Machine Learning for *Ab Initio* Nuclear Structure Calculations

Tobias Wolfgruber

Progress in *Ab Initio* Nuclear Theory
TRIUMF 2023

No-Core Shell Model

- stationary Schrödinger equation as matrix eigenvalue problem

- Slater determinants $|\phi_i\rangle$ constructed from HO basis
  $\rightarrow$ dependency on HO frequency $\hbar\Omega$

- truncate model space by number of excitation quanta N_{max}
  w.r.t. the lowest-energy Slater determinant

No-Core Shell Model

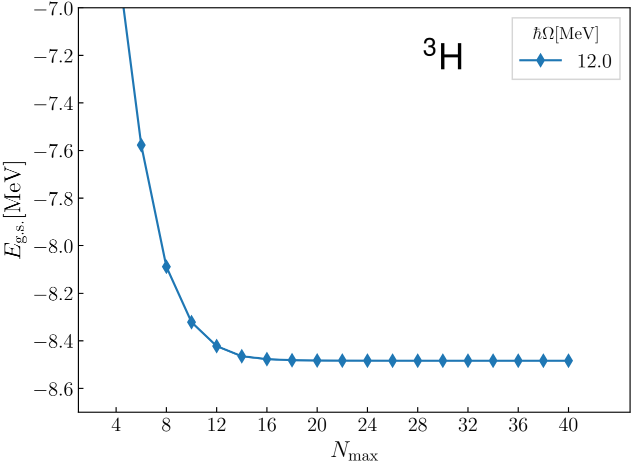
- stationary Schrödinger equation as matrix eigenvalue problem

- Slater determinants $|\phi_i\rangle$ constructed from HO basis
  $\rightarrow$ dependency on HO frequency $\boxed{\hbar\Omega}$

- truncate model space by number of excitation quanta $\boxed{N_{\text{max}}}$
  w.r.t. the lowest-energy Slater determinant

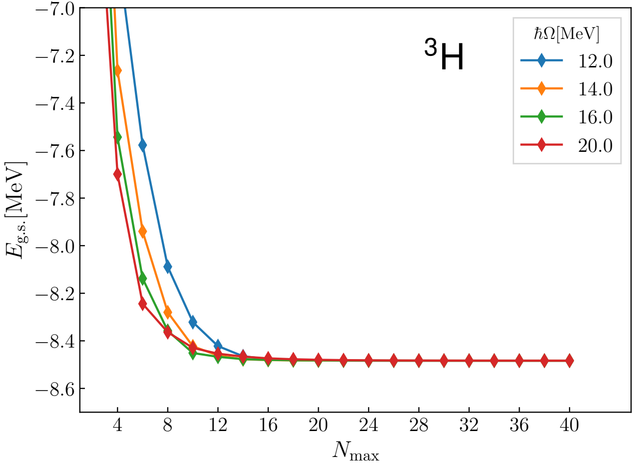
- **convergence controlled by two parameters**

