Journal Club (20250319)

Generative Modeling of Nucleon-Nucleon Interactions

Jiangming Yao (尧江明)

School of Physics and Astronomy, Sun Yat-Sen University, China

Reference: P. S. Wen, J. W. Holt, and Maggie Li, Generative Modeling of Nucleon-Nucleon Interactions, PRL133, 252501 (2024)

Outline

PHYSICAL REVIEW LETTERS **133,** 252501 (2024)

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.

High-precision models of nuclear interaction are essential:

- nuclear structure and dynamics
- hot and dense matter in extreme astrophusial envitoments (neutron stars, core collapse)

"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. "

Evolution of NN interaction

- Nuclear potential arises from quantum chromodynamics (QCD) 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.
- The resolution scale is typically parametrized in the nuclear potential through a momentum-space regulating function parameter Λ that demarcates (界定) the separation between low-energy and high-energy (unresolved) physics.
- 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.
- 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.

B. D. Carlsson et al., Phys. Rev. X 6, 011019 (2016)

This work: Application of Glow model for NN potentials

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 nucleon-nucleon (NN) potential from a continuous distribution in Λ, which when combined with EFT truncation errors and variations in the LECs consistent with data can provide a comprehensive assessment of theory uncertainties.

- We utilize the Generative Flow (Glow) model, 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.

arXiv > stat > arXiv:1807.03039

Search.
Help | A

Statistics > Machine Learning

[Submitted on 9 Jul 2018 (v1), last revised 10 Jul 2018 (this version, v2)]

Glow: Generative Flow with Invertible 1x1 Convolutions

Diederik P. Kingma, Prafulla Dhariwal

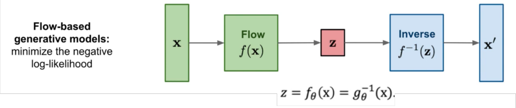
Flow-based generative models (Dinh et al., 2014) are conceptually attractive due to tractability of the exact log-likelihood, tractability of exact latent-variable inference, and parallelizability of both training and synthesis. In this paper we propose Glow, a simple type of generative flow using an invertible 1x1 convolution. Using our method we demonstrate a significant improvement in log-likelihood on standard benchmarks. Perhaps most strikingly, we demonstrate that a generative model optimized towards the plain log-likelihood objective is capable of efficient realistic-looking synthesis and manipulation of large images. The code for our model is available at this https URL

Comments: 15 pages; fixed typo in abstract
Subjects: **Machine Learning (stat.ML)**; Artificial Intelligence (cs.AI); Machine Learning (cs.LG)
Cite as: arXiv:1807.03039 **[stat.ML]**
(or arXiv:1807.03039v2 **[stat.ML]** for this version)
https://doi.org/10.48550/arXiv.1807.03039

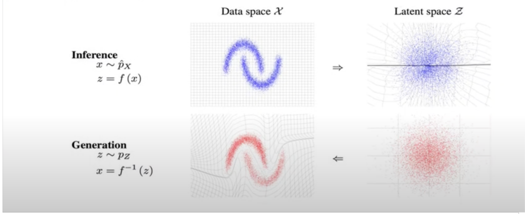
This work: Application of Glow model for NN potentials

- 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 (a reversible flow-based generative model) and Vision Transformer model allows for the extraction of chiral EFT LECs from the generated nuclear potential matrix elements.

The Glow model: basic idea



https://arxiv.org/pdf/1605.08803.pdf



An $i.i.d.$ dataset $D = \{\mathrm{x}^{(1)}, \ldots, \mathrm{x}^{(N)}\}$ with random vector x is described by a model $p_\theta(\mathrm{x})$ with parameters θ. The objective is to minimize the average negative log-likelihood

$$\mathcal{L}(D) = \frac{1}{N}\sum_{i=1}^{N} -log p_\theta(\mathrm{x}^{(i)})$$

Entropy in a random variable $\mathrm{x}^{(i)}$: $\sum log \frac{1}{p_\theta(\mathrm{x}^{(i)})}$ Large probability $p_\theta(\mathrm{x}^{(i)})$ has a low entropy in $\mathrm{x}^{(i)}$.

