study not only the scale dependence of nuclear interactions but also to estimate uncertainties due to missing physics through an analysis of effective field theory truncation errors [4,21–31]. In the framework of chiral EFT, a specific nuclear interaction is obtained by first defining the high-momentum regulating function and then fitting to experimental scattering and bound state data the low-energy constants (LECs) of the theory that characterize unresolved short-distance physics [25,32].

Nuclear potentials from chiral EFT are typically fitted at only a few select values of Λ , from which one can obtain a qualitative estimate of the resolution scale uncertainties. For statistical inference, however, one requires the ability to draw samples of the NN potential from a continuous distribution in Λ , which when combined with EFT truncation errors and variations in the LECs consistent with data [33,34] can provide a comprehensive assessment of theory

*Contact author: pswen2019@physics.tamu.edu

†Contact author: holt@physics.tamu.edu

‡Contact author: maggieli@caltech.edu

uncertainties. For this purpose, we utilize the Generative Flow (Glow) model [35,36], which was originally developed in the field of computer vision for generating realistic images and manipulating their attributes. We adapt and refine the Glow model to develop a generative machine learning model for nuclear potentials. Effectively, we will treat the momentum-space matrix elements of the potential in different partial waves as “images.” We show that the model can recreate the training potentials and generate new physically reasonable nucleon-nucleon potentials over a continuum of cutoff scales. The reliability of the generated potentials is benchmarked by calculating nucleon-nucleon scattering phase shifts. Finally, we show that a combination of the Glow model and Vision Transformer model (ViT) [37] allows for the extraction of chiral EFT LECs from the generated nuclear potential matrix elements.

Methods—The Glow model attempts to learn the properties of a given dataset by constructing an appropriate probability distribution to estimate the probabilities of some features of a sample. Once an appropriate probability distribution is found, the Glow model can generate novel samples by drawing from the distribution through sampling. The Glow model initializes a trainable distribution whose parameters θ are iteratively adjusted such that the likelihood of the distribution can be maximized with respect to θ . In practice, one minimizes [38]

$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_i \log p_\theta(\mathbf{x}_i), \quad (1)$$

where $p_\theta(\mathbf{x}_i)$ is the probability of the sample \mathbf{x}_i , θ represents trainable parameters, and N is the number of samples.

As a flow-based model, Glow couples several layers of transformation functions to form the flow f , which transforms a sample \mathbf{x} from the data space to the latent space variable \mathbf{z} via

$$\mathbf{z} = f_K \circ f_{K-1} \circ \dots \circ f_k \circ \dots \circ f_1(\mathbf{x}), \quad (2)$$

where f_k is the transformation function at the k th layer. In this work, a layer of transformation is composed of three parts: an ActNorm, a 1×1 convolution transformation (1×1 conv) [39], and a rational quadratic spline (RQS) transformation [40,41]. The Glow model incorporates multiple-scale architecture [36] for efficiency purposes. A sequence of operations is performed on each sample at every scale. The first operation, squeezing, changes the shape of a sample from (C, H, W) to $(4 \times C, H/2, W/2)$, where C is the number of channels, and H and W are the height and width, respectively. At the second operation, K layers of transformation functions are coupled. The sample now is transformed according to Eq. (2). Upon reaching the end of the scale, the sample is again split into two halves along the channel dimension. The first half is

treated as part of the fully transformed sample \mathbf{z} in the resulting latent space. Its probability will be calculated by a tractable base distribution model whose parameters are the output of neural networks, and contribute to the total probability $q_\theta(\mathbf{z})$. The remaining half is employed as the input of the neural networks responsible for constructing the base distribution and will continue progressing to the next scale. A tensor with the shape of (C, H, W) traverses through L -layer scales and finally there are still $(2^L \times C, H/2^L, W/2^L)$ elements whose distribution has not been determined. The base distribution to evaluate the probabilities of final elements is built by neural networks whose inputs are the labels of the sample. In our work, the base distributions are Gaussian distributions.

