Precise Neural Network Predictions from the NCSM

Robert Roth

Institut für Kernphysik - Theoriezentrum

SFB 1245

TECHNISCHE UNIVERSITÄT DARMSTADT

HFHF Helmholtz Forschungsakademie Hessen für FAIR

Ab Initio Nuclear Structure Toolbox

Many-Body Solution

Diagonalization & Decoupling
NCSM, IM-SRG, IM-NCSM,...

Pre-Conditioning

Similarity RG Transform
Basis Optimization

$$H\ |\Psi_n\rangle = E_n\,|\Psi_n\rangle$$

Post-Processing

Model-Space Extrapolation
Uncertainty Quantification

Chiral EFT Inputs

Interactions & Currents
NN, 3N, YN, YNN,...

Ab Initio Nuclear Structure Toolbox

Many-Body Solution

Diagonalization & Decoupling
NCSM, IM-SRG, IM-NCSM,...

Pre-Conditioning

Similarity RG Transform
Basis Optimization

$$H\,|\Psi_n\rangle = E_n\,|\Psi_n\rangle$$

Post-Processing

Model-Space Extrapolation
Uncertainty Quantification

Chiral EFT Inputs

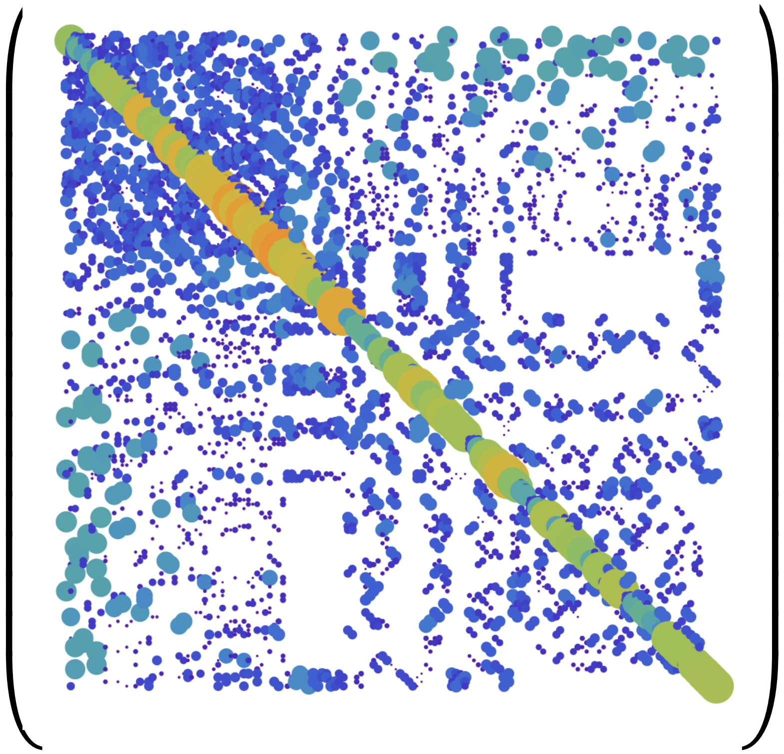
Interactions & Currents
NN, 3N, YN, YNN,...

No-Core Shell Model

Barrett, Vary, Navrátil, Maris, Nogga, Roth,...

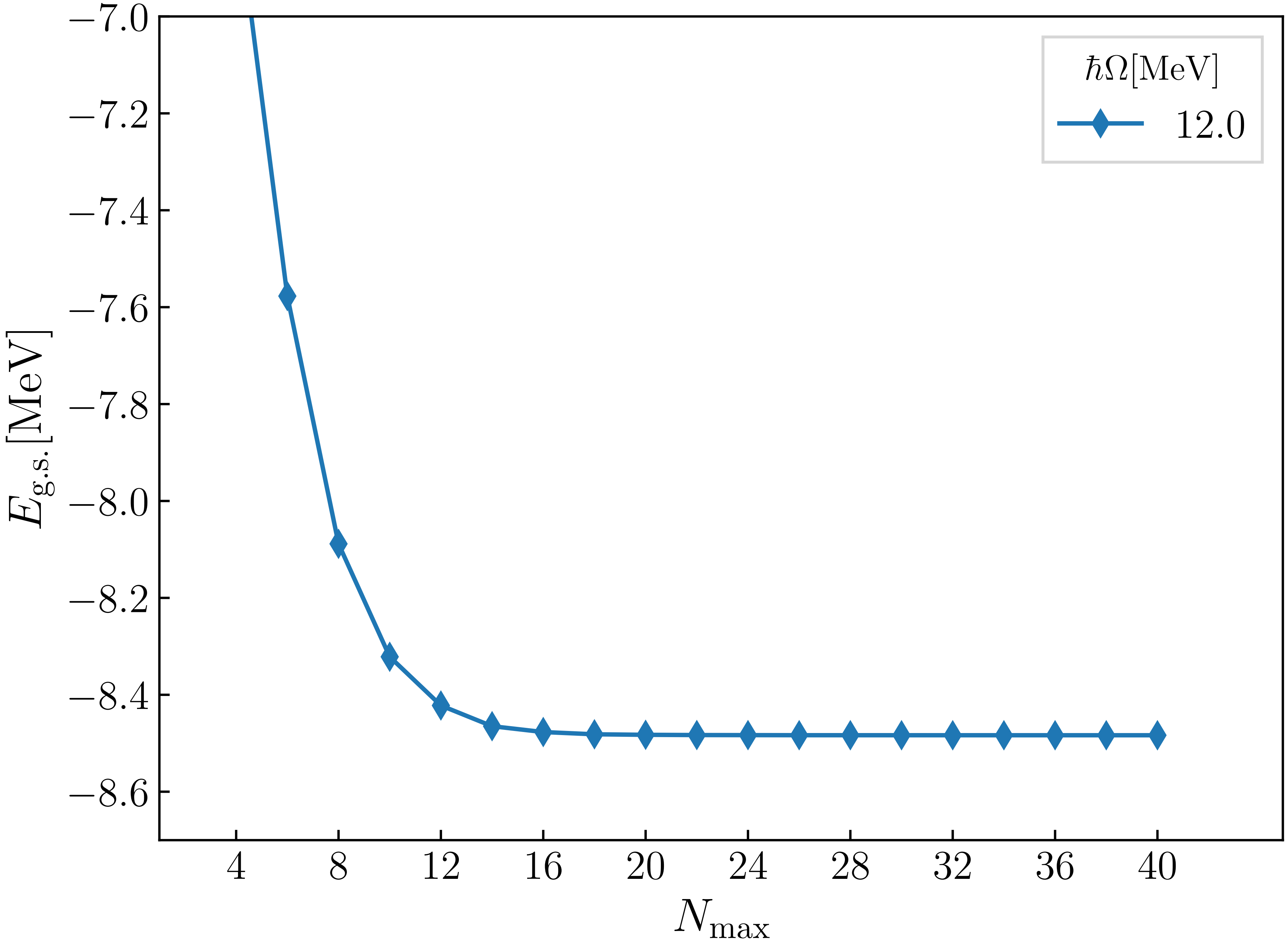
no-core shell model is
universal and powerful ab initio approach for
light nuclei (up to A≈25)

- solve eigenvalue problem of Hamiltonian represented in model space of HO Slater determinants truncated w.r.t. HO excitation energy $N_{\text{max}}\hbar\Omega$



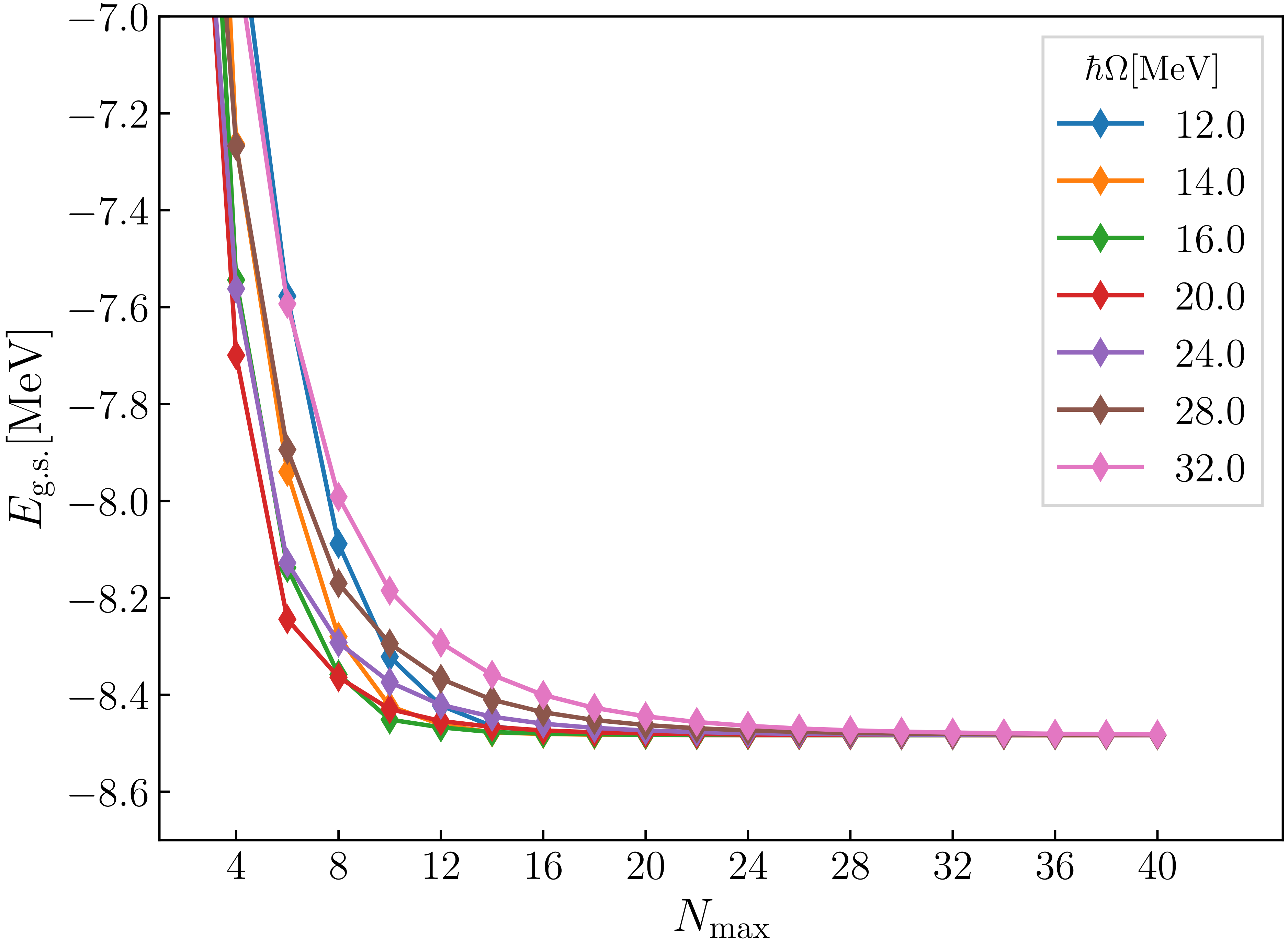
$$\begin{pmatrix} \vdots \\ C^{(n)}_{i'} \\ \vdots \end{pmatrix} = E_n \begin{pmatrix} \vdots \\ C^{(n)}_{i} \\ \vdots \end{pmatrix}$$

NCSM: Simplest Possible Case



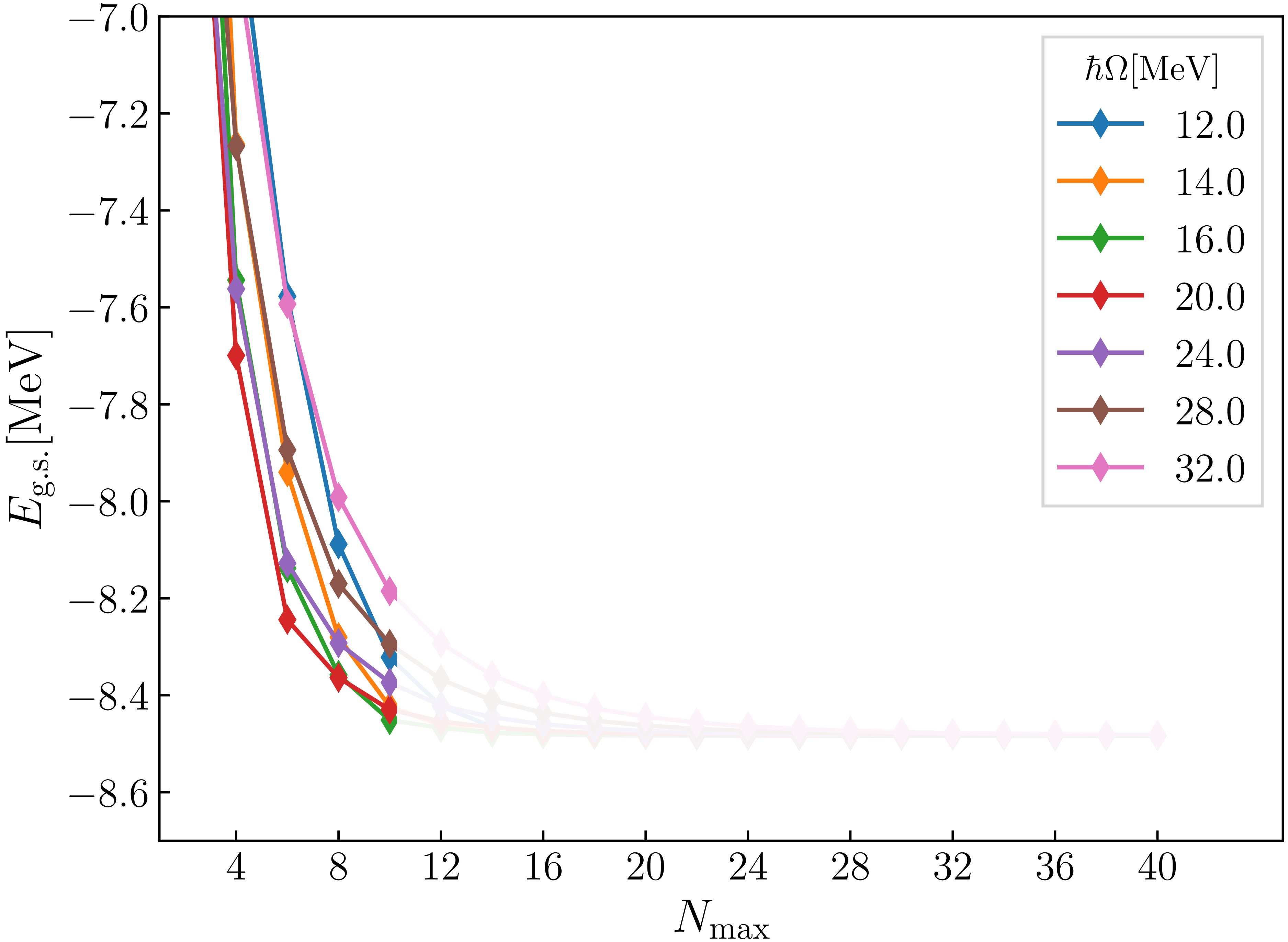
- ground-state energy of ^3H with SRG-evolved NN+3N interaction

NCSM: Simplest Possible Case



- ground-state energy of ^3H with SRG-evolved NN+3N interaction

NCSM: Simplest Possible Case



- for mid-p-shell nuclei N_{max}=8,10 is computational limit
- can we still get an approximation of the converged observable via post-processing?

Model-Space Extrapolations

- try to model the N_{max} and $\hbar\Omega$ dependence and then extrapolate using...

- **heuristic extrapolation schemes**

  Maris, et al., PRC 79, 014308 (2009)

  ...typically based on exponential parametrizations
  ...combined with manual frequency selection/optimization
  ...works OK for energies, not suitable for other observables

- **effective-theory-based IR extrapolations**

  Coon et al., PRC 86, 054002 (2012)
  Furnstahl et al., PRC 86, 031301 (2012)
  More et al., PRC 87, 044326 (2013)
  Furnstahl et al., PRC 89, 044301 (2014)

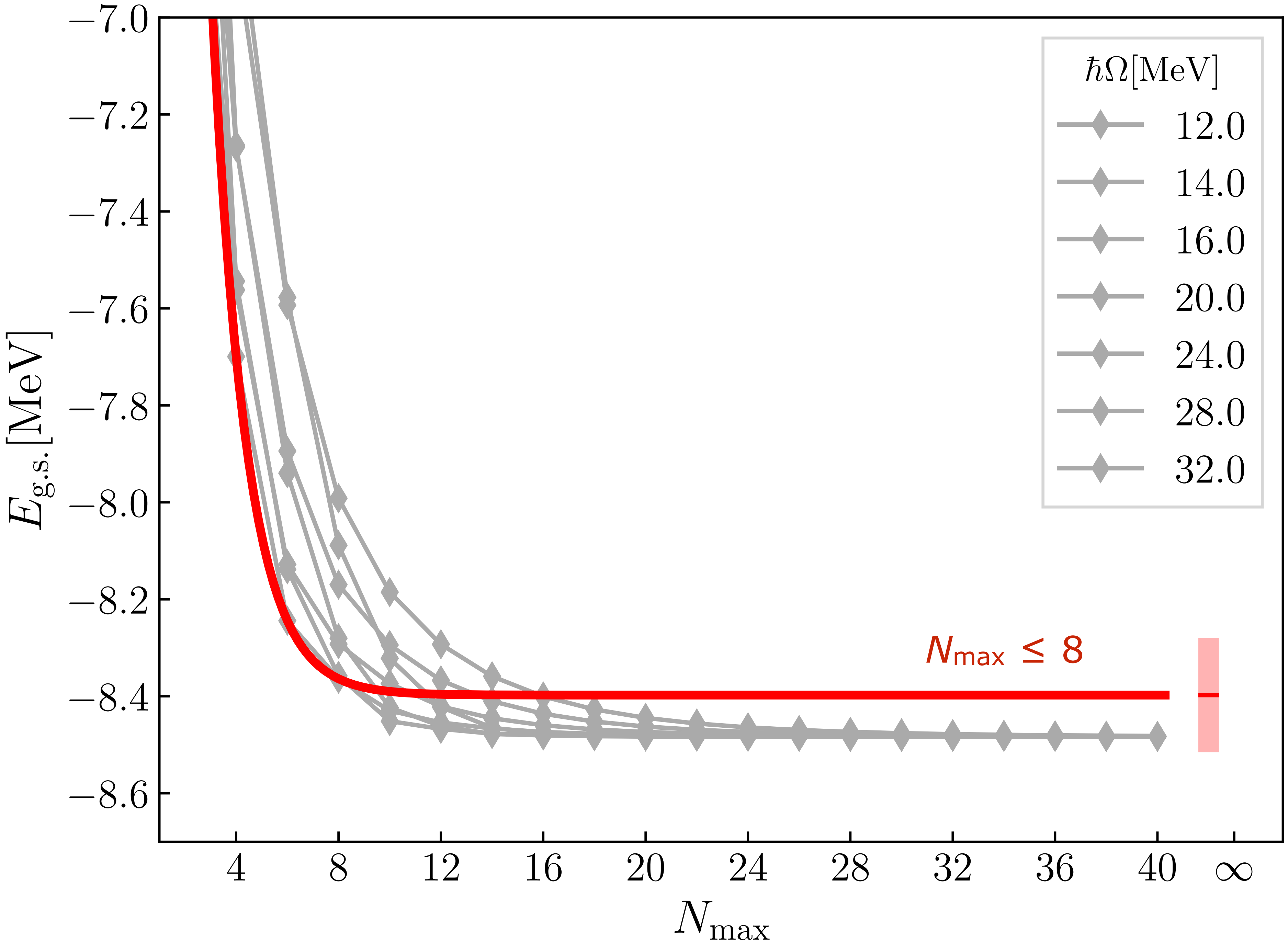
  ...difficult and expensive to obtain enough UV-converged data
  ...not really compatible with N_{max}-type many-body truncations
  ...hardly ever applied in production runs