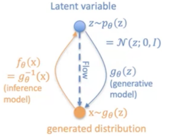
which measures the expectation cost. The optimization is performed by using stochastic gradient descent (SGD) using minibatches of data.

Flow-based generative models define a generative process

$z \sim p_\theta(\mathrm{z})$

$\mathrm{x} \sim g_\theta(\mathrm{z})$

where z is the latent variable and $p_\theta(\mathrm{z})$ has a tractable density, e.g., a Gaussian distribution: $p_\theta(\mathrm{z}) = \mathcal{N}(z; 0, I)$. The mapping function $g_\theta(\mathrm{z})$ is invertible in flow-based generative models, i.e., $z = f_\theta(\mathrm{x}) = g_\theta^{-1}(\mathrm{x})$.



$$p_\theta(\mathrm{x}) = p_\theta(\mathrm{z}) \left|\det\left(\frac{dz}{d\mathrm{x}}\right)\right|$$

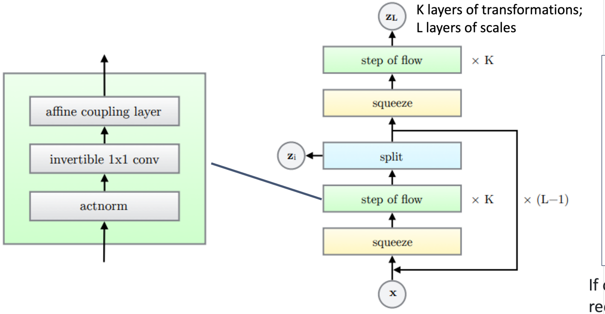
- Starting from a trainable probability distribution function with parameters $\boldsymbol{\theta}$ to estimate the probabilities of some features of a sample.
- Adjusting iteratively the distribution such that the likelihood $p_{\boldsymbol{\theta}}(\boldsymbol{x})$ of the distribution can be maximized with respect to $\boldsymbol{\theta}$.
- Once the probability distribution is found, new samples can be generated by the Glow model from it. That's why it belongs to generative mode.

The Glow model: transformation function

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

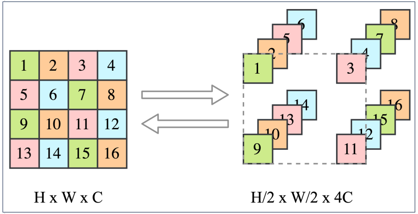
$$\boldsymbol{z} = f_K \circ f_{K-1} \circ \cdots \circ f_k \circ \cdots \circ f_1(\boldsymbol{x}),$$

where f_k is the transformation function at the k-th layer.



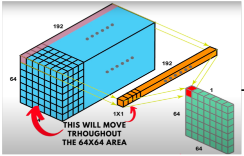
(a) One step of our flow. (b) Multi-scale architecture (Dinh et al., 2016).

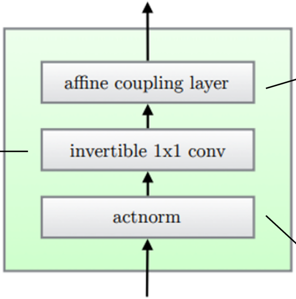
Squeeze and Split



If one wants to remove half of the pixels in an image, one can reduce the height (H) and width (W) of the image by a factor of 2 while scaling the number of channels (C) by 4.

The Glow model: three parts within a layer of transformation





(a) One step of our flow.