Convergence Behavior



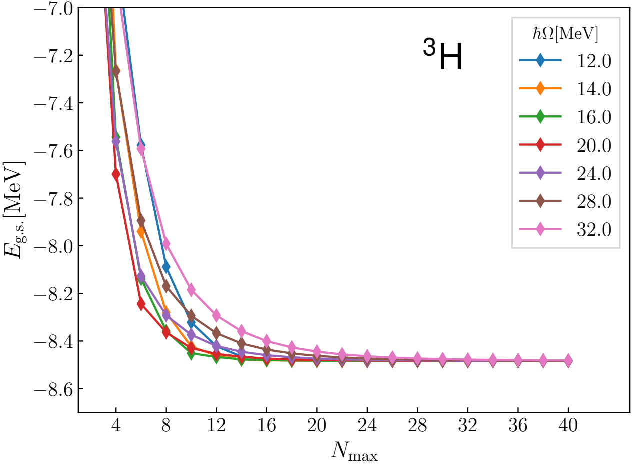
- monotonously decreasing with N_{max}

Convergence Behavior



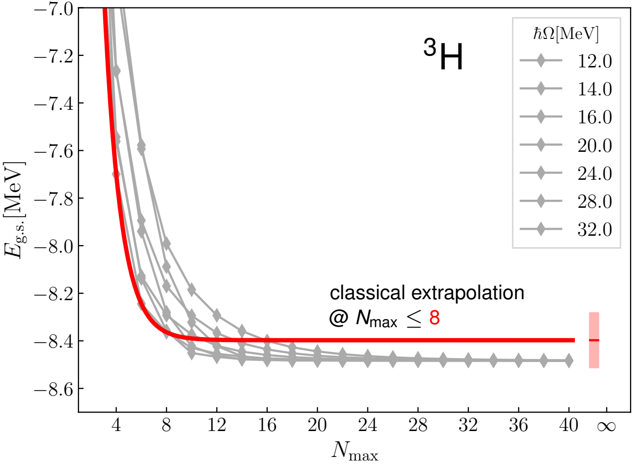
- monotonously decreasing with N_{max}
- different rates of convergence for different HO frequencies

Convergence Behavior



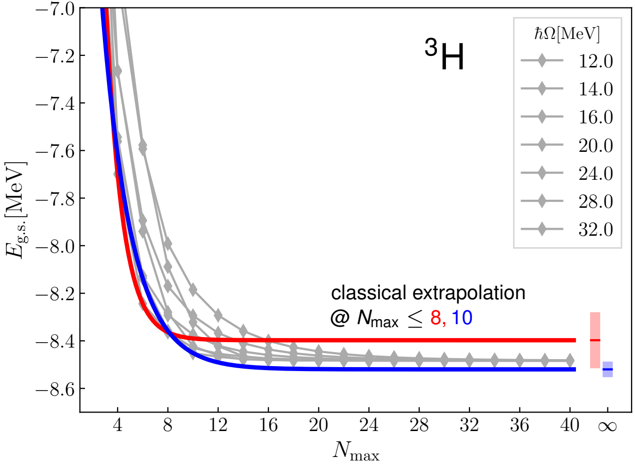
- monotonously decreasing with N_{max}
- different rates of convergence for different HO frequencies
- all sequences share the same limit

Convergence Behavior



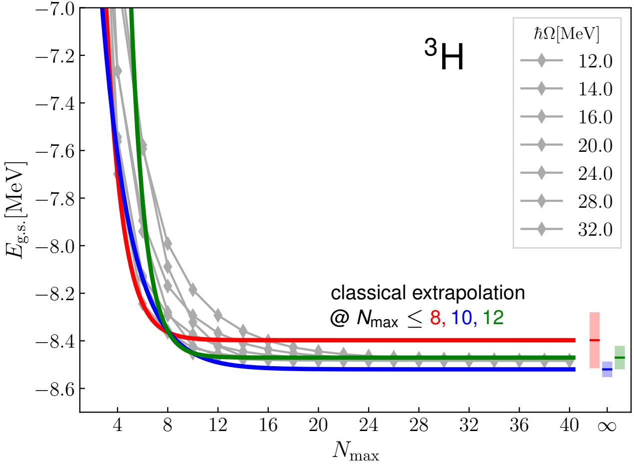
- monotonously decreasing with N_max
- different rates of convergence for different HO frequencies
- all sequences share the same limit

Convergence Behavior



- monotonously decreasing with $N_{\max}$
- different rates of convergence for different HO frequencies
- all sequences share the same limit

Convergence Behavior



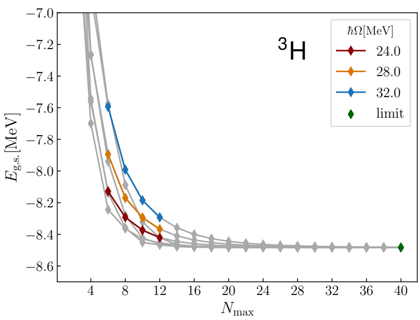
- monotonously decreasing with $N_{\max}$
- different rates of convergence for different HO frequencies
- all sequences share the same limit