- **artificial neural networks**

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

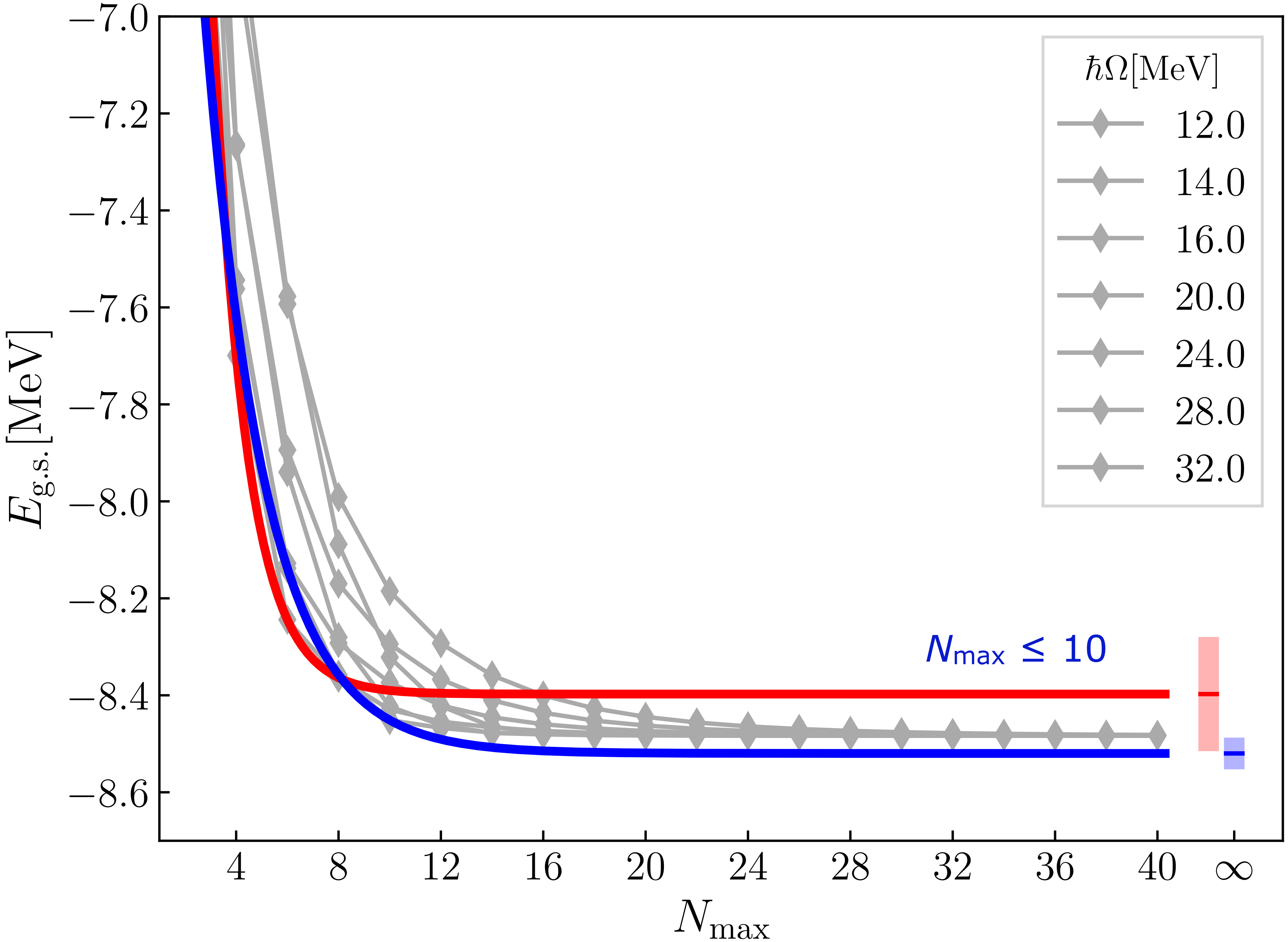
  ...needs retraining of ANN for each interaction, nucleus, state
  ...requires lots of training data for specific application case
  ...ANNs are typically not that great for extrapolations

Model-Space Extrapolations



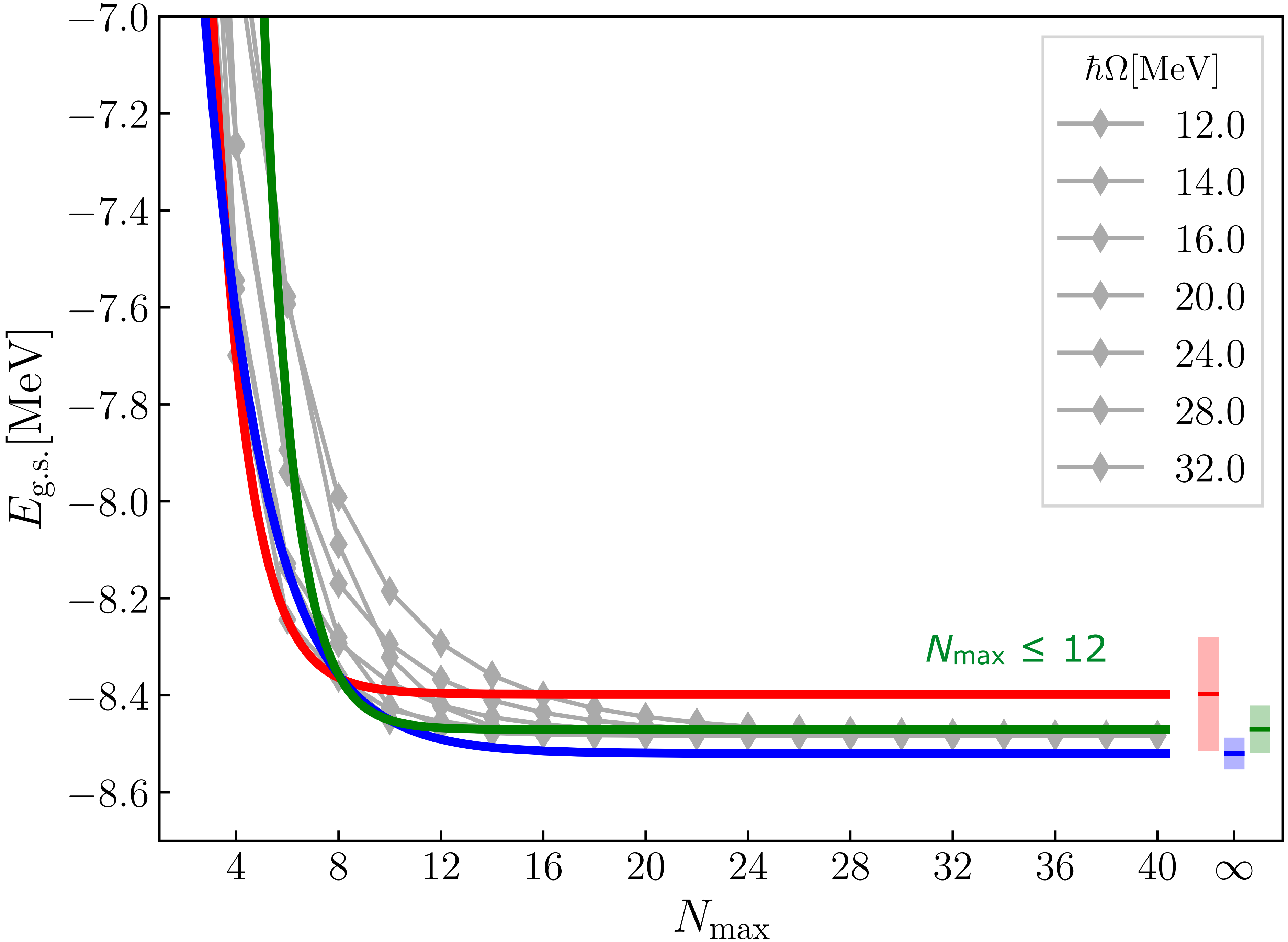
- simple exponential extrapolations might fail
- N_{max} dependence is not exponential

Model-Space Extrapolations



- simple exponential extrapolations might fail
- N_{max} dependence is not exponential

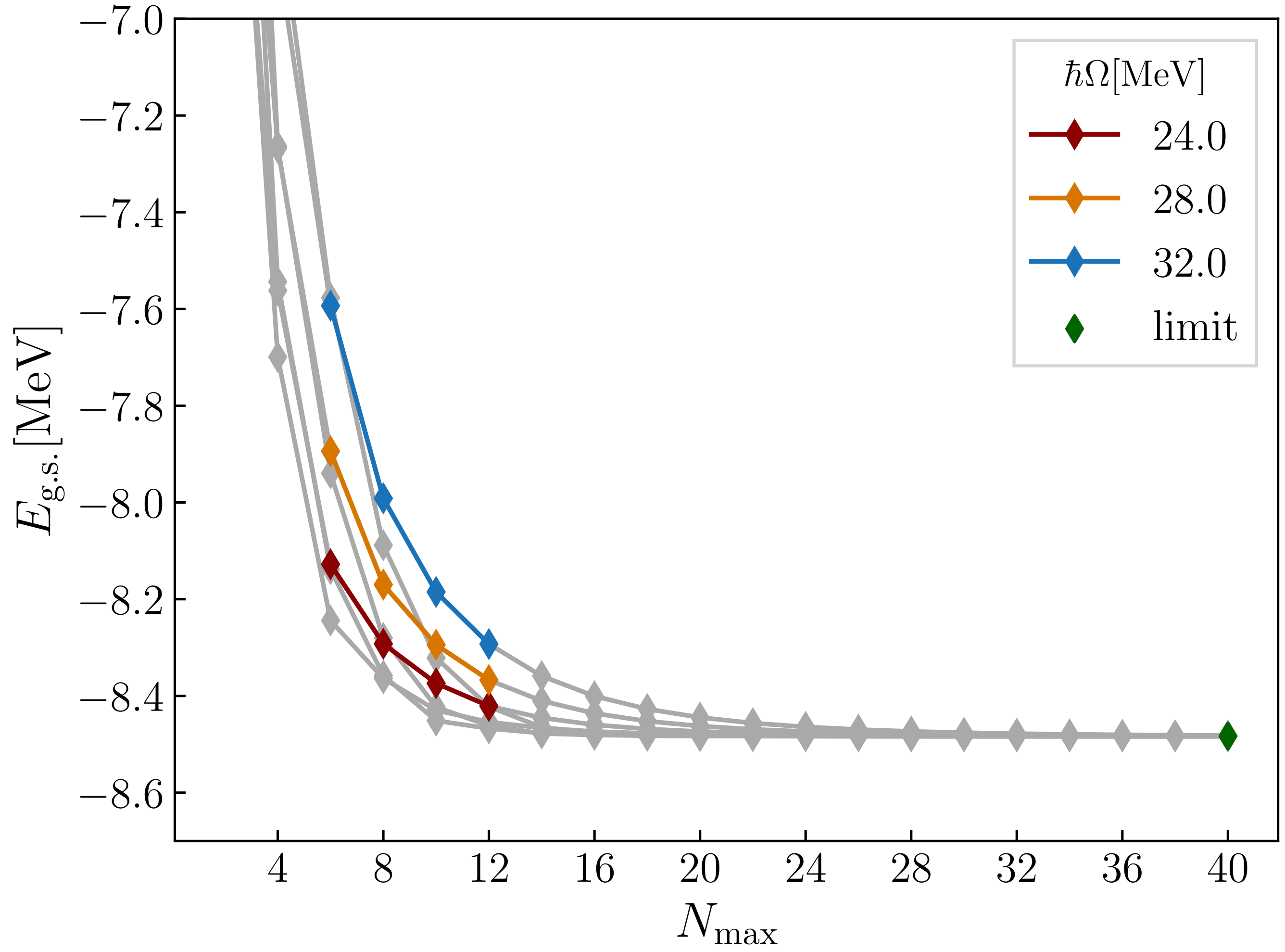
Model-Space Extrapolations



- simple exponential extrapolations might fail
- N_{max} dependence is not exponential

Predicting Converged Result

Knöll, et al.; PLB 839, 137781 (2023); Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

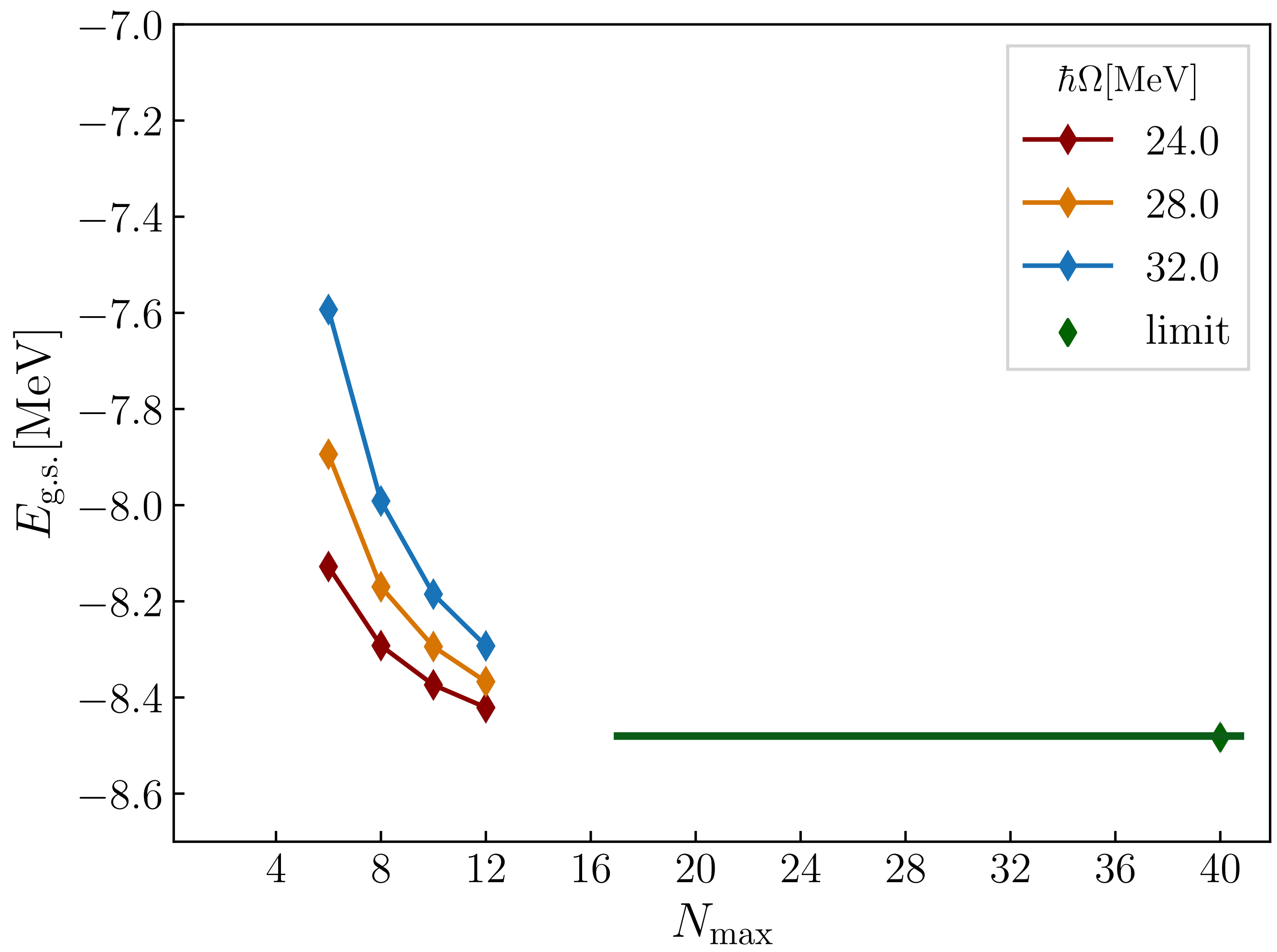


as a practitioner, just seeing the **pattern of converging sequences** allows you to guesstimate the converged result

irrespective of nucleus, interaction,...