To produce samples with new labels, one obtains samples in latent space and transforms them back into data space. For this we apply two different approaches: (1) *latent space sampling* (LSS) and (2) *latent space interpolation* (LSI). In LSS we first draw samples \mathbf{z} in the latent space from a Gaussian that is built according to the labels from the training set. Next, we perform interpolation based on the obtained \mathbf{z} samples and the new label. Finally, we transform the interpolated sample back into the data space. In LSI the samples of \mathbf{z} are obtained by transforming given \mathbf{x} ’s from the data space to the latent space. The next two steps are the same as in LSS approach. The first approach is suitable for drawing samples from distributions with desired labels, while the second approach can be applied to predict instances of samples based on existing samples. For nuclear interaction modeling, we view the partial wave matrix elements as the samples for the Glow model. Therefore, a nuclear potential can be viewed as a sample with shape (C, H, W) , where C is the number of partial wave channels α , and H and W are the number of momentum-space mesh points of p and p' for the potential $V_\alpha(p, p')$.

Results—We conduct two experiments to investigate the Glow model’s capabilities in reconstructing actual nuclear potentials, building distributions, and generating new, realistic ones. Our first experiment employs potentials generated from the similarity renormalization group (SRG) [42,43], where the resolution scale can be freely chosen. In the SRG, a potential V_λ is evolved with a flow parameter λ from an initial potential V by a unitary transformation. In this Letter, the initial potential for the SRG is the next-to-next-to-leading order (n3lo) chiral nuclear potential of Entem and Machleidt [44] with momentum-space cutoff $\Lambda = 500$ MeV. The second experiment attempts to train the Glow model to learn the characteristics of chiral potentials $n\nu\text{lo}\Lambda$ at different orders ν in the chiral expansion and different high-momentum cutoff values Λ . The SRG potential experiments include three partial-wave channels $C = 3$, and the size of the momentum-space mesh grid is $H = W = 32$. In the chiral potential experiments, we include all partial waves that have associated short-distance

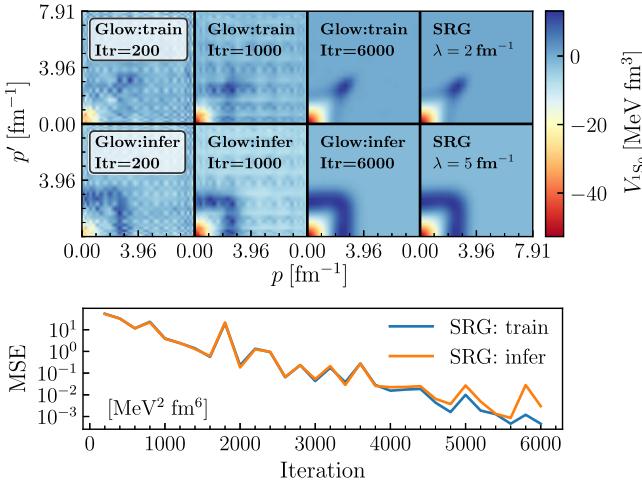


FIG. 1. Top: V_{1S_0} matrix elements of Glow-generated SRG potentials at $\lambda = 2$ and 5 fm^{-1} at different training iterations. The true SRG potential matrix elements are shown on the right side for comparison. Bottom: the mean squared error, calculated based on the difference of all the elements from all the corresponding potentials.

contact terms, which requires $C = 14$, and we set $H = W = 48$. In addition, $(L, K) = (3, 2)$ for the SRG potential experiments, and $(L, K) = (4, 4)$ for the chiral potential experiments.