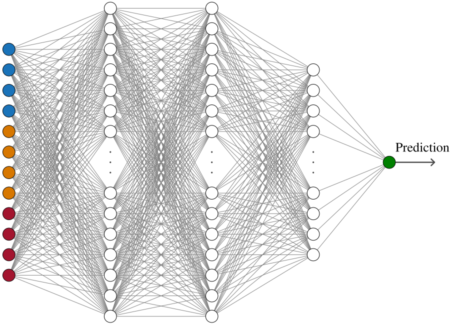
Machine Learning Approach

- previous applications: capture $f(N_{\text{max}}, \hbar\Omega)$

Negoita et al. PR C 99, 054308 (2019)
Jiang et al. PR C 100, 054326 (2019)

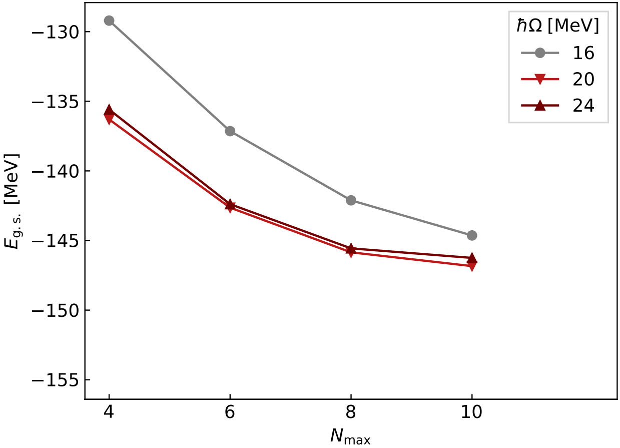
- now: directly predict converged value from available calculations
  - include information of multiple frequencies





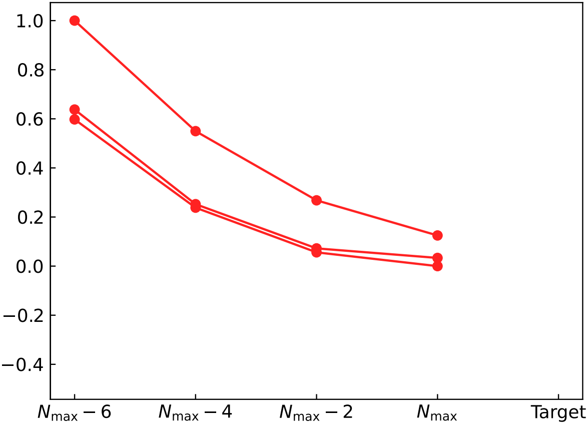
Knöll, TW et al. PLB 839, 137781 (2023)