Predicting Converged Result

Knöll, et al.; PLB 839, 137781 (2023); Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

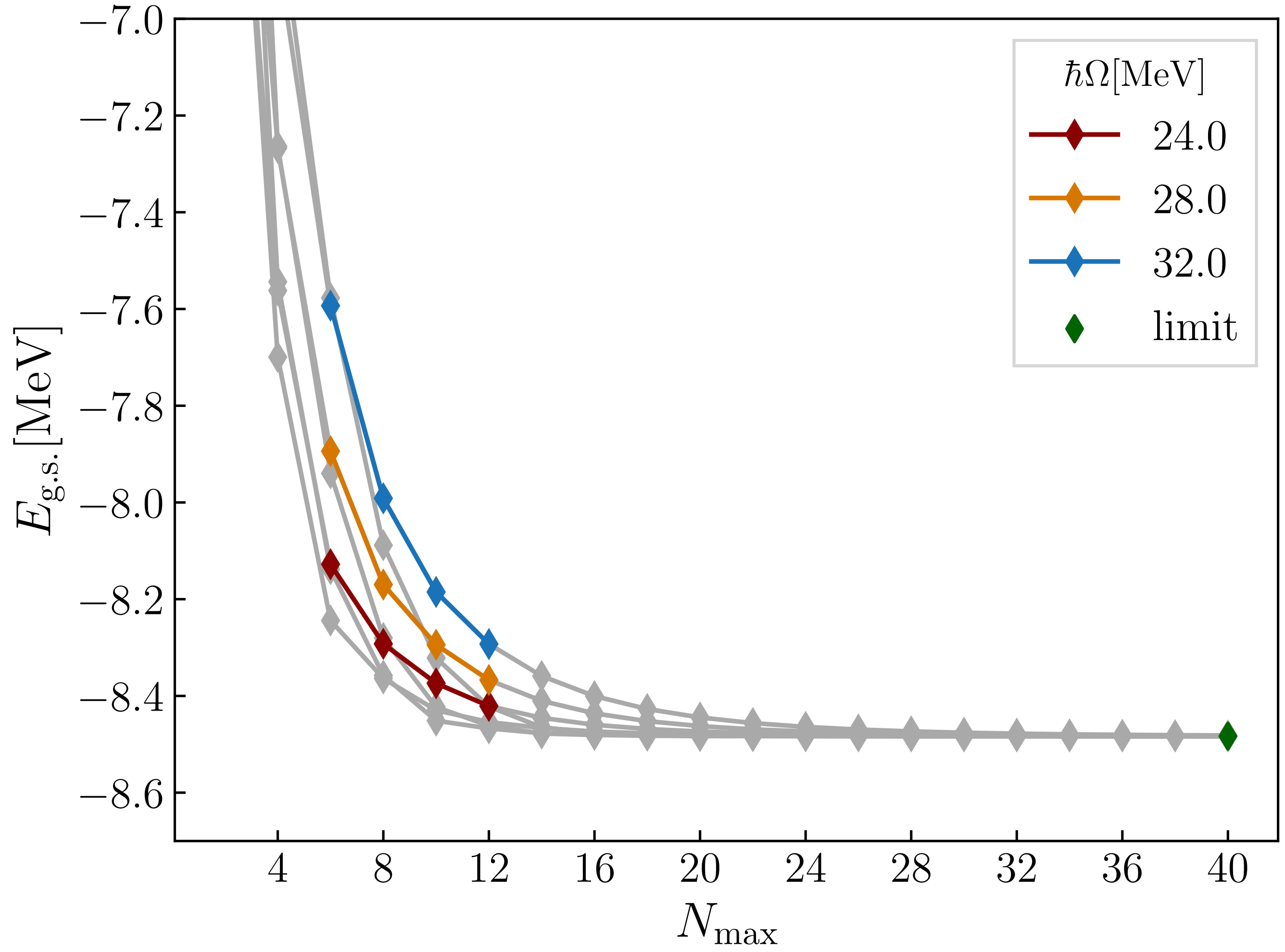


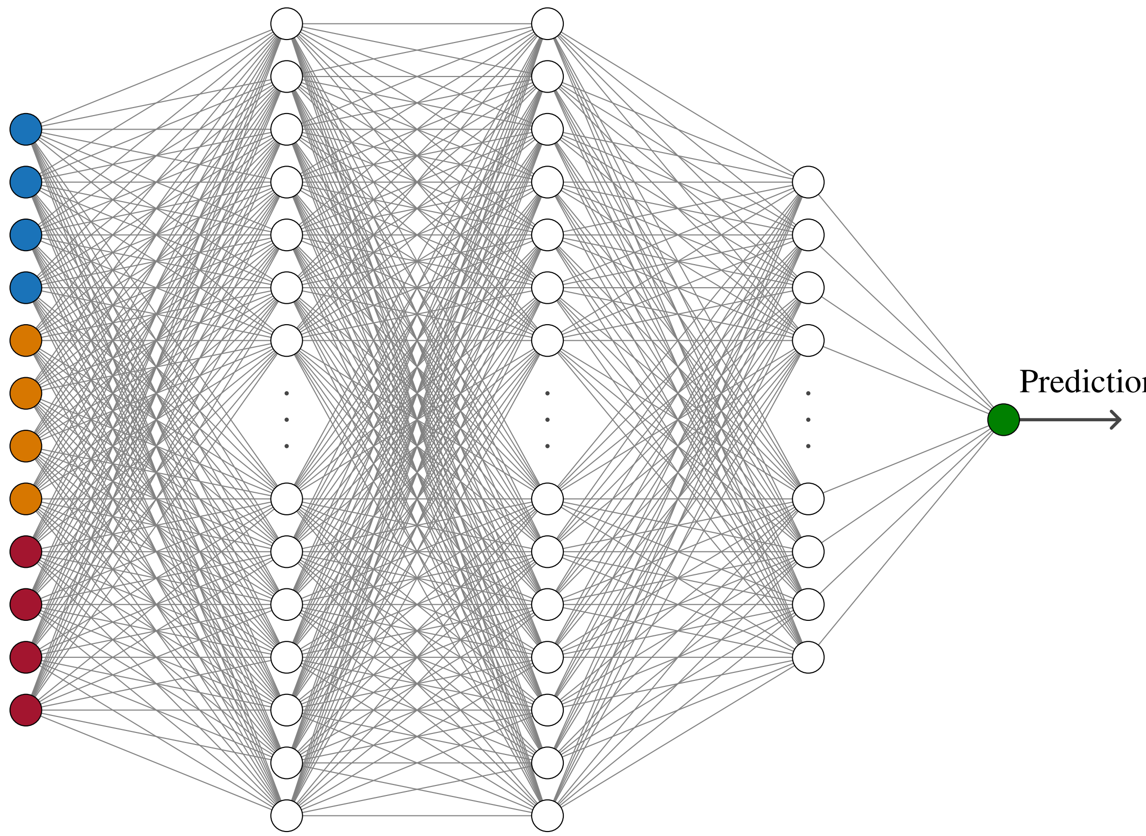
as a practitioner, just seeing the **pattern of converging sequences** allows you to guesstimate the converged result

irrespective of nucleus, interaction,...

Predicting Converged Result

Knöll, et al.; PLB 839, 137781 (2023); Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)



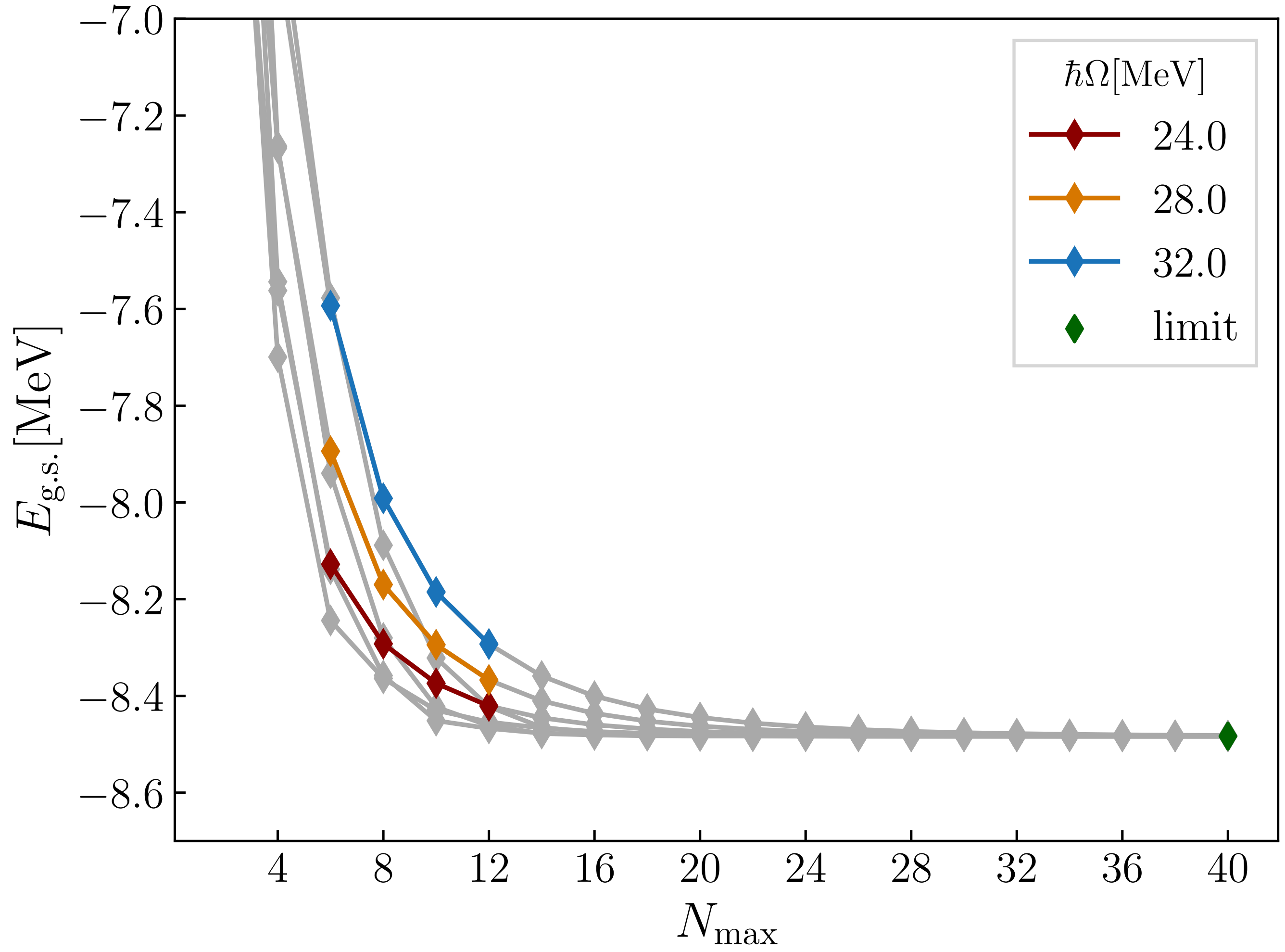


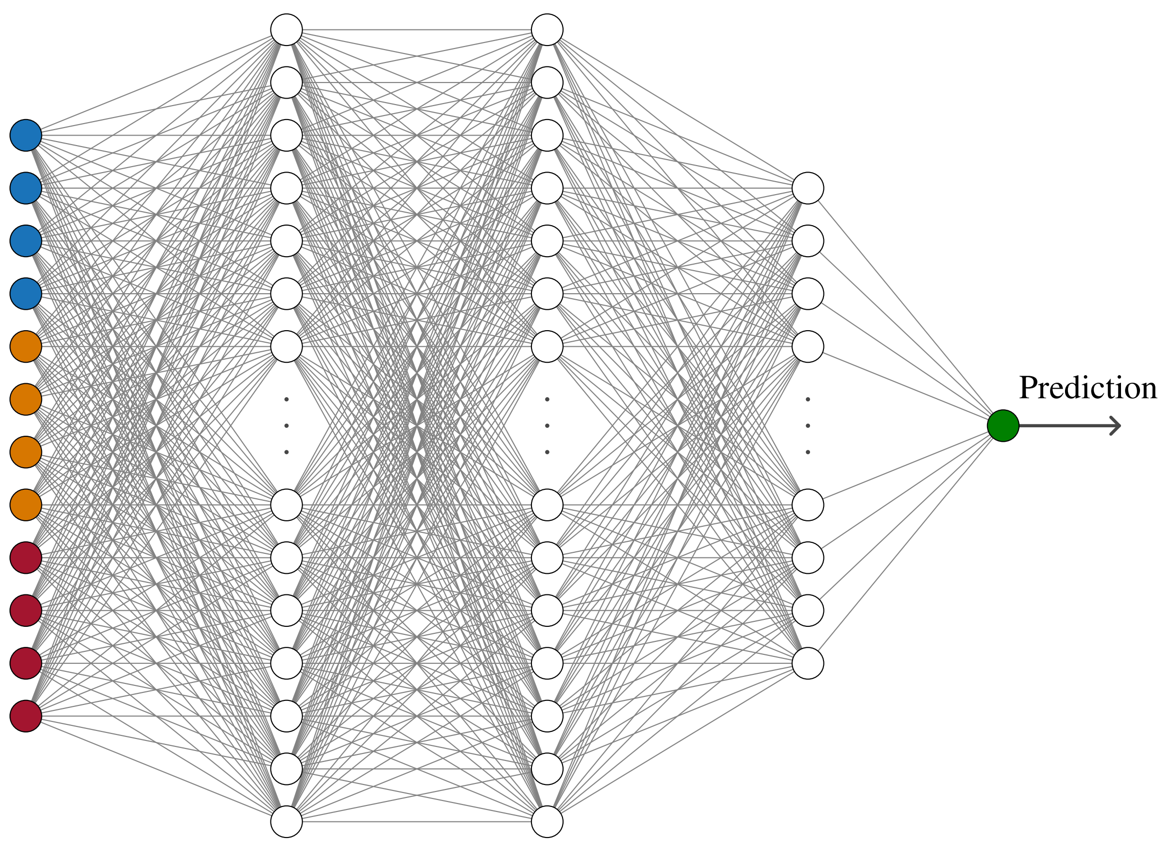
train **neural network** to identify the convergence patterns and to **predict the converged value** on this basis

ANNs are great for pattern recognition

Predicting Converged Result

Knöll, et al.; PLB 839, 137781 (2023); Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

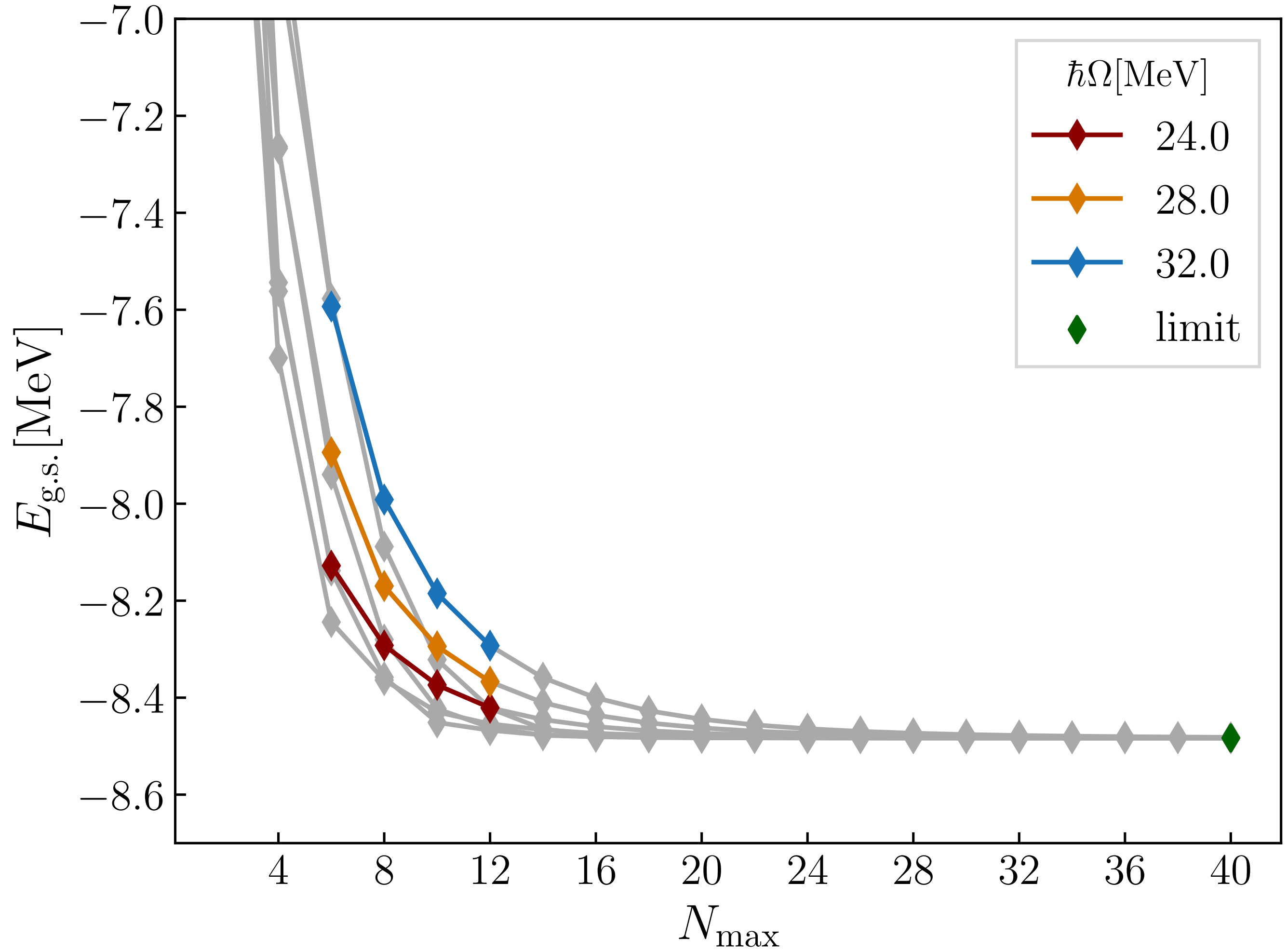


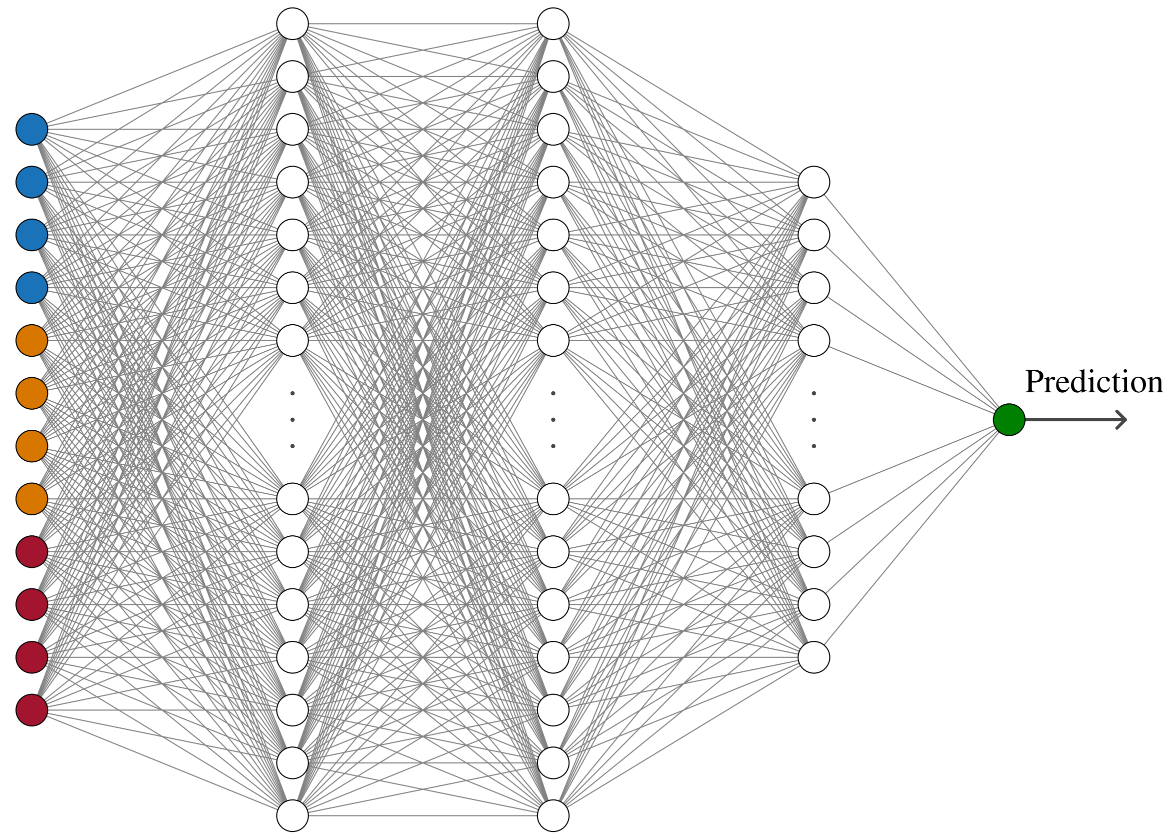


- create data **samples** of 4 consecutive N_{max} and 3 different $\hbar\Omega$
- feed 12 values of observable into dense, deep, feed-forward network
- output is a single value for the converged observable
- ANN has not information about actual interaction, nucleus, N_{max}, $\hbar\Omega$, etc.