To train a Glow model for the SRG, we gather potentials with $\lambda \in \{2, 3, 4, 7, 8, 10, 11, 12\} \text{ fm}^{-1}$. We then employ the Glow model to deduce potentials with $\lambda = 5, 6, 9, 13 \text{ fm}^{-1}$ using LSI throughout the training to evaluate the model's inference capability. When using LSI, the input potentials are the actual SRG potentials used for training. We calculate the mean square error (MSE) by measuring the difference between all the matrix elements of Glow-generated potentials and the actual potentials from SRG. In the bottom panel of Fig. 1, we show the MSE as a function of training iteration. During the training process, both the reconstructed potentials (blue line) and the inferred potentials (orange line) converge toward the true potentials generated by the SRG, as indicated by the decreasing MSE. The top panel of Fig. 1 shows the $\lambda = 2 \text{ fm}^{-1}$ (training) and $\lambda = 5 \text{ fm}^{-1}$ (inferred) Glow-generated SRG potential matrix elements in the 1S_0 partial wave at different iterations (three columns on the left) and the actual potentials from SRG (rightmost column). Initially, the Glow model only generates Gaussian noise. Upon completion of the training, the potentials obtained from the Glow model and the SRG appear indistinguishable.

Next, we train a separate Glow model based on a set of six chiral potentials with truncation orders $\nu \in \{2, 3\}$ and cutoffs $\Lambda \in \{450, 500, 550\} \text{ MeV}$ [45]. Since the chiral potentials generated by the Glow model are given in terms of their partial-wave matrix elements, we have also trained a ViT model [37,46] to deduce the associated LECs and

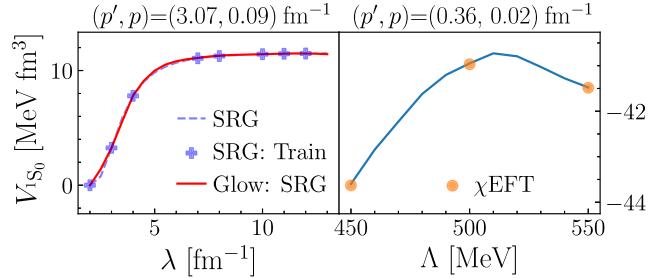


FIG. 2. SRG (left) and chiral (right) off-diagonal potential matrix elements in the 1S_0 partial-wave channel. The plus and dot markers as well as the dashed lines are from the exact potentials, while solid lines are generated by trained Glow models using LSI and LSS for SRG and chiral potentials, respectively.

value of Λ from Glow-generated potentials. The ViT model is a widely used machine learning model that is well-suited for regression and classification tasks. In the present case, the ViT model is trained on a large dataset of (unphysical) chiral potentials with LECs randomly sampled from a uniform distribution.

In Fig. 2, we show the cutoff dependence of selected 1S_0 off-diagonal matrix elements of the Glow-generated SRG potentials (left panel) and chiral potentials (right panel). The Glow-generated SRG potentials are obtained from LSI, where the inputs for LSI are the actual preexisting potentials, while the Glow-generated chiral potentials are generated via LSS. In the left panel, the plus symbols mark the actual values of the SRG potential matrix elements at the values of λ used for training, which naturally lie exactly on the dashed-blue line that indicates the actual SRG matrix elements as a function of λ . The matrix elements from the Glow model are shown with the solid red line, which lies nearly on top of the blue dashed line of the actual SRG matrix elements. We note that the Glow model is not naively interpolating between training points, since the points between $\lambda = 4$ and 7 fm^{-1} would then be much farther from the blue-dashed SRG line. In addition, from the left panel of Fig. 2 we see that the Glow model produces the expected [47] decay $V_\lambda(p_1, p_2) \approx V(p_1, p_2) e^{-(p_1^2 - p_2^2)^2/4}$ of the far off-diagonal matrix element $\langle p' = 3.07 \text{ fm}^{-1} | V_{1S_0} | p = 0.09 \text{ fm}^{-1} \rangle$ with respect to λ . Similarly, in the right panel of Fig. 2, we see that the chiral EFT matrix elements $\langle p' = 0.36 \text{ fm}^{-1} | V_{1S_0} | p = 0.02 \text{ fm}^{-1} \rangle$ generated by the Glow model across a continuum of Λ values follow a smooth contour that passes through the actual values at $\Lambda = 450, 500, 550 \text{ MeV}$.