^{16}O Prediction



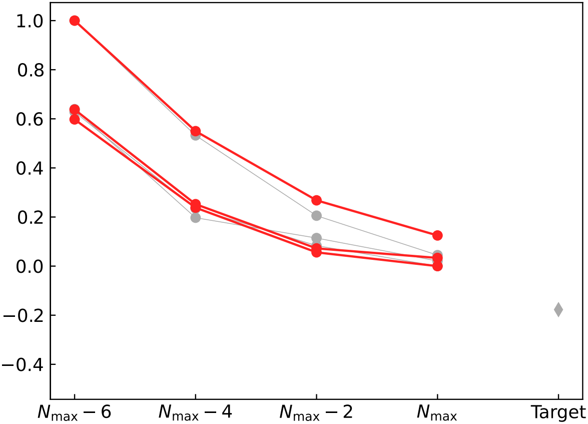
- ^{16}O sequence

^{16}O Prediction



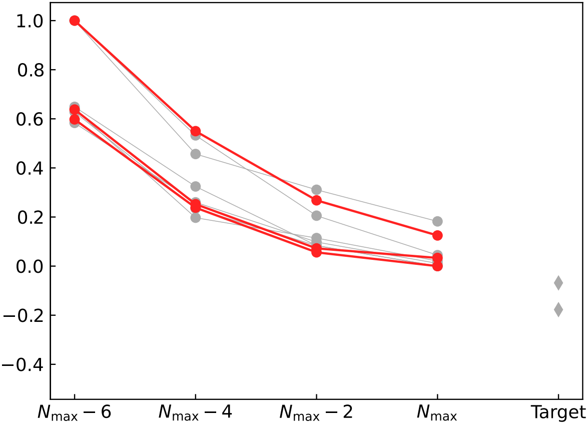
- ^{16}O ~~sequence~~ sample
- no $N_{\max}$ or $\hbar\Omega$ information
- New: MINMAX norm $\rightarrow$ rescale to [0, 1]

^{16}O Prediction



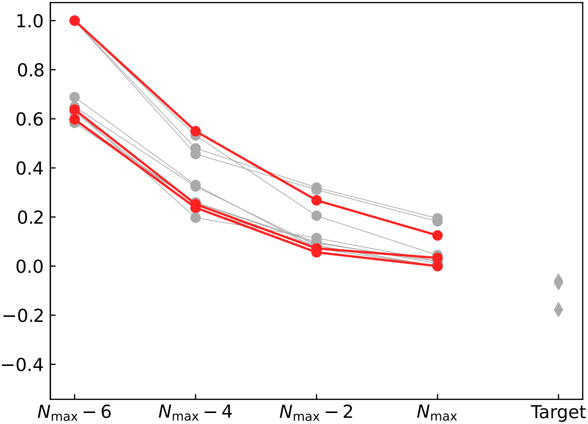
- ^{16}O ~~sequence~~ sample
- no $N_{\max}$ or $\hbar\Omega$ information
- New: MINMAX norm $\rightarrow$ rescale to [0, 1]
- very similar to ^{2}H training samples

^{16}O Prediction



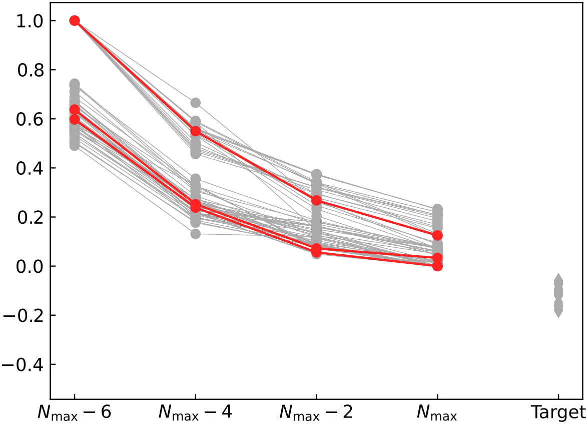
- ^{16}O ~~sequence~~ sample
- no N_{max} or $\hbar\Omega$ information
- New: MINMAX norm → rescale to [0, 1]
- very similar to ^{2}H training samples

^{16}O Prediction



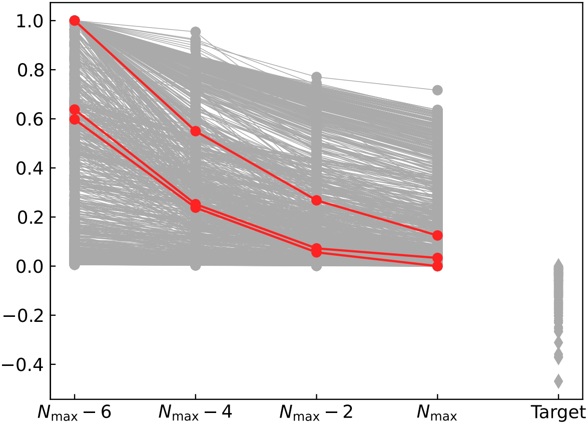
- ^{16}O ~~sequence~~ sample
- no N_{max} or $\hbar\Omega$ information
- New: MINMAX norm → rescale to [0, 1]
- very similar to ^{2}H training samples