Predicting Converged Result

Knöll, et al.; PLB 839, 137781 (2023); Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)



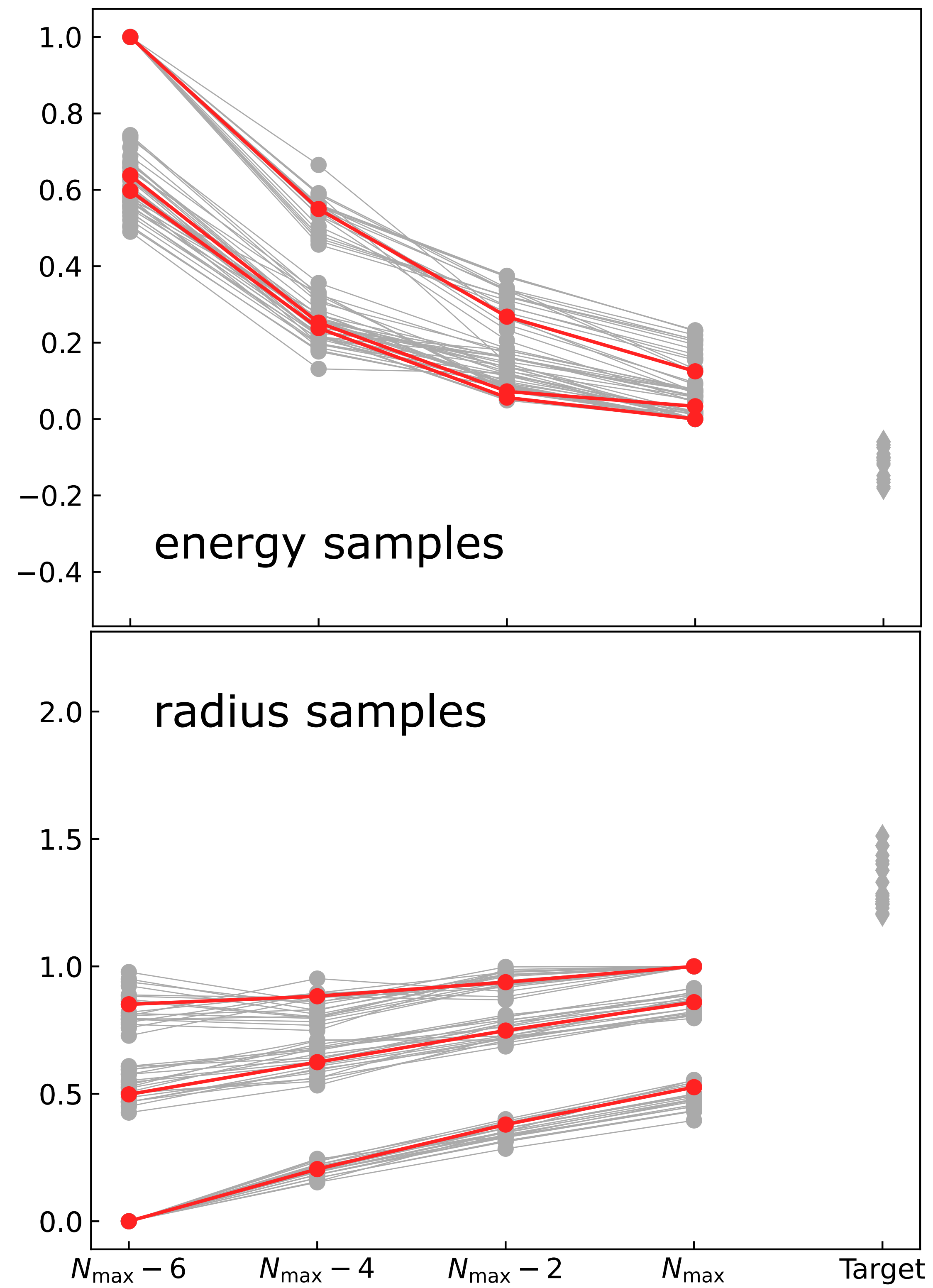


- 3 hidden layers with (48,48,24) neurons using ReLU activation function
  ➔ approx. 4200 weight/bias parameters
- hyper-parameter optimization yields this as robust topology
- use normalization of input data to eliminate nucleus-specific scales

Training Data, Samples, Normalization

Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

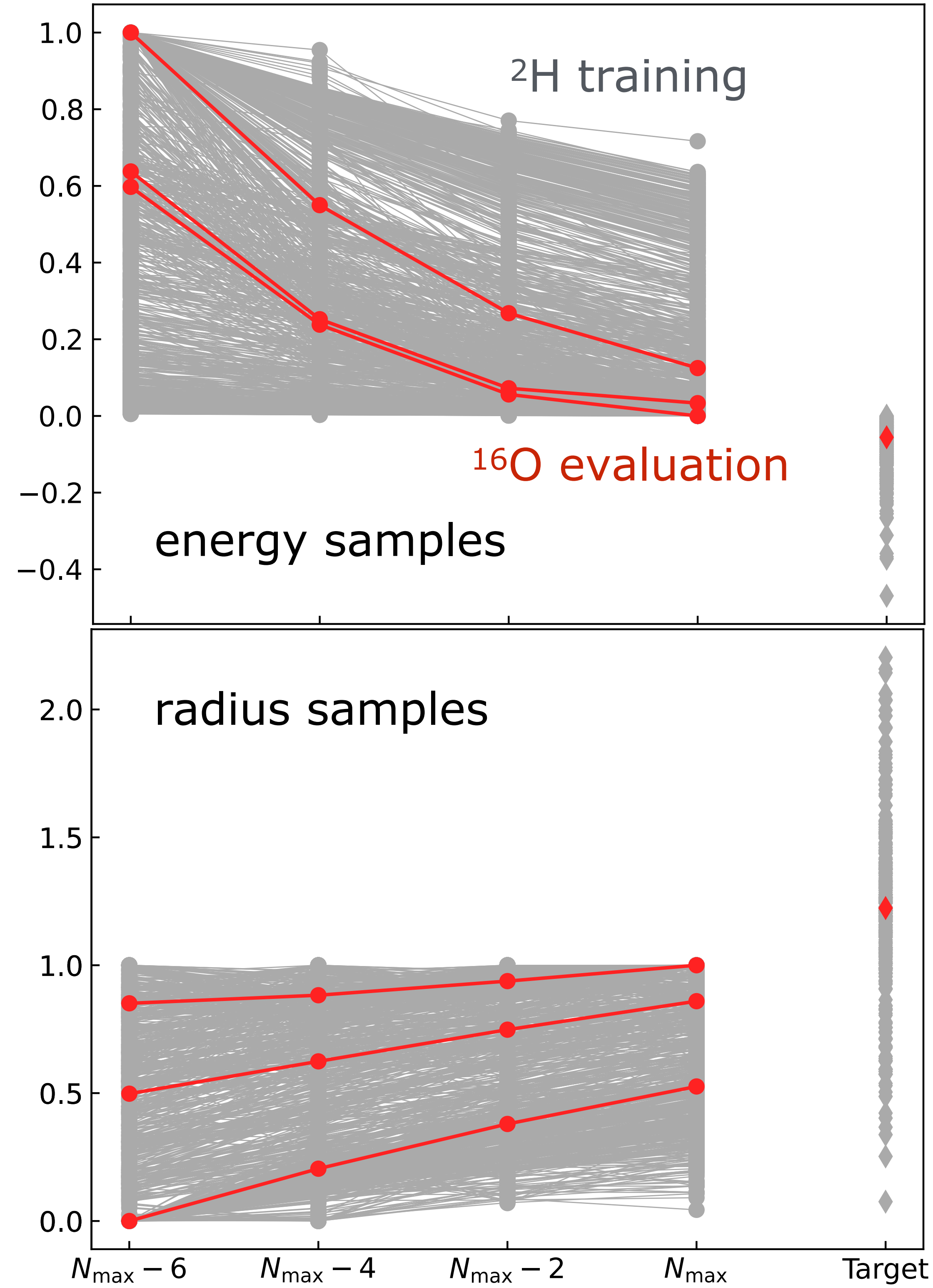
- large library of **training data** obtained in Jacobi-NCSM calculations for
  - ^{2}H, ^{3}H, ^{4}He
  - 9 different NN+3N interactions (non-local, 3 orders, 3 cutoffs)
  - 5 different SRG flow parameter (α = 0, 0.02, 0.04, 0.08, 0.16 fm^4)
  - approx. 350000 unique samples
- **sample-wise normalization** of training and evaluation data to eliminate nucleus-specific scale
- **universality** visible at the level normalized samples



Training Data, Samples, Normalization

Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

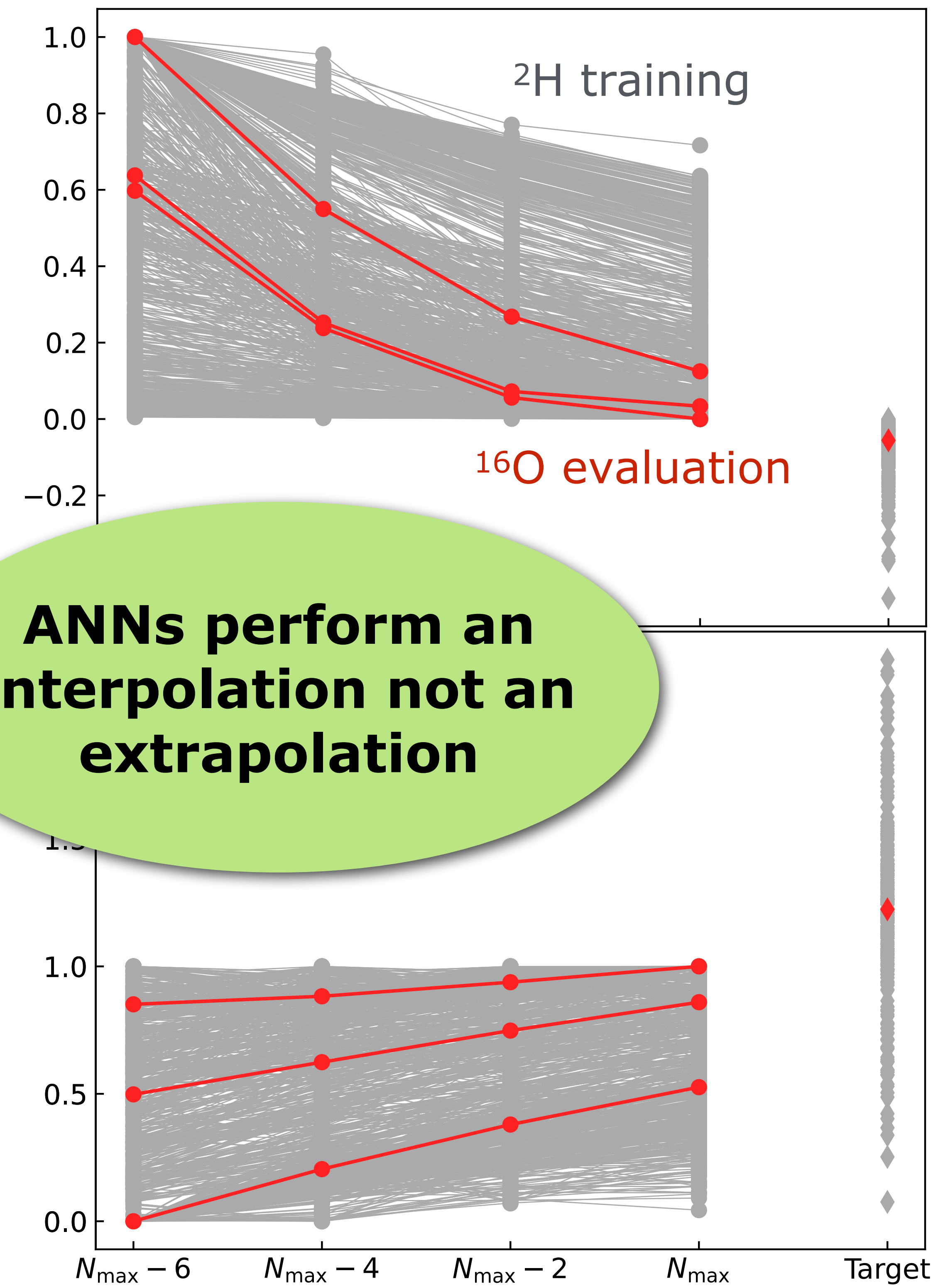
- large library of **training data** obtained in Jacobi-NCSM calculations for
  - ^2H, ^3H, ^4He
  - 9 different NN+3N interactions (non-local, 3 orders, 3 cutoffs)
  - 5 different SRG flow parameter (α = 0, 0.02, 0.04, 0.08, 0.16 fm^4)
  - approx. 350000 unique samples
- **sample-wise normalization** of training and evaluation data to eliminate nucleus-specific scale
- **universality** visible at the level normalized samples



Training Data, Samples, Normalization

Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

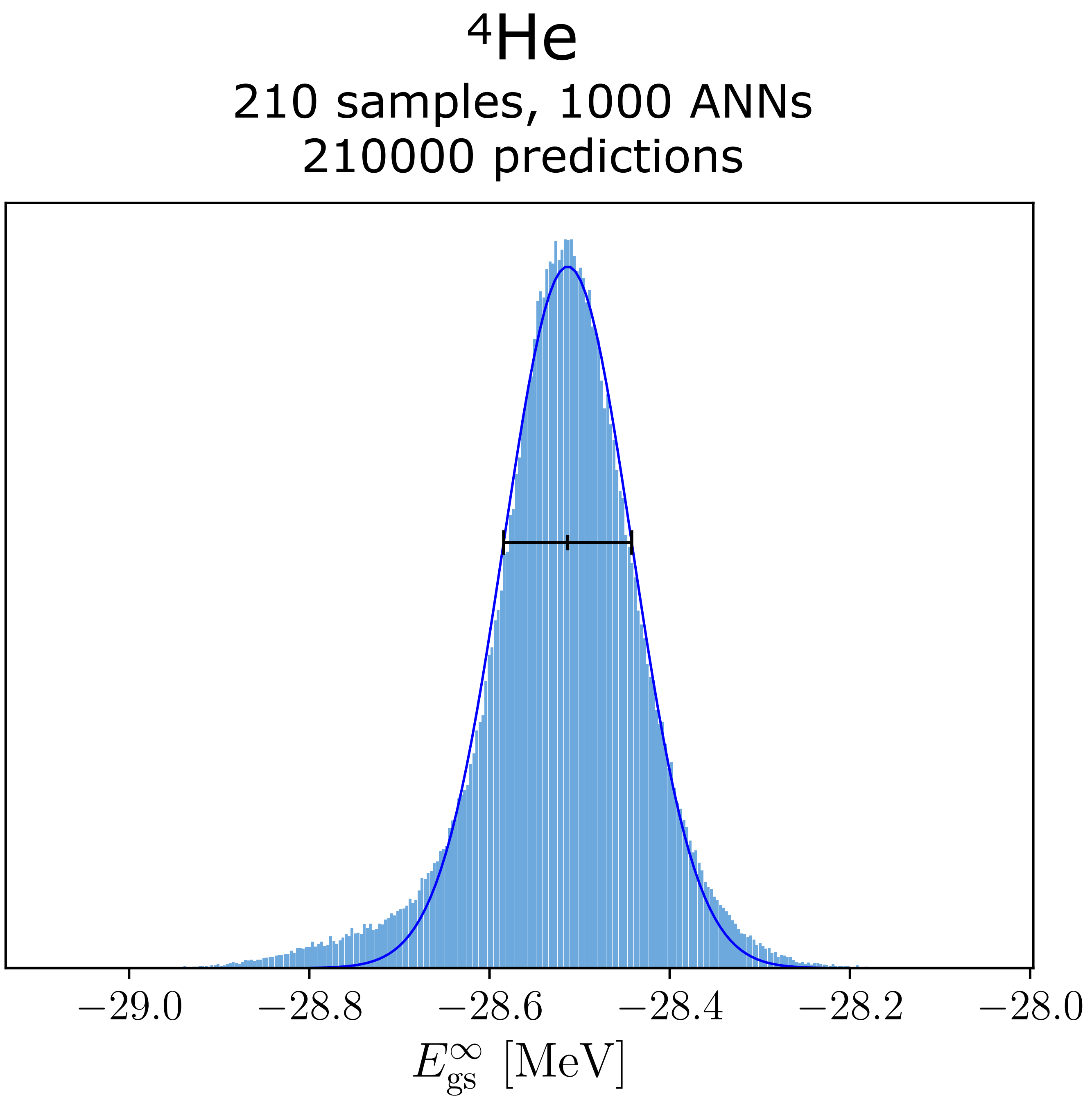
- large library of **training data** obtained in Jacobi-NCSM calculations for
  - ^{2}H, ^{3}H, ^{4}He
  - 9 different NN+3N interactions (non-local, 3 orders, 3 cutoffs)
  - 5 different SRG flow parameter (α = 0, 0.02, 0.04, 0.08, 0.16 fm^4)
  - approx. 350000 unique samples
- **sample-wise normalization** of training and evaluation data to eliminate nucleus-specific scale
- **universality** visible at the level normalized samples