Affine Coupling is a method for implementing a normalizing flow (where we stack a sequence of invertible bijective transformation functions). Affine coupling is one of these bijective transformation functions. Specifically, it is an example of a reversible transformation where the forward function, the reverse function and the log-determinant are computationally efficient. For the forward function, we split the input dimension into two parts:

$$\mathbf{x}_a, \mathbf{x}_b = \text{split}(\mathbf{x})$$

The second part stays the same $\mathbf{x}_b = \mathbf{y}_b$, while the first part $\mathbf{x}_a$ undergoes an affine transformation, where the parameters for this transformation are learnt using the second part $\mathbf{x}_b$ being put through a neural network. Together we have:

$$(\log \mathbf{s}, \mathbf{t}) = \text{NN}(\mathbf{x}_b)$$

$$\mathbf{s} = \exp(\log \mathbf{s})$$

$$\mathbf{y}_a = \mathbf{s} \odot \mathbf{x}_a + \mathbf{t}$$

$$\mathbf{y}_b = \mathbf{x}_b$$

$$\mathbf{y} = \text{concat}(\mathbf{y}_a, \mathbf{y}_b)$$

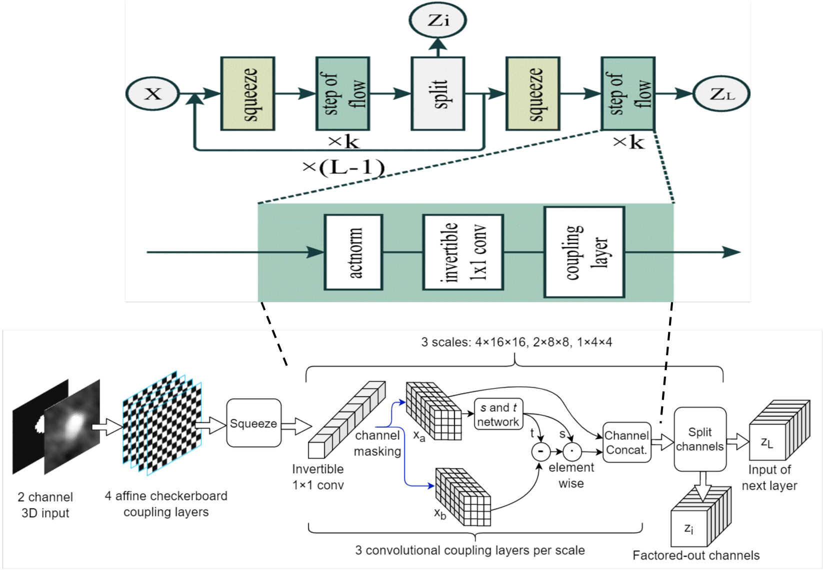
Normalization layer similar to batch norm, except that the mean and standard deviation statistics are trainable parameters rather than estimated from the data