^{16}O Prediction



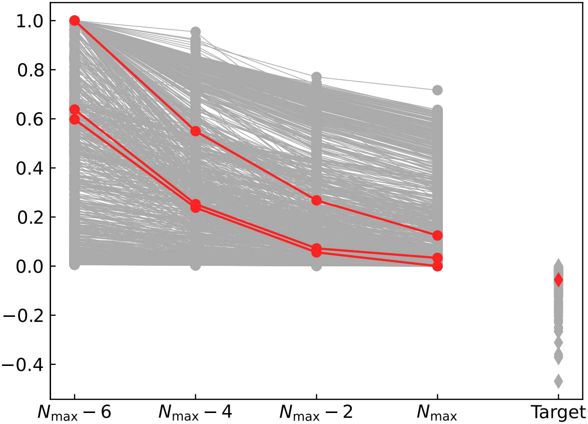
- ^{16}O ~~sequence~~ sample
- no $N_{\max}$ or $\hbar\Omega$ information
- New: MINMAX norm $\rightarrow$ rescale to [0, 1]
- very similar to ^{2}H training samples

^{16}O Prediction



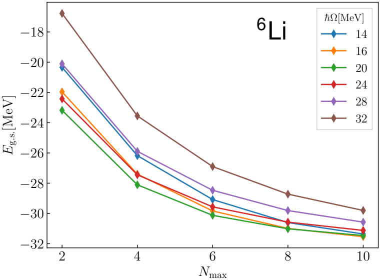
- ^{16}O ~~sequence~~ sample
- no N_{max} or $\hbar\Omega$ information
- New: MINMAX norm $\rightarrow$ rescale to [0, 1]
- very similar to ^{2}H training samples
- total of 340 200 samples across ^{2}H, ^{3}H, and ^{4}He

^{16}O Prediction



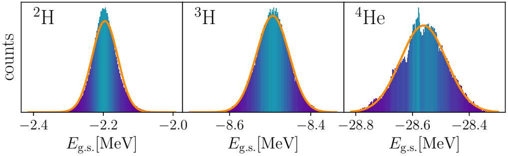
- ^{16}O ~~sequence~~ sample
- no $N_{\max}$ or $\hbar\Omega$ information
- New: MINMAX norm → rescale to [0, 1]
- very similar to ^{2}H training samples
- total of 340 200 samples across ^{2}H, ^{3}H, and ^{4}He

Statistical Evaluation

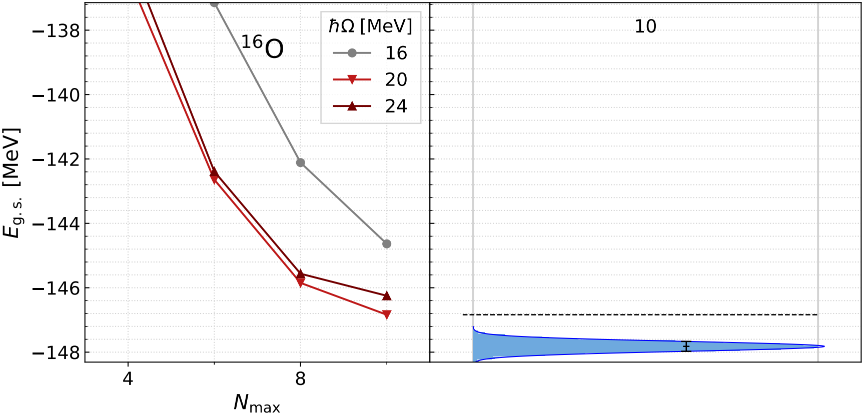


- different family of interactions
  Maris et al. PR C 103, 054001 (2021)
- construction of evaluation samples analogoulsy to training samples
- different predictions from one ANN
- turn to statistical approach