ANNs perform an interpolation not an extrapolation

Statistical Evaluation

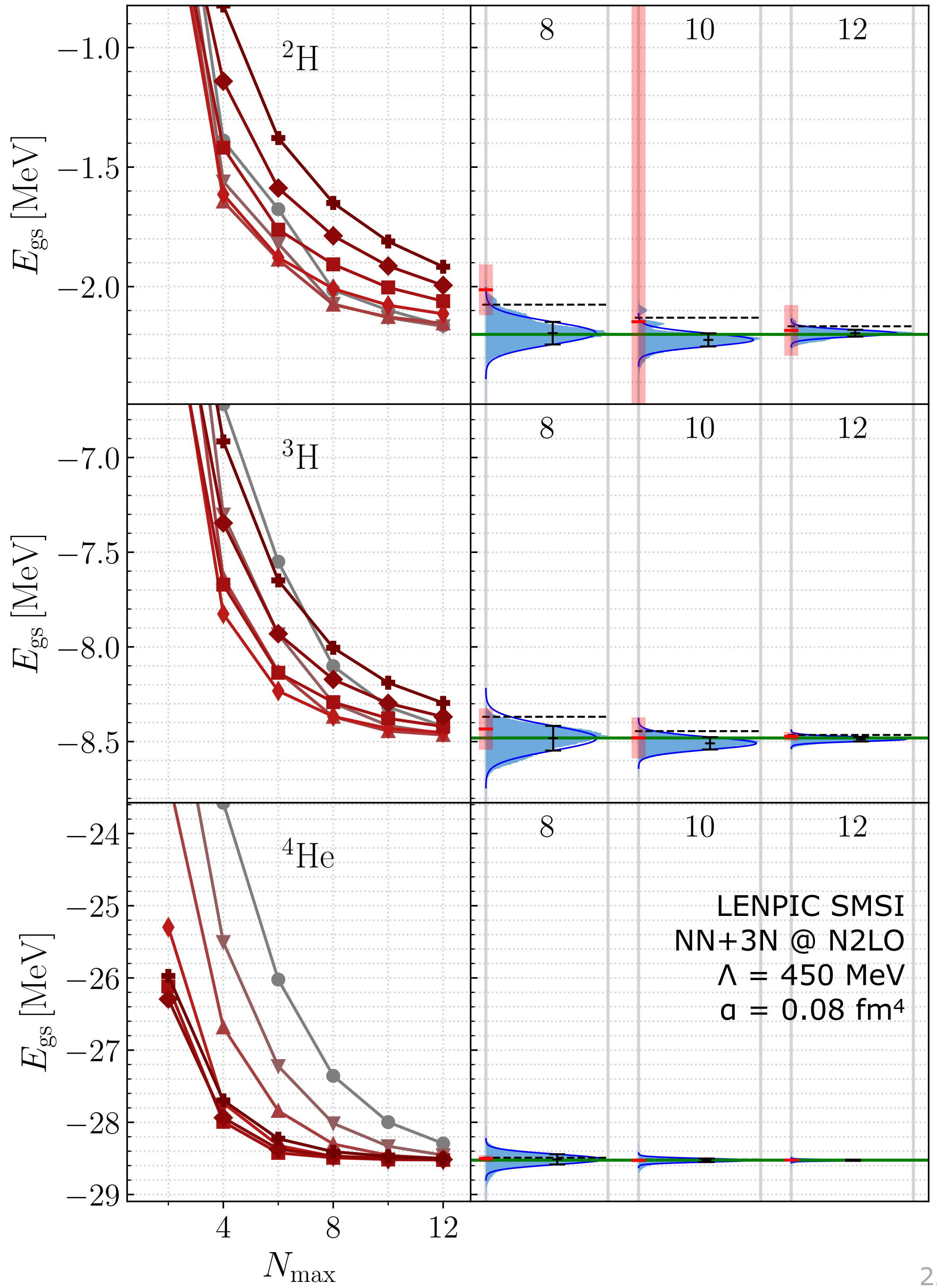
- obtain multiple predictions of converged observable from
  - different samples constructed from **evaluation data**
  - **several ANN realizations** from separate, randomly initialized training runs
- distribution of predictions provides statistically meaningful estimate for **model-space uncertainty**
- typically close to Gaussian distribution, use mean and standard deviation



Examples

Ground-State Energies

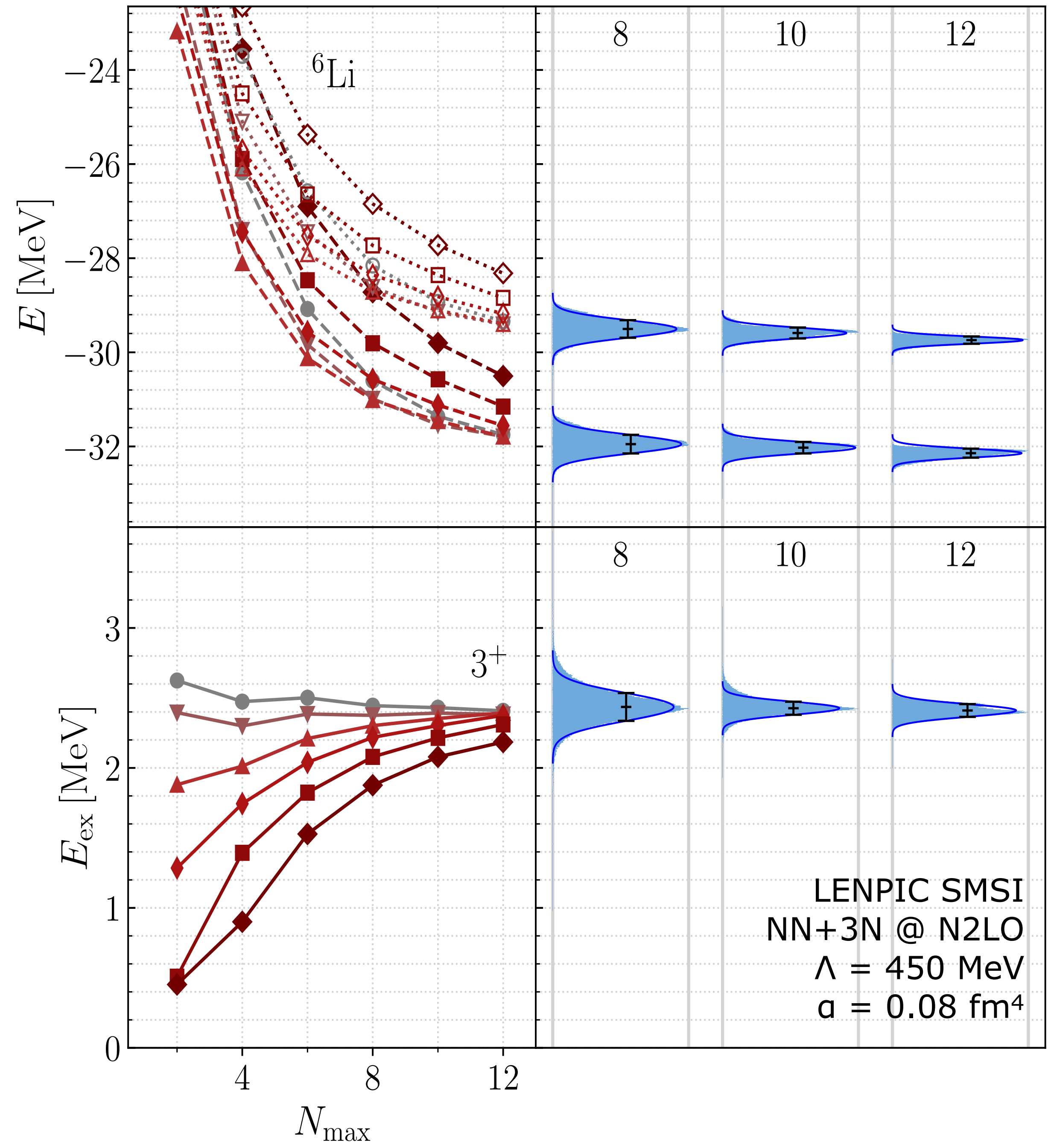
- training yields **standard set of 1000 energy-ANNs** that is used for all evaluations of energies
- application to NCSM data obtained with different NN+3N interaction not seen in training
- separate samples by largest $N_{\max}$=8,10,12 to assess robustness and consistency of predictions
- predictions always agree with exact result within 1σ



Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

Difference Observables

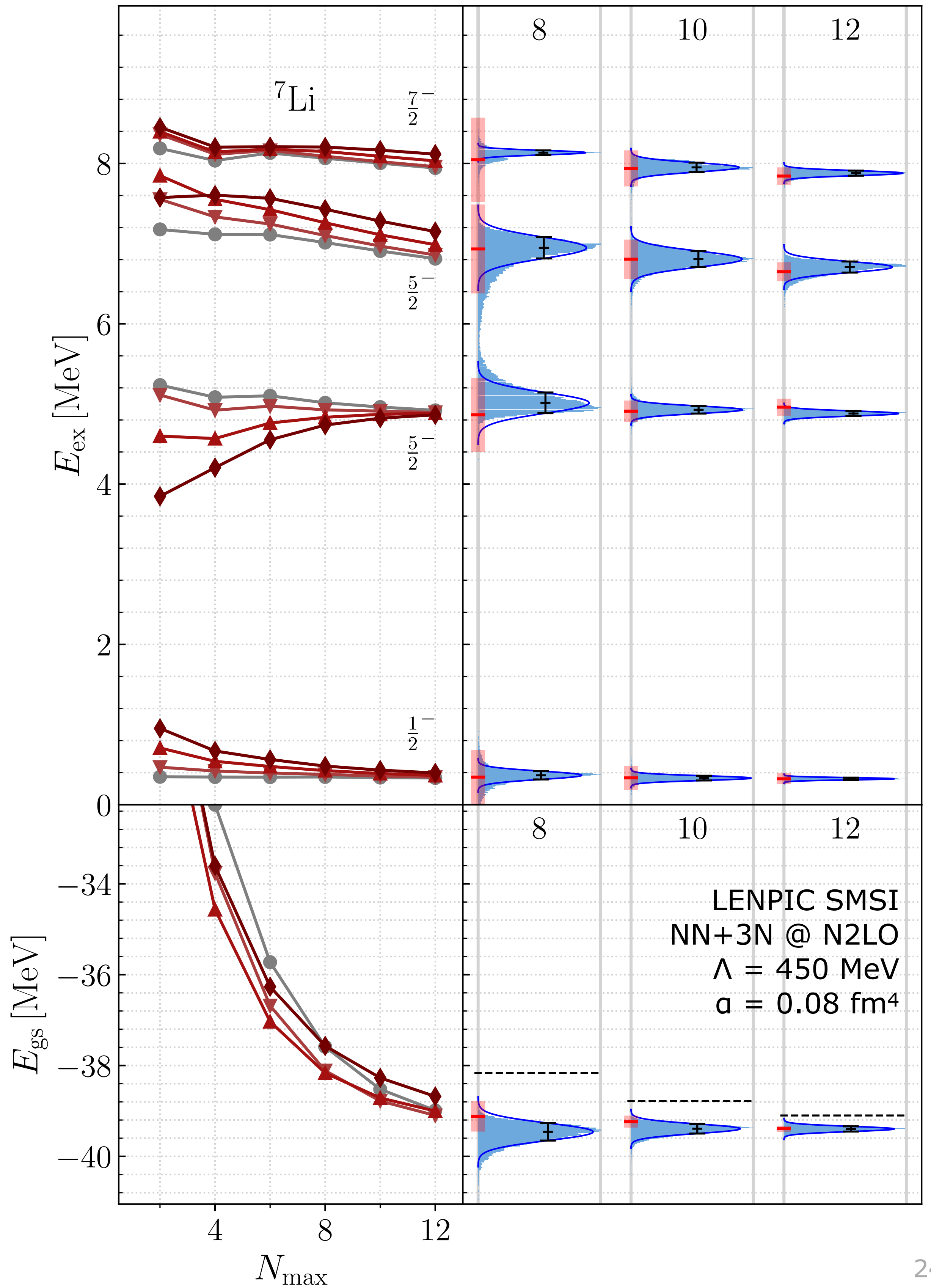
- use the same **standard set of 1000 energy-ANNs** to predict excited-state energies and excitation energies
- for each sample use ANN to predict converged ground-state and excited state energy
- sample-wise difference yields excitation energy
- construct distribution by looping over all samples and ANNs
- for difference quantities correlations in the convergence patterns are taken into account



Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

Excitation Energies

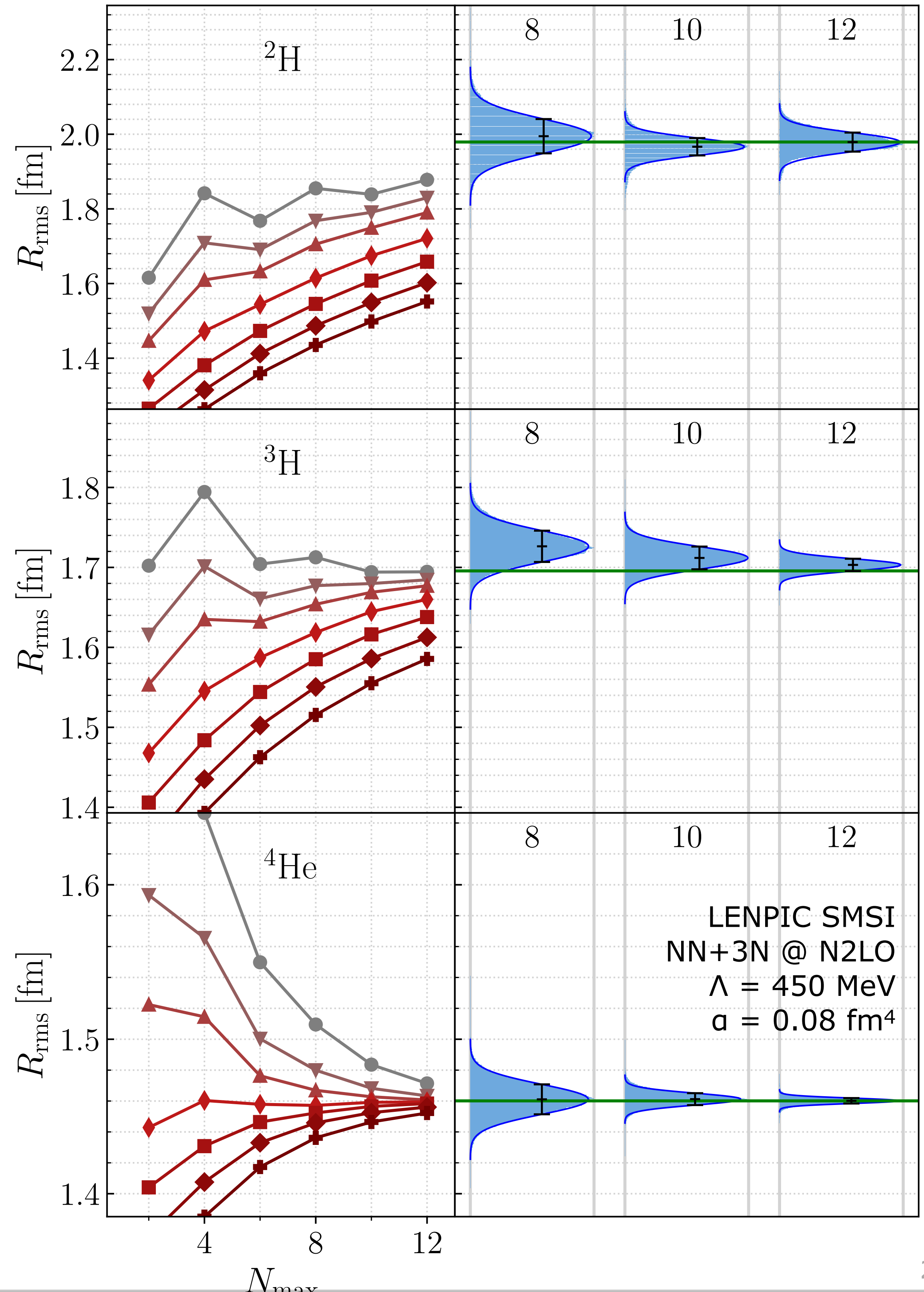
- full access to ground-state and excitation energies based on one set of energy-ANNs
- robust and consistent predictions for p-shell spectra
- no need for separate excited state to excitation energy ANNs



Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

Root-Mean-Square Radii

- training yields **standard set of 1000 radius-ANNs** that is used for all evaluations of radii
- much more difficult and varied convergence patterns
- still robust and consistent predictions that almost always agree with exact result within 1σ

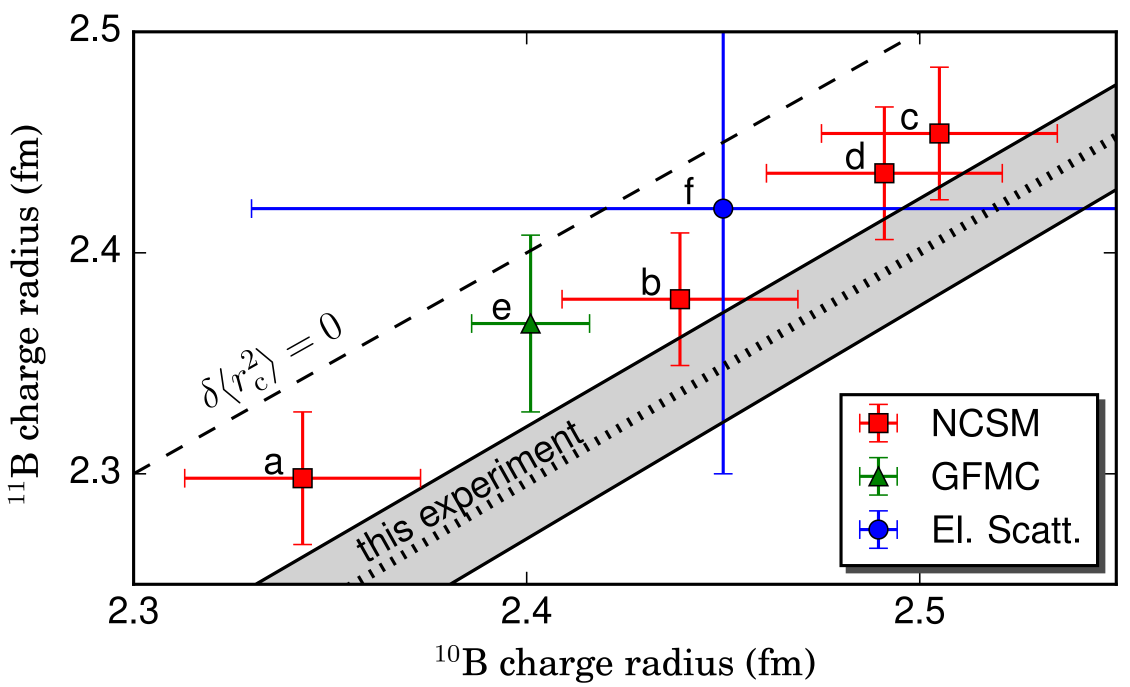


Wolfgruber, Knöll, Roth; arXiv:2310.05256 (2023)