The Glow model can also be used to obtain LEC distributions either (i) starting from scratch by directly fitting to phase shifts or (ii) from existing LEC distributions at different resolution scales. To demonstrate (i), in the left panel of Fig. 3 we start by training a Glow model at three different cutoffs, $\Lambda = \{450(\text{cyan}), 500(\text{orange}), 550(\text{magenta})\} \text{ MeV}$, using a wide distribution (top panel,

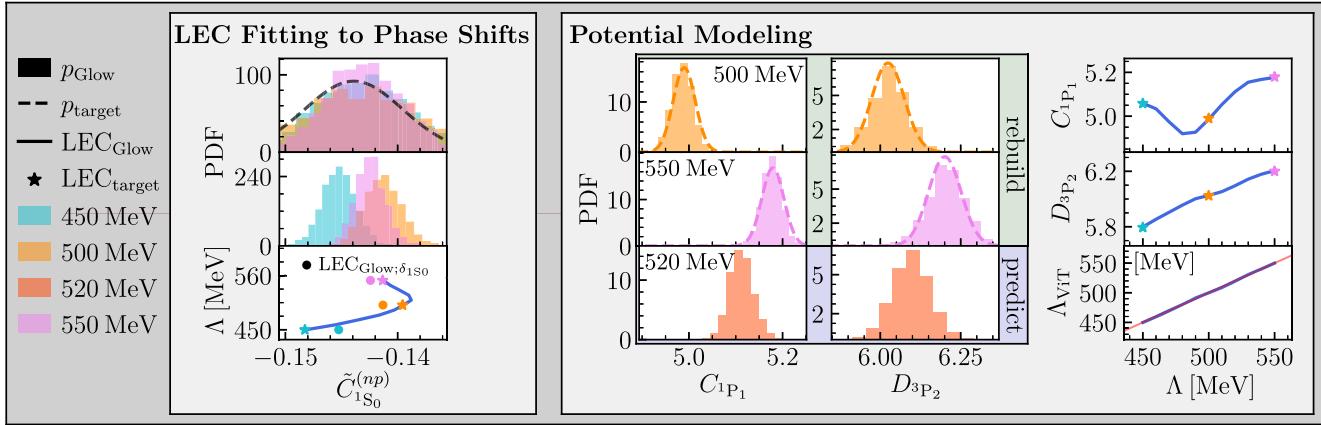


FIG. 3. Left panel: distributions for \tilde{C}_{1S_0} before (top row) and after (middle row) phase shift training. The mean values of the distributions after phase shift training are shown as circles, while the literature values are shown as stars. Right panel: The first two columns show LEC distributions predicted using LSS for n3lo chiral potentials with $\Lambda = 500, 550, 520$ MeV. We show for comparison the input distributions as dashed lines. The rightmost column shows the evolution of the LEC distribution mode as a function of the cutoff, where the stars are the LECs values from the actual chiral n3lo potentials. The bottom-right panel shows the ViT predicted cutoff vs the cutoff input label into the Glow model. A red diagonal line (bottom right) is plotted to gauge the quality of the Glow-learned and ViT-extracted cutoffs.

dashed line) for the LEC $\tilde{C}_{1S_0}^{(np)}$. The Glow model is able to rebuild this wide distribution at all three values of Λ as shown by the different colored distributions in the top panel. The pretrained Glow model is then further trained to minimize the difference between the calculated and experimental phase shifts δ_{1S_0} in the neutron-proton channel up to the laboratory energy of 300 MeV. As shown in the middle panel of the left subfigure in Fig. 3, after the phase shift training, the distributions for the $C_{1S_0}^{(np)}$ LEC extracted by the ViT model converge separately for the three values of Λ . Therefore, the Glow model can obtain LEC distributions for different cutoffs from scratch. In the bottom panel of the left subfigure, we show the peak values of $C_{1S_0}^{(np)}$ determined from the above method (circle symbols) compared to the actual values from the literature (star symbols). The values are not identical, since the literature values were obtained by fitting to experimental scattering data and deuteron properties. Nevertheless, the LEC values are quite similar and show the same trend as Λ is varied. Finally, the obtained potentials produce high-quality phase shifts as shown with the triangle symbols in the left panel of Fig. 4.