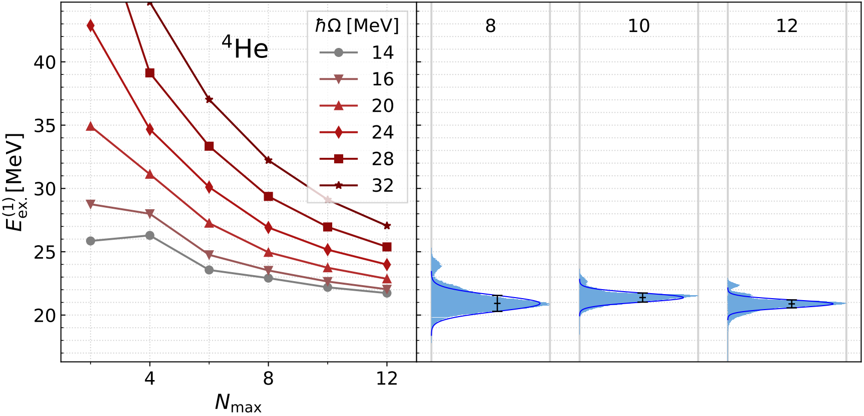
- apply 1000 ANN
- prediction and uncertainty from Gaussian fit



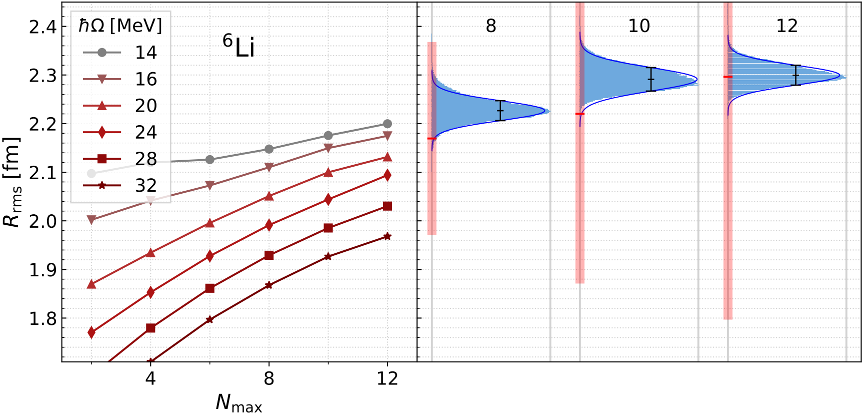
^{16}O Ground-State Energy



^{4}He Excitation Energy



^{6}Li Radius


ħΩ [MeV]
14
16
20
24
28
32
^{6}Li
8
10
12
2.4
2.3
2.2
2.1
2.0
1.9
1.8
R_{rms} [fm]
4
8
12
$N_{\max}$

Thank you for your attention!

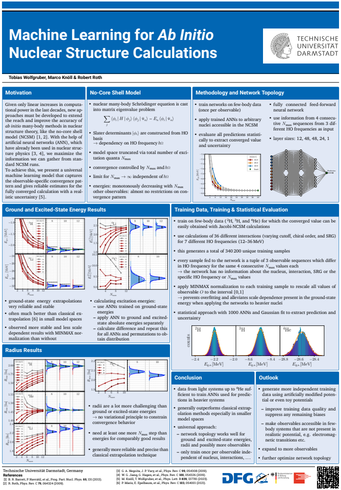
- **thanks to my group and collaborators**

  P. Falk, K. Katzenmeier, **M. Knöll**, L. Mertes, J. Müller, **R. Roth**, L. Wagner, C. Wenz

Bundesministerium für Bildung und Forschung

Hessisches Kompetenzzentrum für Hochleistungsrechnen

SFB 1245



Machine Learning for *Ab Initio* Nuclear Structure Calculations

Tobias Wolfgruber, Marco Knöll & Robert Roth

Motivation

Given only linear increases in computational power in the last decades, new approaches must be developed to extend the reach and improve the accuracy of *ab initio* many-body methods in nuclear structure theory, like the no-core shell model (NCSM) [1, 2]. With the help of artificial neural networks (ANN), which have already been used in nuclear structure physics [3, 4], we maximize the information we can gather from standard NCSM runs.

To achieve this, we present a universal machine learning model that captures the observable-specific convergence pattern and gives reliable estimates for the fully converged calculation with a realistic uncertainty [5].