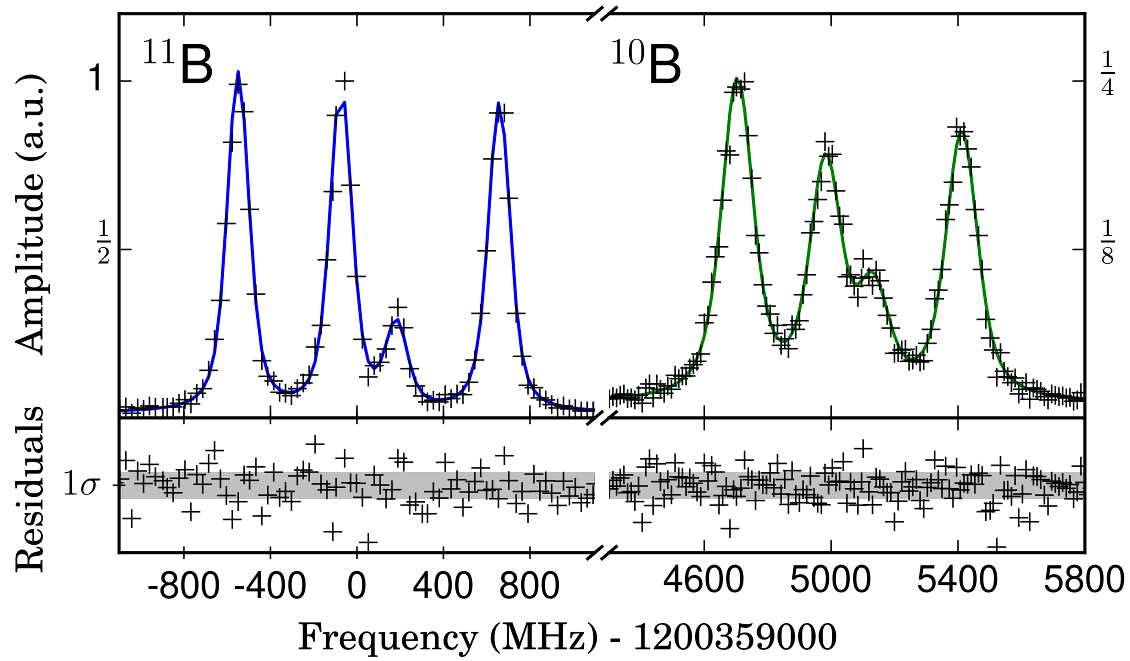
Boron Isotopes

Boron Isotopes

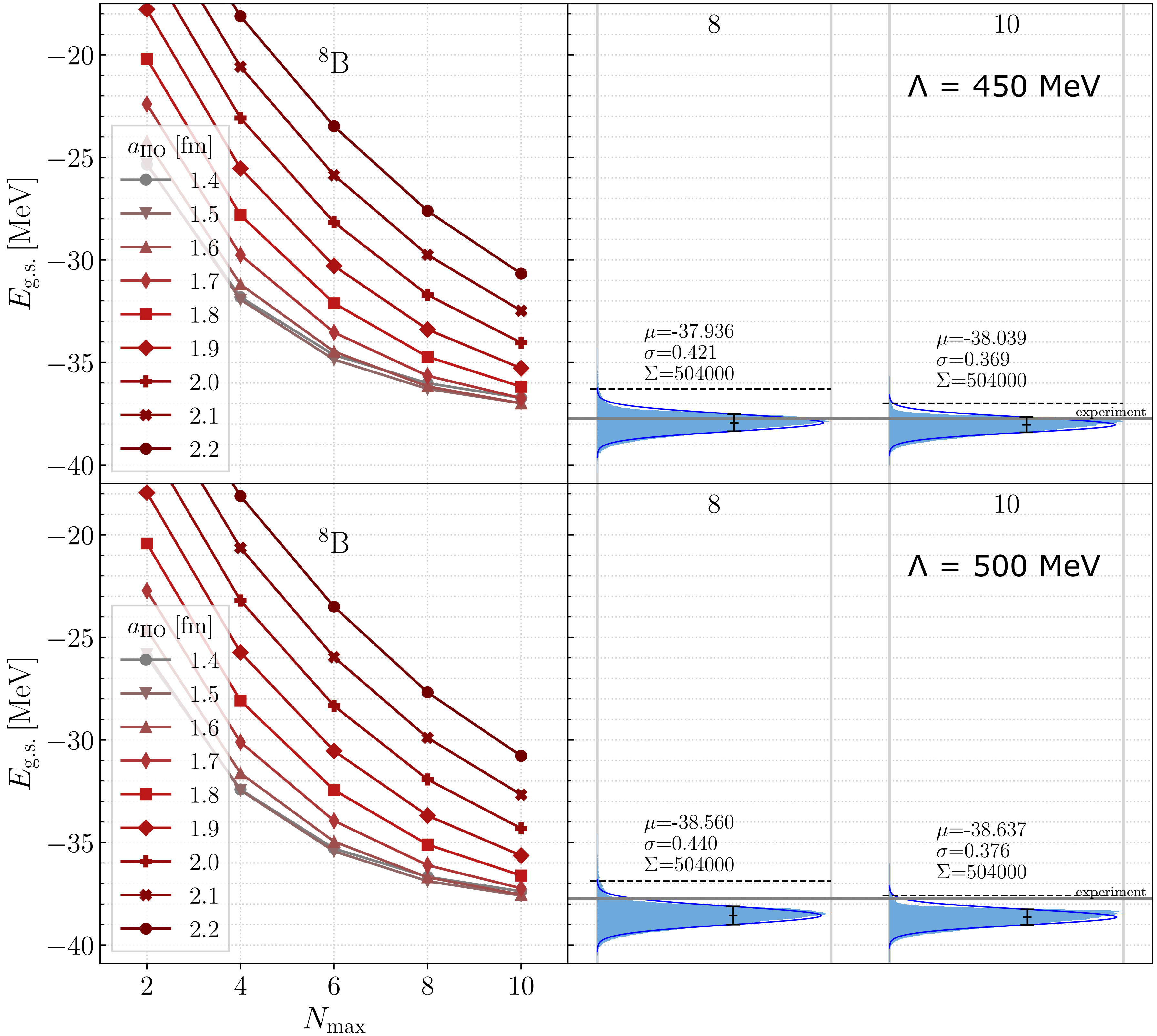
- nice example of mid-p-shell isotopic chain with interesting spectroscopy
- proton-dripline nucleus ^{8}B is one of few **proton-halo** candidates
- **precision measurements** of proton-radius differences by atomic laser spectroscopy via isotope shift (Nörtershäuser et al.)
  - precision measurement of ^{11}B-^{10}B radius difference exists; improved measurement at COALA beamline at TUDa later this year
  - planned experiment on ^{8}B-$^{10,11}B$ radius difference at ANL



B. Maaß, T. Hüther, et al.; PRL 122, 182501 (2019)



^{8}B: Ground-State Energy



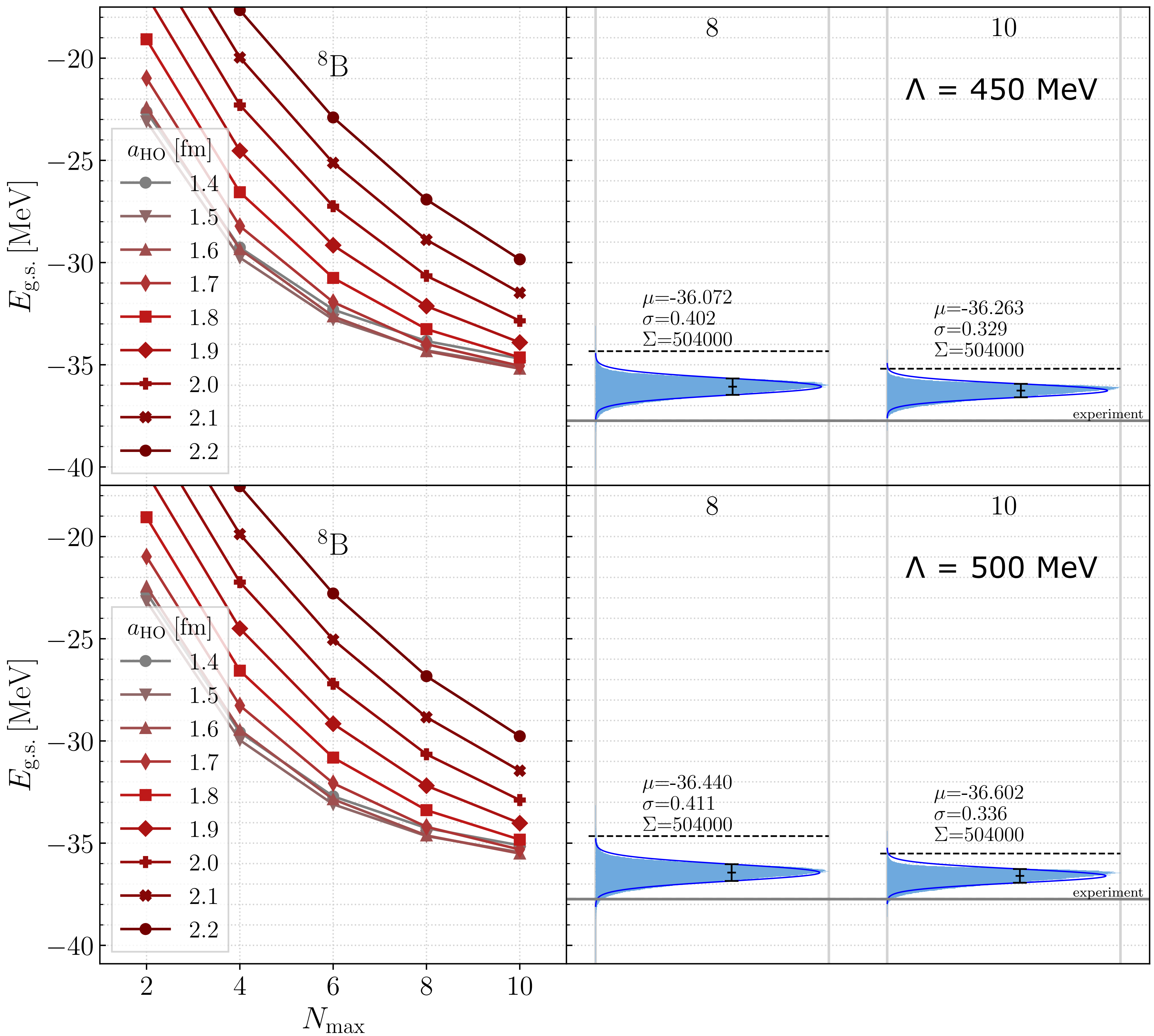
LENPIC SMSI

NN @ N2LO
3N @ N2LO

Λ = 450/500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{8}B: Ground-State Energy



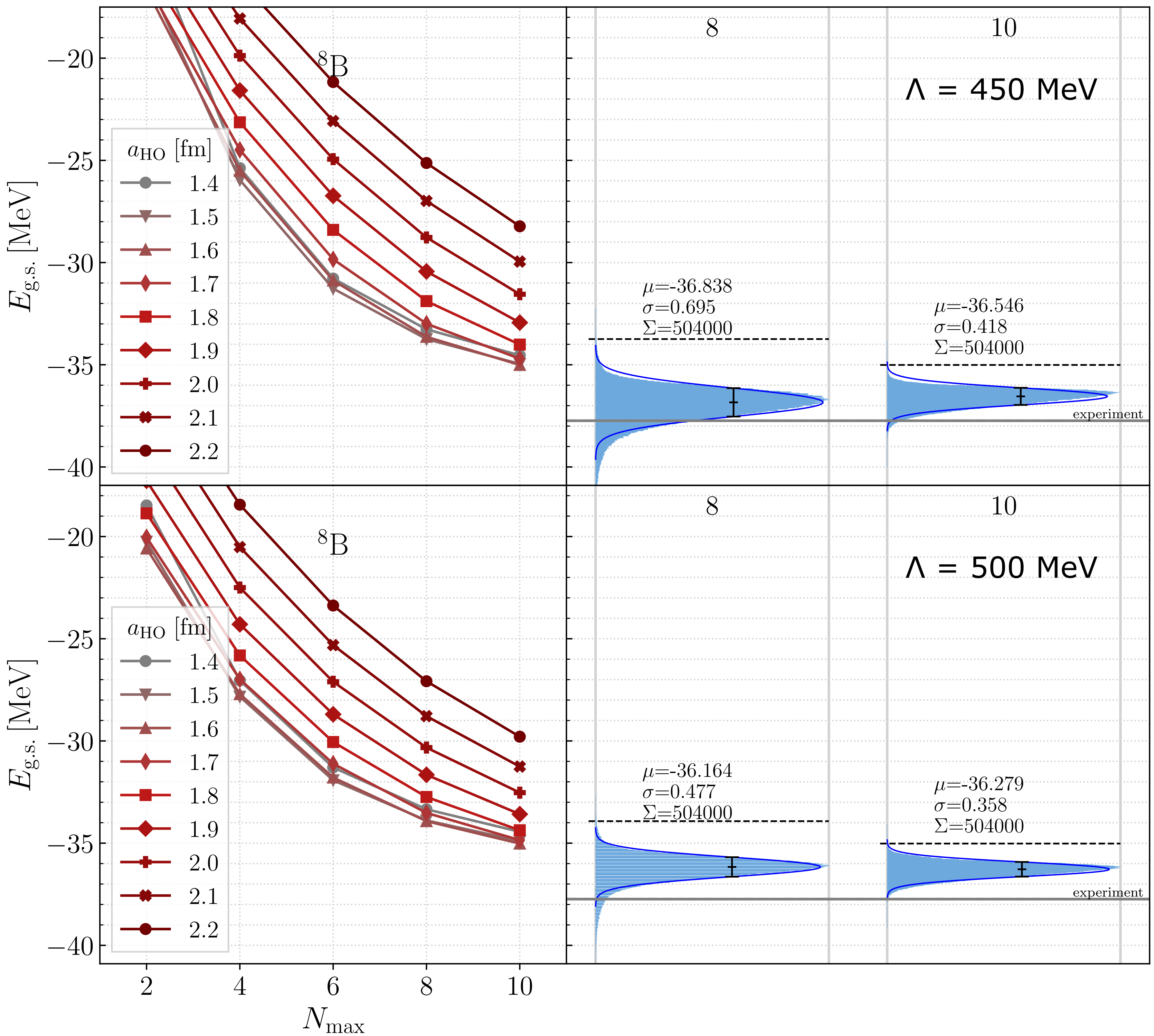
LENPIC SMSI

NN @ N4LO+
3N @ N2LO

Λ = 450/500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{8}B: Ground-State Energy



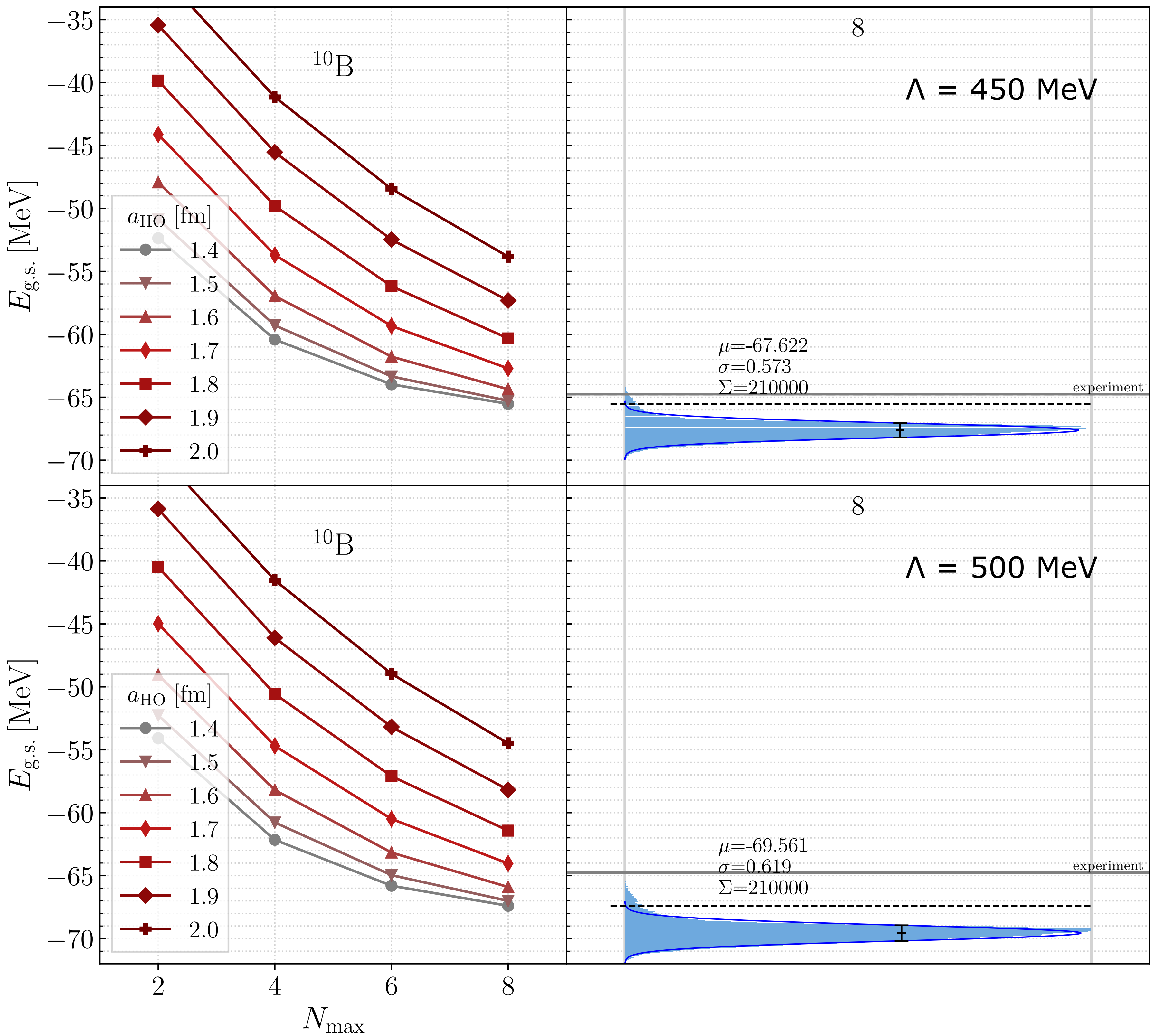
MBO Non-Local

NN @ N3LO
3N @ N3LO

Λ = 450/500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{10}B: Ground-State Energy



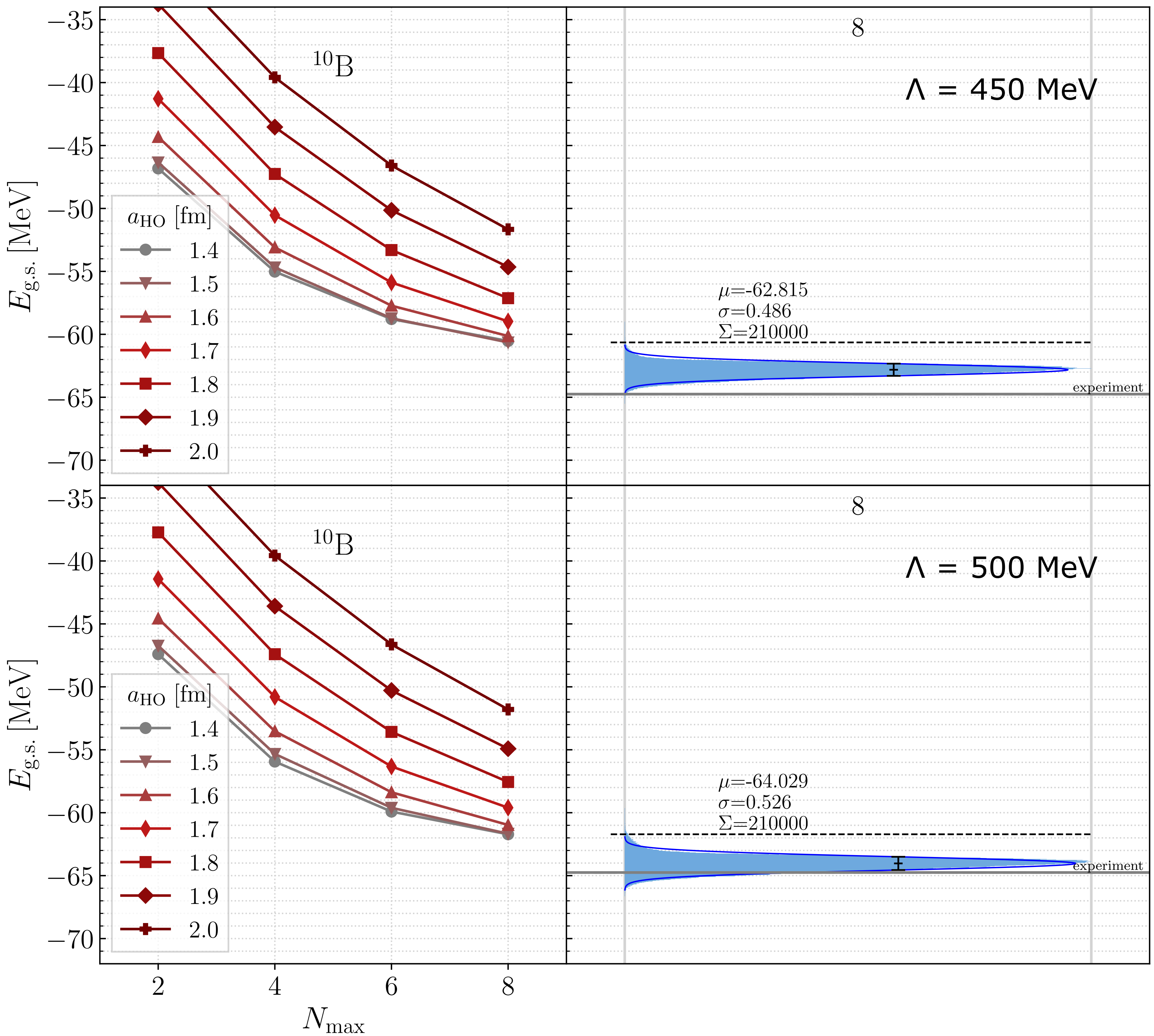
LENPIC SMSI

NN @ N2LO
3N @ N2LO

Λ = 450/500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{10}B: Ground-State Energy