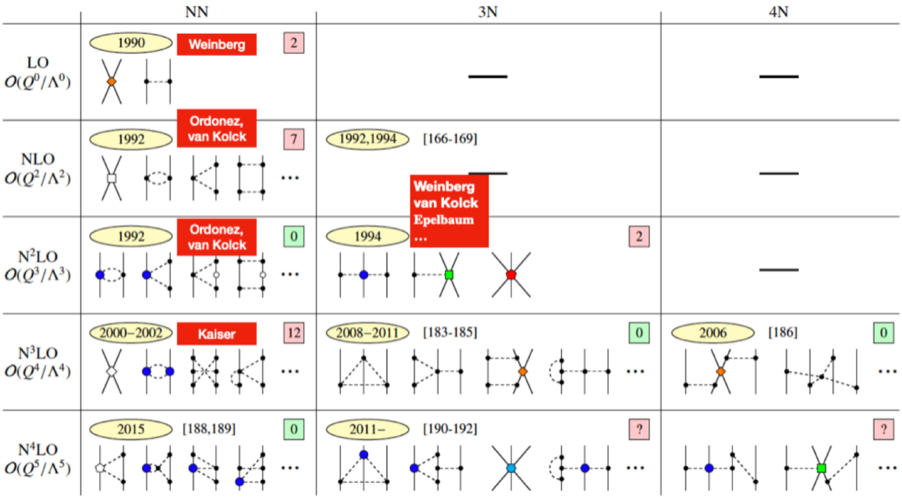
Forward $y = x * \sigma + \mu$

Backward $x = (y - \mu)/\sigma$

Log-det: $\log\det \text{ActNorm} = h \times w \times \sum \log(|\sigma|)$

The Glow model: the entire flow

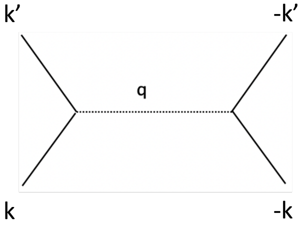




K. Hebeler, Phys. Rep. 890, 1 (2020)

The NN potential V_{NN} in momentum space

$$\langle \boldsymbol{k}_1', \boldsymbol{k}_2' | V(1,2) | \boldsymbol{k}_1, \boldsymbol{k}_2 \rangle = V^{(2)}(\boldsymbol{k}', \boldsymbol{k}) \delta(\mathbf{K} - \mathbf{K}')$$



The NN potential is usually multiplied with a regulator function

$$V^{(2)}(\boldsymbol{k}', \boldsymbol{k}) \to f^{\Lambda}_{NN}(k') V^{(2)}(\boldsymbol{k}', \boldsymbol{k}) f^{\Lambda}_{NN}(k)$$

where $f^{\Lambda}_{NN} = \exp[-(k/\Lambda)^n]$, with $n = 6$.

The UV momentum-space cutoff is here chosen as $\Lambda = 450, 500, 550$ MeV, respectively.

$$\begin{aligned} V^{\rm LO}(k', k) &= C_c + C_\tau \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 + C_\sigma \vec{\sigma}_1 \cdot \vec{\sigma}_2 \\ &\quad + C_{\sigma\tau} \vec{\sigma}_1 \cdot \vec{\sigma}_2 \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 \\ &\quad - \frac{g_A^2}{4 f_\pi^2} \frac{1}{q^2 + m_\pi^2} \vec{\sigma}_1 \cdot \vec{\sigma}_2 \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 \end{aligned}$$

where the momentum transfer $q = |\boldsymbol{k} - \boldsymbol{k}'|$.

Partial wave expansion for the two-body state

$$|\boldsymbol{k}\rangle = \sum_{l,m_l} Y^*_{lm_l}(\hat{k})|klm_l\rangle$$

In the $(lS)J$-coupled basis

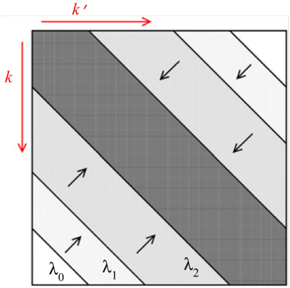
$$\langle k'(l'S)J|\, V_{NN}\, |k(lS)J\rangle = \frac{2}{\pi}\int drr^2 j_{l'}(k'r)j_l(kr)\langle(l'S')J|V_{NN}(r)|(lS)J\rangle.$$

- Apply unitary transformations to decouple high and low-momentum states

$$H_s = U_s H U_s^\dagger \equiv T_{\rm rel} + V_s$$

from which one finds the flow equation

$$\frac{dH_s}{ds} = [\eta_s, H_s], \quad \eta_s = [T_{\rm rel}, H_s]$$



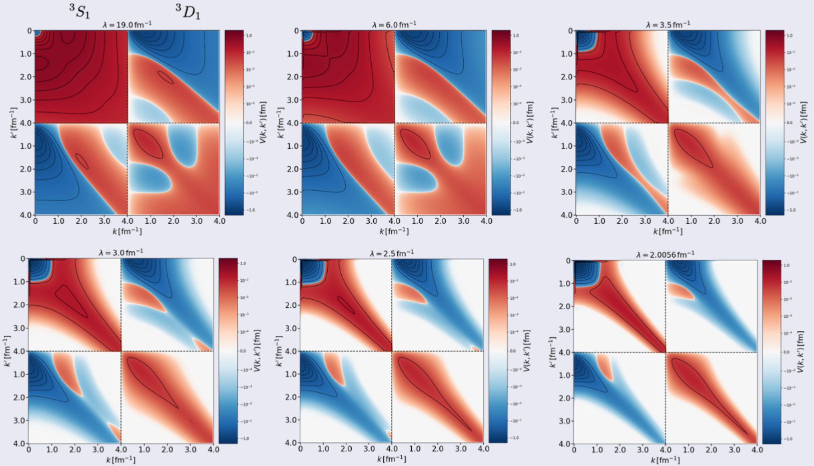
The flow parameter s is usually replaced with $\lambda = s^{-1/4}$ in units of fm^{-1} (a measure of the spread of off-diagonal strength).