No-Core Shell Model

- nuclear many-body Schrödinger equation is cast into matrix eigenvalue problem

$$\sum_j \langle\phi_i|H|\phi_j\rangle\langle\phi_j|\Psi_n\rangle = E_n\langle\phi_i|\Psi_n\rangle$$

- Slater determinants $|\phi_i\rangle$ are constructed from HO basis
  → dependency on HO frequency $\hbar\Omega$
- model space truncated via total number of excitation quanta $N_{\max}$
- convergence controlled by $N_{\max}$ and $\hbar\Omega$
- limit for $N_{\max} \to \infty$ independent of $\hbar\Omega$
- energies: monotonously decreasing with $N_{\max}$
  other observables: almost no restrictions on convergence pattern

Methodology and Network Topology

- train networks on few-body data (once per observable)
- apply trained ANNs to arbitrary nuclei accessible in the NCSM
- evaluate all predictions statistically to extract converged value and uncertainty
- fully connected feed-forward neural network
- use information from 4 consecutive $N_{\max}$ sequences from 3 different HO frequencies as input
- layer sizes: 12, 48, 48, 24, 1

Ground and Excited-State Energy Results

- ground-state energy extrapolations very reliable and stable
- often much better than classical extrapolation [6] in small model spaces
- observed more stable and less scale dependent results with MINMAX normalization than without
- calculating excitation energies:
  - use ANNs trained on ground-state energies
  - apply ANN to ground and excited-state absolute energies separately
  - calculate difference and repeat this for all ANNs and permutations to obtain distribution

Radius Results

- radii are a lot more challenging than ground or excited-state energies
  → no variational principle to constrain convergence behavior
- need at least one more $N_{\max}$ step than energies for comparably good results
- generally more reliable and precise than classical extrapolation technique

Training Data, Training & Statistical Evaluation

- train on few-body data (^{2}H, ^{3}H, and ^{4}He) for which the converged value can be easily obtained with Jacobi-NCSM calculations
- use calculations of 36 different interactions (varying cutoff, chiral order, and SRG) for 7 different HO frequencies (12–36 MeV)
- this generates a total of 340 200 unique training samples
- every sample fed to the network is a tuple of 3 observable sequences which differ in HO frequency for the same 4 consecutive $N_{\max}$ values each
  → the network has no information about the nucleus, interaction, SRG or the specific HO frequency or $N_{\max}$ values
- apply MINMAX normalization to each training sample to rescale all values of observable O to the interval [0,1]
  → prevents overfitting and alleviates scale dependence present in the ground-state energy when applying the networks to heavier nuclei
- statistical approach with 1000 ANNs and Gaussian fit to extract prediction and uncertainty

Conclusion

- data from light systems up to ^{4}He sufficient to train ANNs used for predictions in heavier systems
- generally outperforms classical extrapolation methods especially in smaller model spaces
- universal approach:
  - network topology works well for ground and excited-state energies, radii and possibly more observables
  - only train once per observable independent of nucleus, interactions, …

Outlook

- generate more independent training data using artificially modified potential or even toy potentials
  - improve training data quality and suppress any remaining biases
  - make observables accessible in few-body systems that are not present in realistic potential, e.g. electromagnetic transitions etc.
- expand to more observables
- further optimize network topology

Technische Universität Darmstadt, Germany

References

[1] B. R. Barrett, P. Navrátil, et al., Prog. Part. Nucl. Phys. **69**, 131 (2013).
[2] R. Roth, Phys. Rev. C **79**, 064324 (2009).
[3] G. A. Negoita, J. P. Vary, et al., Phys. Rev. C **99**, 054308 (2019).
[4] W. G. Jiang, G. Hagen, et al., Phys. Rev. C **100**, 054326 (2019).
[5] M. Knöll, T. Wolfgruber, et al., Phys. Lett. B **839**, 137781 (2023).
[6] P. Maris, E. Epelbaum, et al., Phys. Rev. C **103**, 054001 (2021).