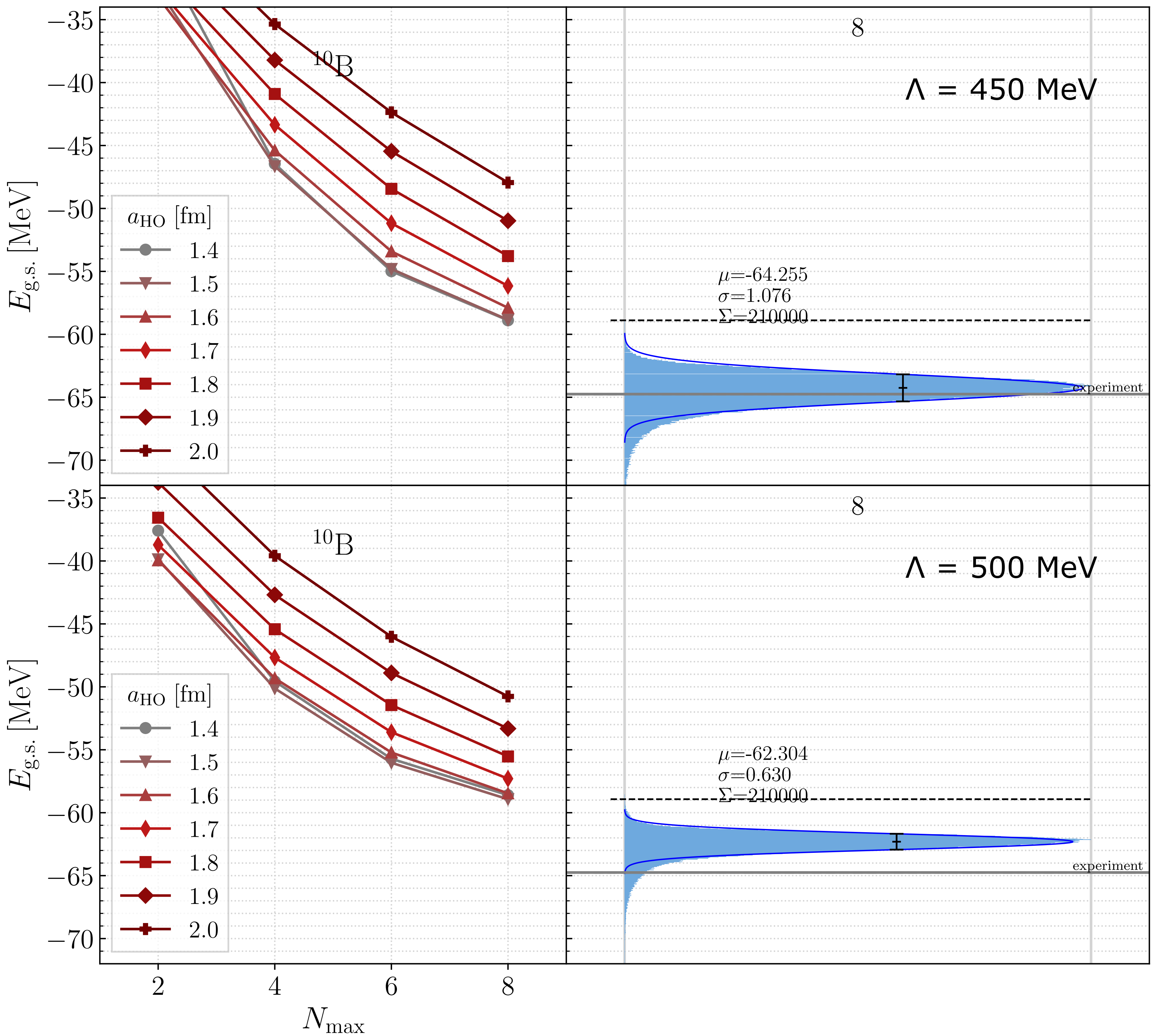
LENPIC SMSI

NN @ N4LO+
3N @ N2LO

Λ = 450/500 MeV
α = 0.08 fm^4

all NCSM runs
by Pieter Maris

^{10}B: Ground-State Energy



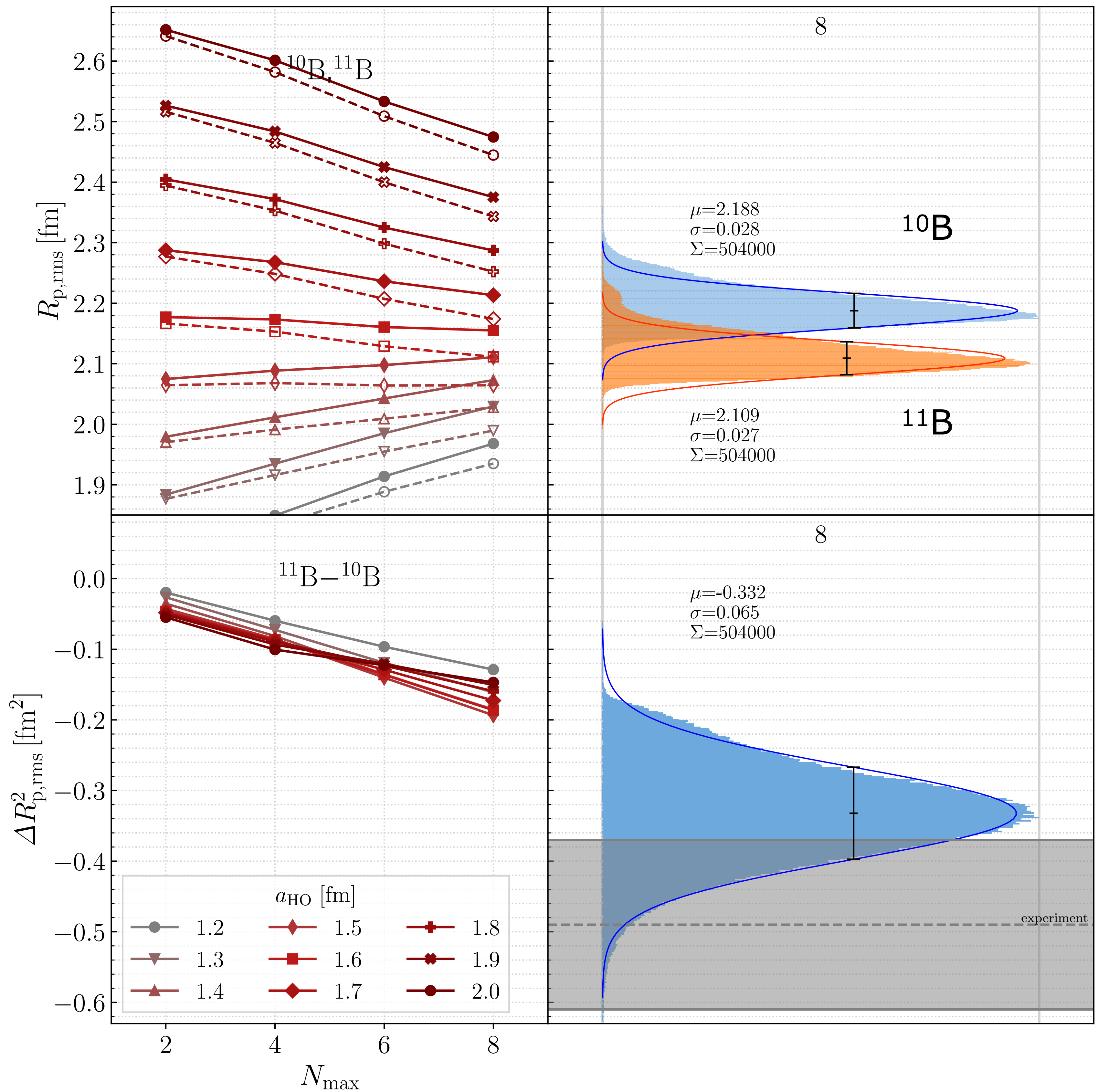
MBO Non-Local

NN @ N3LO
3N @ N3LO

Λ = 450/500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{10,11}B: Ground-State Proton Radii



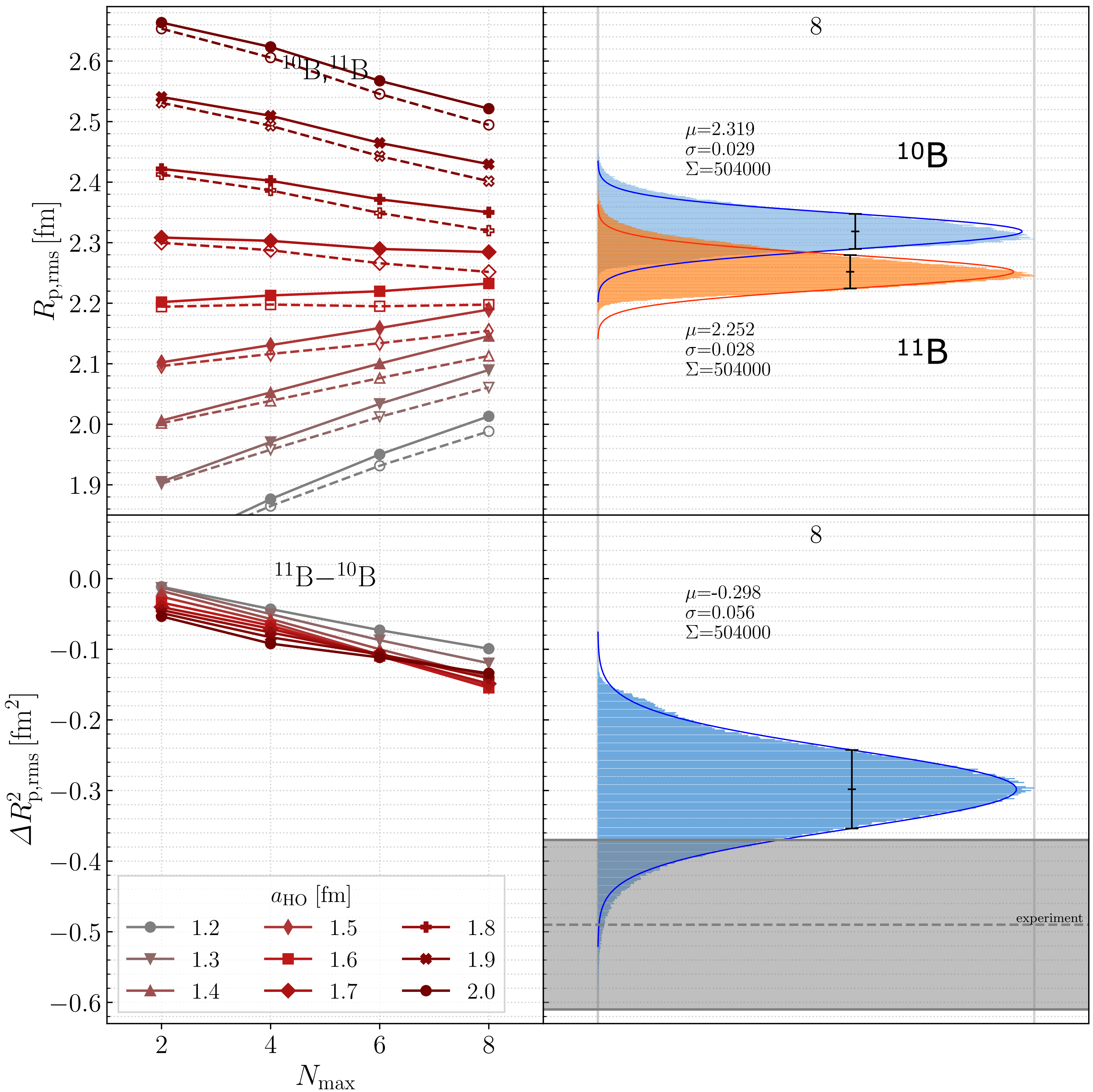
LENPIC SMSI

NN @ N2LO
3N @ N2LO

Λ = 450 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{10,11}B: Ground-State Proton Radii



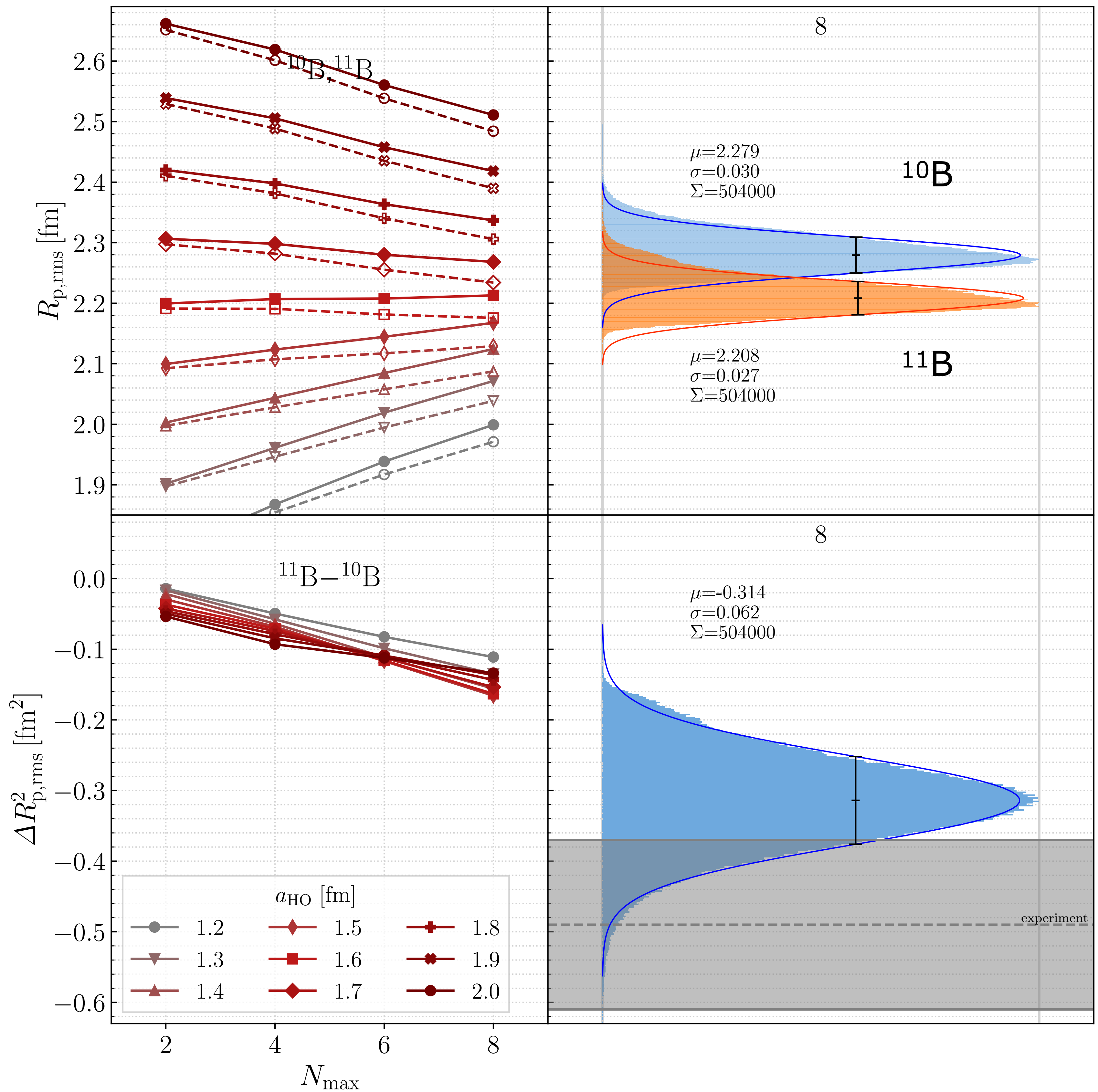
LENPIC SMSI

NN @ N4LO+
3N @ N2LO

Λ = 450 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{10,11}B: Ground-State Proton Radii