In the right panel of Fig. 3, we show the conditional distributions of LECs for given Λ s in the first two columns, where the potentials are sampled via LSS. Again, we find that the Glow model has the ability to rebuild the LEC distribution at the training values of $\Lambda = 500, 550$ MeV (top two rows). We also show in the third row the predicted distribution $p(\text{LECs}|\Lambda = 520$ MeV). The central point of the predicted distribution is positioned within the range between the centers of $p(\text{LECs}|\Lambda = 500$ MeV) and $p(\text{LECs}|\Lambda = 550$ MeV). Finally, the right column shows

the continuous evolution of the peak in the LEC distributions between $\Lambda = 450$ – 550 MeV, where the actual values are shown as stars for comparison. We also show in the bottom-right panel of Fig. 3 the relation between the cutoff Λ_{Glow} as the input label for the Glow model and the cutoff Λ_{ViT} inferred by the ViT model. The near identity between the two Λ values shows that the Glow model can generate chiral potentials with the desired value of Λ and that the ViT model can accurately extract Λ from the potential matrix elements.

To verify the quality of the chiral potentials generated by the Glow model, in Fig. 4 we show the calculated phase shifts in selected partial-wave channels. From Fig. 4 we observe that the Glow-n3lo Λ potentials give good phase shift results when compared to experimental data used to fit the actual chiral potentials. The ability of the Glow model to extrapolate outside the Λ training range is illustrated by the phase shifts for the Glow-generated potential at $\Lambda = 435$ MeV. Furthermore, even though the values of Λ are uniformly sampled over the range $450 < \Lambda < 550$, the phase shifts need not be uniformly distributed.

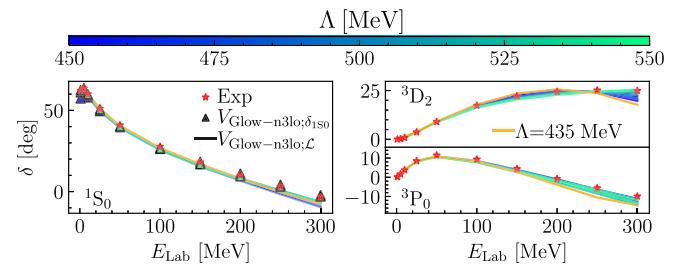


FIG. 4. Phase shifts for np scattering. The colored solid lines correspond to Glow-n3lo Λ potentials. Red stars correspond to the phase shift analysis [48] of NN scattering data.

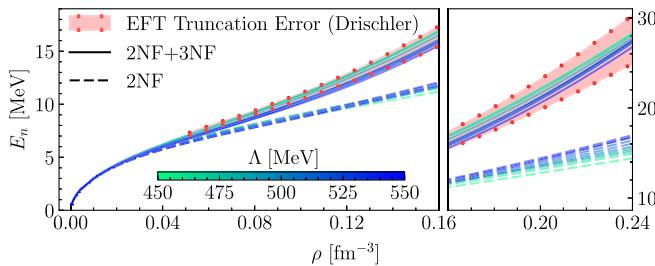


FIG. 5. Pure neutron matter equation of state with different cutoffs. The EFT truncation error (red region) from Ref. [49] is plotted for comparison.

Finally, as a test case, we study the influence of the resolution scale on the zero-temperature neutron matter equation of state calculated up to second order in perturbation theory, shown in Fig. 5 for two-body forces alone and including the n2lo chiral three-body force. Below saturation density, the uncertainties due to the choice of cutoff scale are larger than those arising from the truncation in the EFT expansion [49] shown as the red band. However, beyond nuclear saturation density the EFT expansion parameter increases and one observes the EFT truncation errors to grow stronger than those due to the resolution scale. Again, we can find that the distribution of the energy is not a simple uniform distribution with respect to the cutoff.