Evolution of the potential

$$\frac{dV_s(k,k')}{ds} = -(k^2 - k'^2)^2 V_s(k,k') + \frac{2}{\pi}\int_0^\infty q^2 dq (k^2 + k'^2 - 2q^2) V_s(k,q) V_s(q,k')$$

The SRG evolution of NN potential

The *np* potential of AV18 with different momentum resolutions

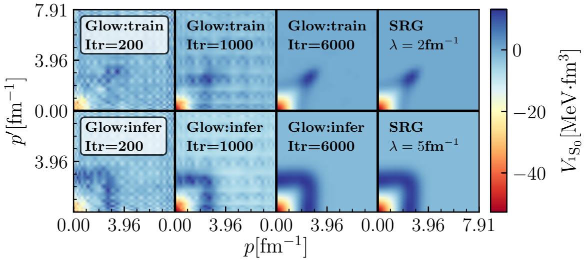


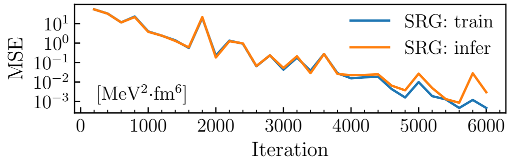
The NN potentials in the Glow model and SRG

We conduct two experiments to investigate the Glow model's capabilities in reconstructing actual nuclear potentials, building distributions, and generating new, realistic ones.

1. N3LO chiral nuclear potential of Entem and Machleidt with momentum-space cutoff $\Lambda = 500$ MeV, evolved with the Similarity Renormalization Group (SRG) to $\lambda \in \{2, 3, 4, 7, 8, 10, 11, 12\}\,\mathrm{fm}^{-1}$. The SRG potential experiments include three partial-wave channels $C = 3$, and the size of the momentum-space mesh grid is $H = W = 32$, and $(L, K) = (3, 2)$.
2. A set of six chiral potentials with chiral expansion orders $\nu \in \{2, 3\}$ and cutoffs $\Lambda \in \{450, 500, 550\}\,\mathrm{MeV}$. We include all partial waves that have associated short-distance contact terms, which requires $C = 14$, and we set $H = W = 48$, and $(L, K) = (4, 4)$.

The NN potential in the Glow model and SRG





- (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.

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

The prediction of LECs and cutoffs Λ with the ViT model

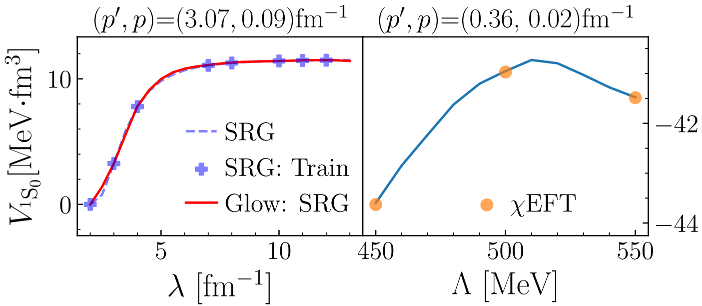
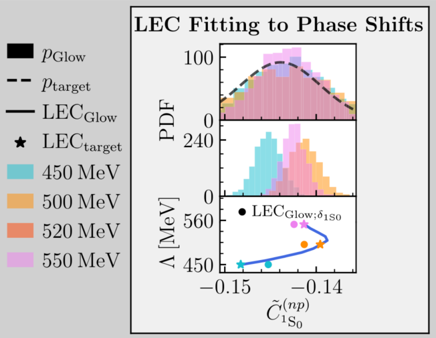