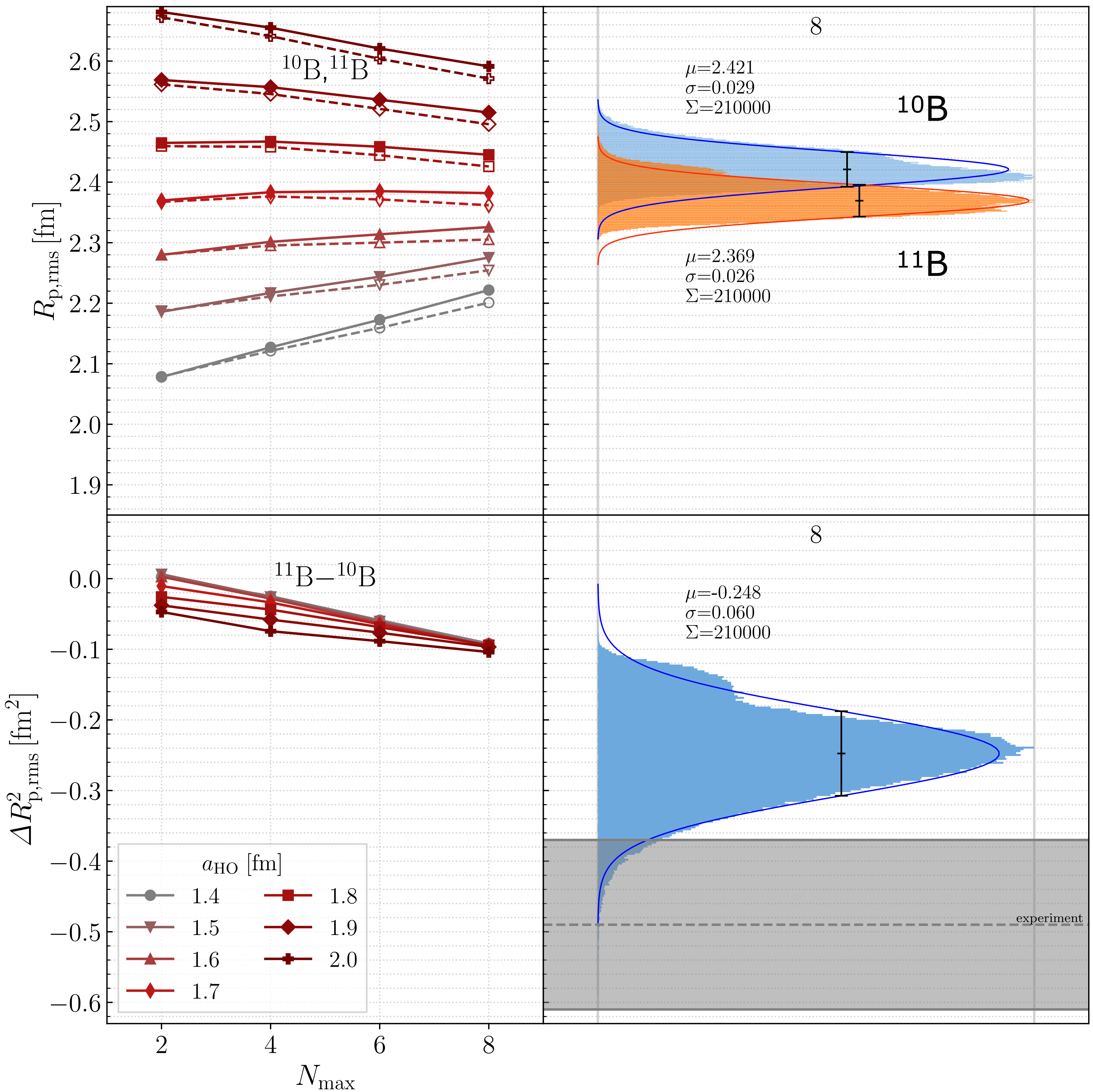
LENPIC SMSI

NN @ N4LO$^+$
3N @ N2LO

Λ = 500 MeV
α = 0.08 fm^4

all NCSM runs
by Pieter Maris

^{10,11}B: Ground-State Proton Radii



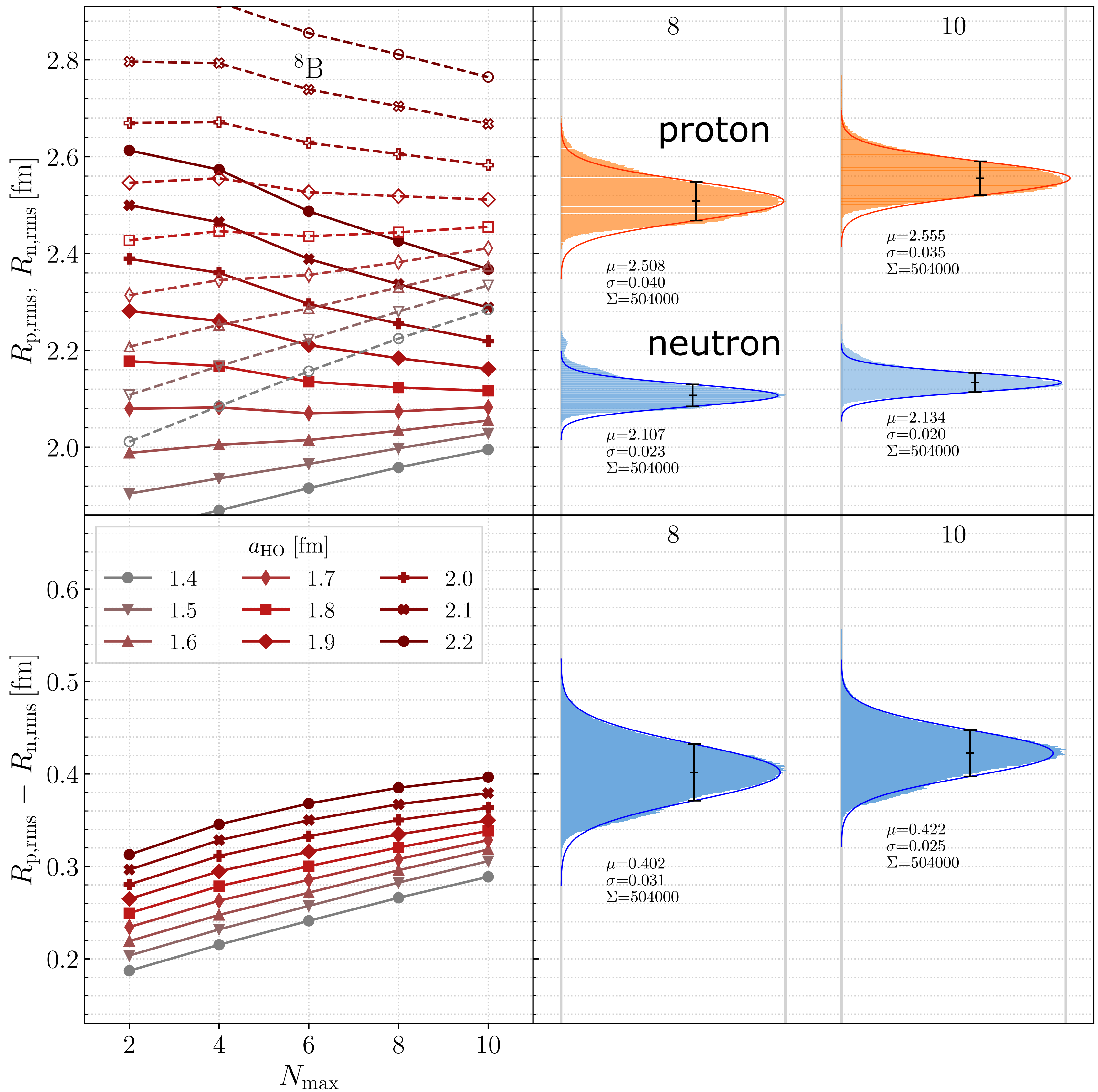
MBO Non-Local

NN @ N3LO
3N @ N3LO

Λ = 500 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{8}B: Ground-State Proton & Neutron Radii



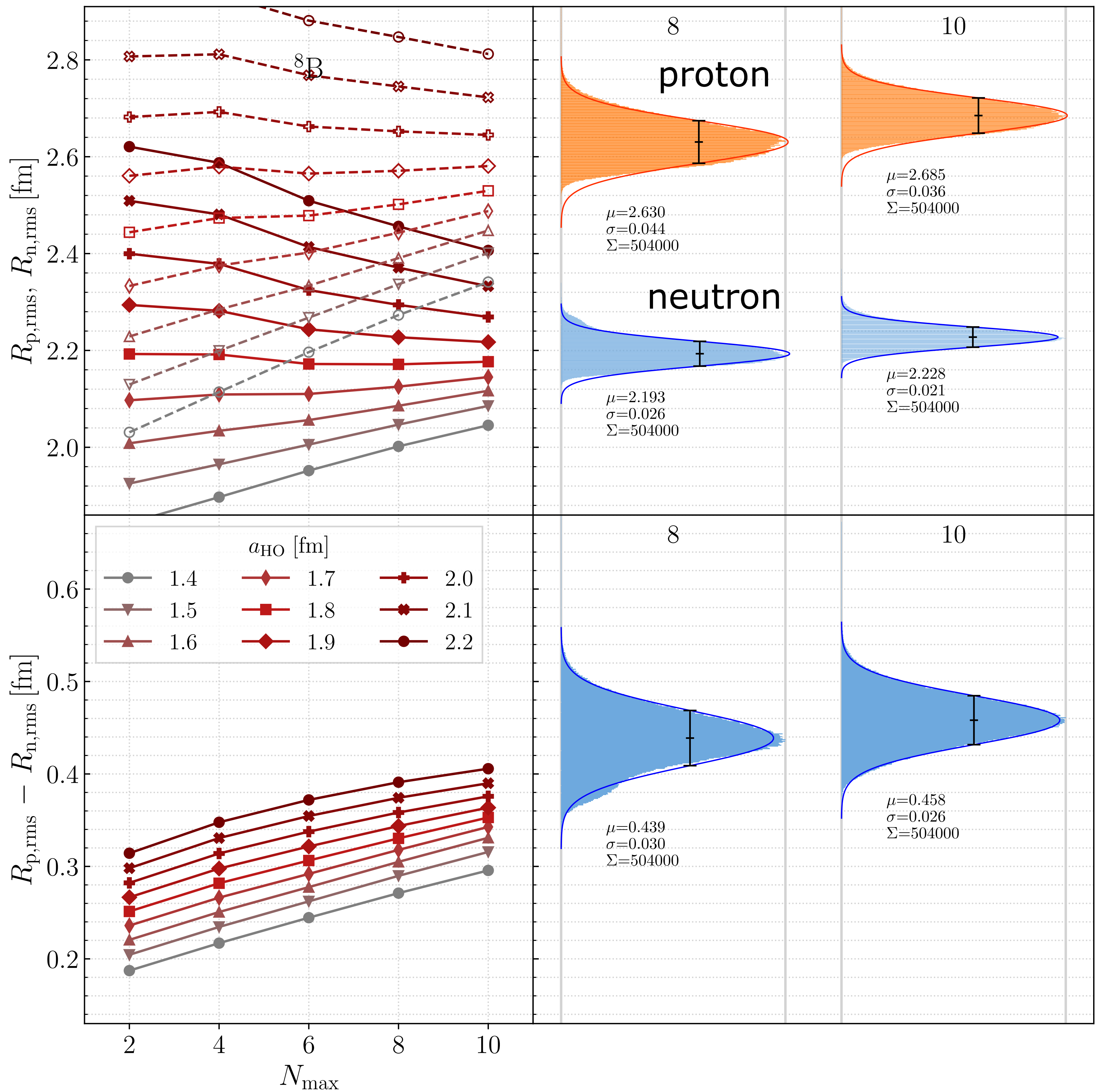
LENPIC SMSI

NN @ N2LO
3N @ N2LO

Λ = 450 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{8}B: Ground-State Proton & Neutron Radii



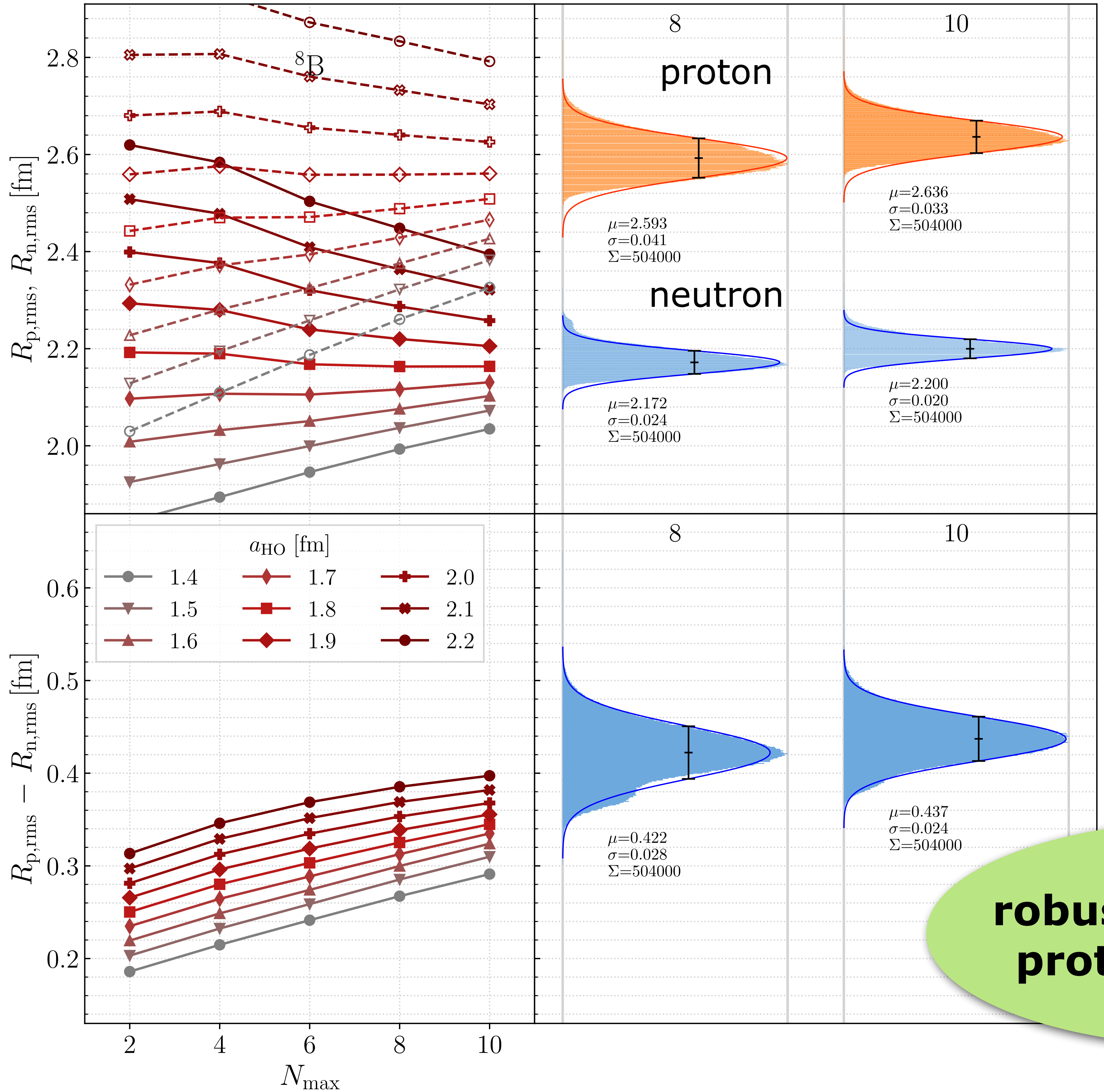
LENPIC SMSI

NN @ N4LO+
3N @ N2LO

Λ = 450 MeV
α = 0.08 fm⁴

all NCSM runs
by Pieter Maris

^{8}B: Ground-State Proton & Neutron Radii



LENPIC SMSI

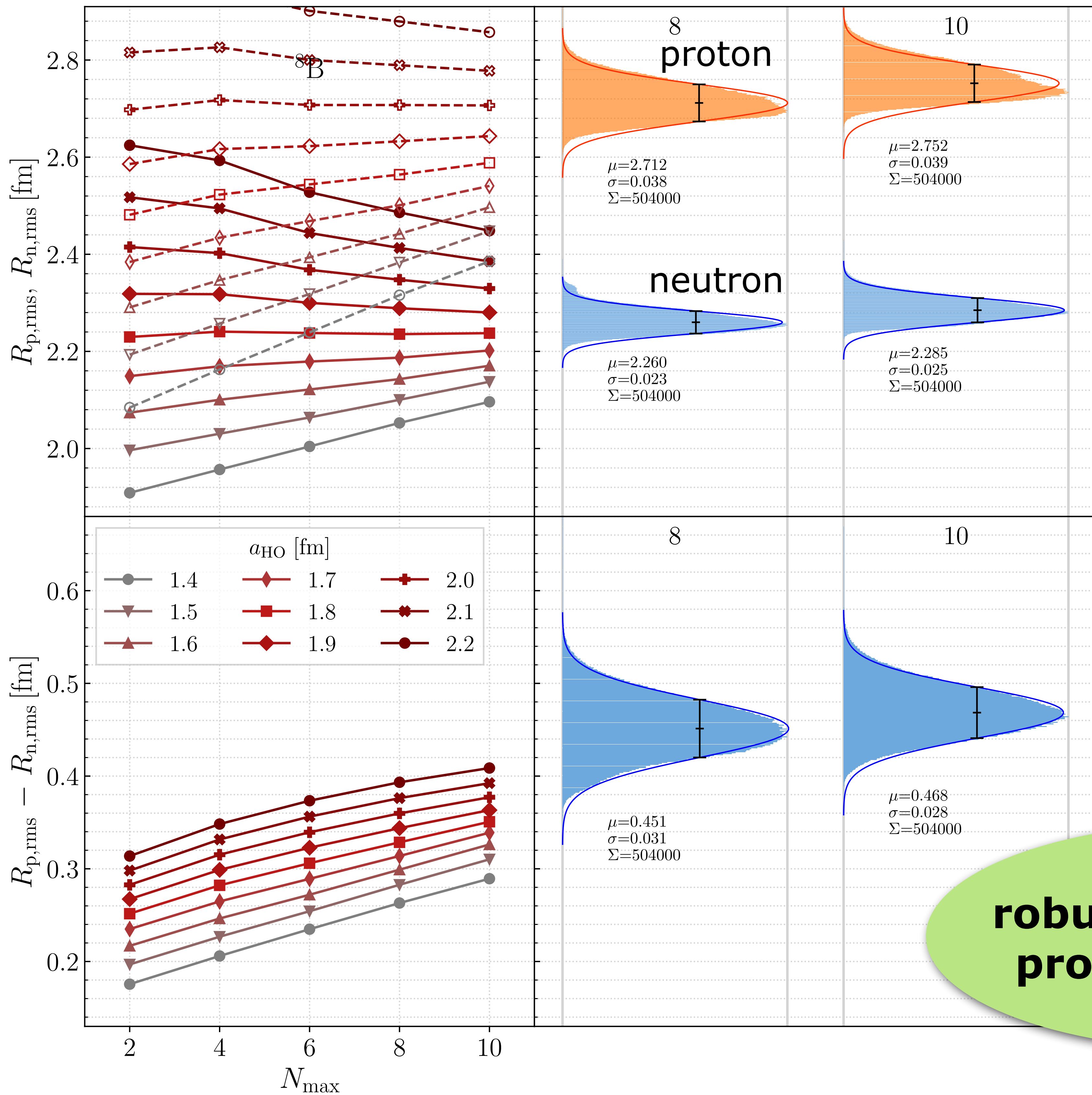
NN @ N4LO^{+}
3N @ N2LO

Λ = 500 MeV
α = 0.08 fm^{4}

all NCSM runs
by Pieter Maris

robust prediction of proton halo for ^{8}B

^{8}B: Ground-State Proton & Neutron Radii



MBO Non-Local

NN @ N3LO
3N @ N3LO

Λ = 500 MeV
α = 0.08 fm^4

all NCSM runs
by Pieter Maris

robust prediction of proton halo for ^{8}B

Work in Progress

Boron Radii

- extract two-dimensional confidence regions for pairs of observables, e.g., $R_{p,rms}(^{10}B)$ vs. $R_{p,rms}(^{11}B)$ or $R_{p,rms}(^{8}B)$ vs. $R_{n,rms}(^{8}B)$
- include chiral order-by-order uncertainty via Bayesian scheme built on ANN distributions
- extend to other observables and other halo systems

Artificial Neural Networks

- use synthetic nuclei in training process
- use flow-parameter variation as input dimension for ANN
- extension to electromagnetic observables, particularly E2 moments and transition strengths
- develop transformer networks to represent and exploit correlations among observables