Summary and outlook—In this Letter, we have extended the Glow machine learning model to generate novel instances of the nucleon-nucleon interaction by training the neural network on existing interactions in the literature. We have shown that the Glow model can accurately reconstruct the training nuclear potentials and build a continuous distribution of potentials over a range of resolution scales, all while reproducing nucleon-nucleon scattering phase shift data. The Glow model enables the generation of realistic nucleon-nucleon potentials in a matter of seconds, and therefore it can play an important role for more reliable estimations of nuclear many-body uncertainties that arise due to the arbitrary choice of resolution scale in the nucleon-nucleon interaction. The treatment of nuclear three-body forces within the Glow model is expected to be a straightforward extension of the methods developed in this study and will be pursued in future work. These tools add to the growing body of recent literature exploring machine learning models for uncertainty quantification and the renormalization group.

Acknowledgments—We thank Takayuki Miyagi for providing us with codes for generating SRG NN potentials. Work supported by the National Science Foundation under Grants No. PHY1652199 and No. PHY2209318. Portions of this research were conducted with the advanced

computing resources provided by Texas A&M High Performance Research Computing.

Data availability—Source code for this project can be found at <https://github.com/pswen2019/Glow-nuclear-potential.git>.

- [1] K. Hebeler, J. D. Holt, J. Menendez, and A. Schwenk, *Annu. Rev. Nucl. Part. Sci.* **65**, 457 (2015).
- [2] A. Ekström, C. Forssén, G. Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock, *Front. Phys.* **11**, 1129094 (2023).
- [3] K. Hebeler, J. M. Lattimer, C. J. Pethick, and A. Schwenk, *Phys. Rev. Lett.* **105**, 161102 (2010).
- [4] C. Drischler, J. W. Holt, and C. Wellenhofer, *Annu. Rev. Nucl. Part. Sci.* **71**, 403 (2021).
- [5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, *arXiv: 1406.2661*.
- [6] D. P. Kingma and M. Welling, *arXiv:1312.6114*.
- [7] J. Ho, A. Jain, and P. Abbeel, *arXiv:2006.11239*.
- [8] J. Kim, S. Kim, J. Kong, and S. Yoon, *arXiv:2005.11129*.
- [9] H.-Y. Hu, S.-H. Li, L. Wang, and Y.-Z. You, *Phys. Rev. Res.* **2**, 023369 (2020).
- [10] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, *Inf. Fusion* **76**, 243 (2021).
- [11] M. Yang, P. Wang, M. Fan, D. Lu, Y. Cao, and G. Zhang, *arXiv:2404.00502*.
- [12] V. Böhm, F. Lanusse, and U. Seljak, *arXiv:1910.10046*.
- [13] S. K. Bogner, T. T. S. Kuo, and A. Schwenk, *Phys. Rep.* **386**, 1 (2003).
- [14] S. K. Bogner, R. J. Furnstahl, and A. Schwenk, *Prog. Part. Nucl. Phys.* **65**, 94 (2010).
- [15] J. E. Lynn, J. Carlson, E. Epelbaum, S. Gandolfi, A. Gezerlis, and A. Schwenk, *Phys. Rev. Lett.* **113**, 192501 (2014).
- [16] J. W. Holt and N. Kaiser, *Phys. Rev. C* **95**, 034326 (2017).
- [17] A. M. Gasparyan and E. Epelbaum, *Phys. Rev. C* **107**, 034001 (2023).
- [18] S. Weinberg, *Physica (Amsterdam)* **96A**, 327 (1979).
- [19] E. Epelbaum, H.-W. Hammer, and Ulf.-G. Meißner, *Rev. Mod. Phys.* **81**, 1773 (2009).
- [20] R. Machleidt and D. R. Entem, *Phys. Rep.* **503**, 1.
- [21] E. Epelbaum, H. Krebs, and U. G. Meißner, *Eur. Phys. J. A* **51**, 53 (2015).
- [22] R. J. Furnstahl, D. R. Phillips, and S. Wesolowski, *J. Phys. G* **42**, 034028 (2015).
- [23] R. J. Furnstahl, N. Klco, D. R. Phillips, and S. Wesolowski, *Phys. Rev. C* **92**, 024005 (2015).
- [24] S. Wesolowski, N. Klco, R. J. Furnstahl, D. R. Phillips, and A. Thapaliya, *J. Phys. G* **43**, 074001 (2016).
- [25] B. D. Carlsson, A. Ekström, C. Forssén, D. F. Strömberg, G. R. Jansen, O. Lilja, M. Lindby, B. A. Mattsson, and K. A. Wendt, *Phys. Rev. X* **6**, 011019 (2016).
- [26] C. Drischler, K. Hebeler, and A. Schwenk, *Phys. Rev. Lett.* **122**, 042501 (2019).
- [27] S. Wesolowski, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips, *J. Phys. G* **46**, 045102 (2019).