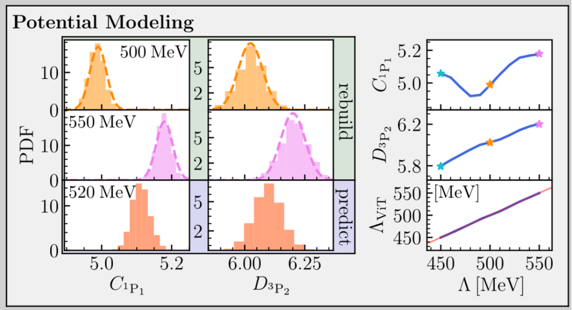
Figure: SRG (left) and chiral (right) off-diagonal potential matrix elements in the $^1S_0$ partial-wave channel. The plus and dot markers as well as the dashed lines are from the exact potentials, while solid lines are generated by trained Glow models using LSI (for SRG) and LSS (for chiral potentials),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

LSS: latent space sampling; LSI: latent space interpolation.

The prediction of LECs and cutoffs Λ with the ViT model



- 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.
- Taining a Glow model at three different cutoffs $\Lambda = \{450, 500, 550\,(\text{magenta})\}$ MeV using a wide distribution (top panel, dashed line) for the LEC $\tilde{C}^{(np)}_{^1S_0}$. 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.
- After the phase shift training, the distributions for the $C^{(np)}_{^1S_0}$ LEC extracted by the ViT model converge separately for the three values of Λ.



- The Glow model has the ability to rebuild the LEC distribution at the training values of $\Lambda = 500, 550$ MeV (top two rows), and to predicte distribution $p(\mathrm{LECs}|\Lambda = 520\,\mathrm{MeV})$.
- The continuous evolution of the peak in the LEC distributions between $\Lambda = 450 - 550\,\mathrm{MeV}$ 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.

The predicted phase shifts

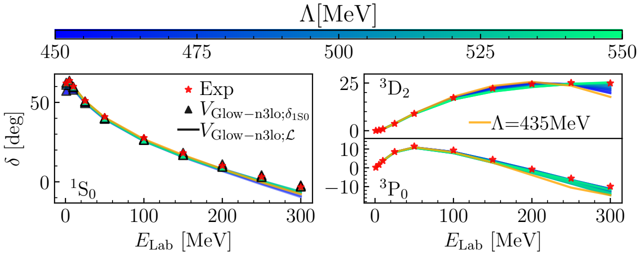


Figure: Phase shifts for *np* scattering. The colored solid lines correspond to Glow-n3loΛ potentials. Red stars correspond to the phase shift analysis of NN scattering data.

- The Glow-n3loΛ 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 $\Lambda = 435\,\mathrm{MeV}$.

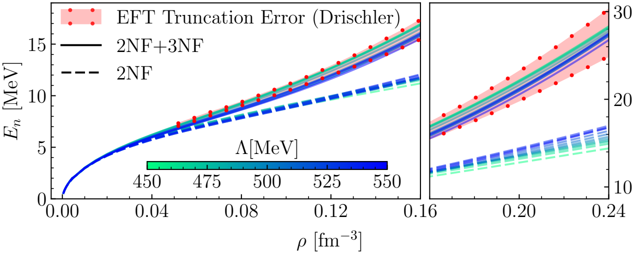


Figure: Pure neutron matter equation of state with different cutoffs. The EFT truncation error (red region) from (Drischler 2020) is plotted for comparison.

- Below saturation density, the uncertainties due to the choice of cutoff scale are larger than those arising from the truncation in the EFT expansion shown as the red band.
- Beyond nuclear saturation density, the EFT expansion parameter increases, and the truncation errors to grow stronger than those due to the resolution scale.

Summary and Outlook

- In this work, 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.