[28] C. Drischler, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips, *Phys. Rev. Lett.* **125**, 202702 (2020).

[29] C. Drischler, J. A. Melendez, R. J. Furnstahl, and D. R. Phillips, *Phys. Rev. C* **102**, 054315 (2020).

[30] C. Drischler, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips, *Phys. Rev. Lett.* **125**, 202702 (2020).

[31] J. A. Melendez, S. Wesolowski, and R. J. Furnstahl, *Phys. Rev. C* **96**, 024003 (2017).

[32] D. R. Entem, N. Kaiser, R. Machleidt, and Y. Nosyk, *Phys. Rev. C* **91**, 014002 (2015).

[33] I. Svensson, A. Ekström, and C. Forssén, *Phys. Rev. C* **107**, 014001 (2023).

[34] I. Svensson, A. Ekström, and C. Forssén, *Phys. Rev. C* **109**, 064003 (2024).

[35] D. P. Kingma and P. Dhariwal, [arXiv:1807.03039](https://arxiv.org/abs/1807.03039).

[36] L. Dinh, J. Sohl-Dickstein, and S. Bengio, [arXiv:1605.0880](https://arxiv.org/abs/1605.0880).

[37] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, [arXiv:2010.11929](https://arxiv.org/abs/2010.11929).

[38] C. Winkler, D. Worrall, E. Hoogeboom, and M. Welling, [arXiv:1912.00042](https://arxiv.org/abs/1912.00042).

[39] L. Dinh, D. Krueger, and Y. Bengio, [arXiv:1410.8516](https://arxiv.org/abs/1410.8516).

[40] C. Gao, J. Isaacson, and C. Krause, *Mach. Learn. Sci. Technol.* **1**, 045023 (2020).

[41] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, [arXiv:1906.04032](https://arxiv.org/abs/1906.04032).

[42] S. K. Bogner, R. J. Furnstahl, and R. J. Perry, *Phys. Rev. C* **75**, 061001(R) (2007).

[43] E. D. Jurgenson, P. Navrátil, and R. J. Furnstahl, *Phys. Rev. Lett.* **103**, 082501.

[44] D. R. Entem and R. Machleidt, *Phys. Rev. C* **68**, 041001(R) (2003).

[45] D. R. Entem, R. Machleidt, and Y. Nosyk, *Phys. Rev. C* **96**, 024004 (2017).

[46] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, [arXiv:1706.03762](https://arxiv.org/abs/1706.03762).

[47] S. K. Bogner, R. J. Furnstahl, and A. Schwenk, *Prog. Part. Nucl. Phys.* **65**, 94 (2010).

[48] R. Navarro Pérez, J. E. Amaro, and E. Ruiz Arriola, *Phys. Rev. C* **89**, 064006 (2014).

[49] C. Drischler, J. A. Melendez, R. J. Furnstahl, and D. R. Phillips, *Phys. Rev. C* **102**, 054315 (2020).