TECHNISCHE UNIVERSITÄT DARMSTADT

AB INITIO THEORY AND MACHINE LEARNING FOR HYPERNUCLEI

Dem Fachbereich Physik
der Technischen Universität Darmstadt

zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation von

Marco Knöll, M.Sc.

aus Dieburg

Referent: Prof. Dr. Robert Roth
Korreferent: Prof. Achim Schwenk, Ph.D.

Darmstadt 2024

Ab Initio Theory and Machine Learning for Hypernuclei

Ab Initio Theorie und Maschinelles Lernen für Hyperkerne

Genehmigte Dissertation von Marco Knöll, M.Sc. aus Dieburg

Darmstadt, Technische Universität Darmstadt

Referent: Prof. Dr. Robert Roth

Korreferent: Prof. Achim Schwenk, Ph.D.

Tag der mündlichen Prüfung: 01.07.2024

Jahr der Veröffentlichung der Dissertation auf TUprints: 2024
URN: urn:nbn:de:tuda-tuprints-276547
URL: https://tuprints.ulb.tu-darmstadt.de/27654

Veröffentlicht unter CC BY-NC-ND 4.0 International
https://creativecommons.org/licenses/

Abstract

Systems with strangeness, such as hypernuclei, offer a unique window into the dynamics of the strong force beyond traditional nucleonic systems. Building on recent developments of the hypernuclear no-core shell model with realistic baryonic interactions from chiral effective field theory, this work presents advancements on the three major frontiers of ab initio hypernuclear structure theory: the refinement of the hyperon-nucleon interaction, precision calculations with quantified uncertainties, and the push to medium-mass hypernuclei.

Firstly, we present a novel extrapolation tool supplementing the no-core shell model that is based on artificial neural networks. Due to their pattern recognition abilities, they hold great promise in enhancing precision and uncertainty estimation in (hyper)nuclear many-body calculations, giving access to extrapolation procedures and meaningful many-body uncertainties for observables beyond ground-state energies. Moreover, they demonstrate the universality of observable-specific convergence patterns through applications across nucleonic and hyperonic systems.

Secondly, we address the chiral hyperon-nucleon interaction, which is poorly constrained due to the scarce experimental scattering data available. Based on precise many-body calculations we demonstrate the potential of ground-state and spectroscopic data for p-shell hypernuclei as additional constraints for the hyperon-nucleon interaction. We present an optimized interaction at leading order, which is then applied across a range of p-shell hypernuclei. Through only minor adjustments of the low-energy constants we alleviate the previous overbinding of the hyperon and significantly improve the agreement with experimental data in light hypernuclei.

Lastly, we present pioneering work on ab initio calculations of medium-mass hypernuclei. In a first step, we extend the concept of natural orbitals as an optimized single-particle basis to hypernuclei. We find that the associated wavefunctions yield great insight into the structure of hypernuclei, allowing us to identify ${}^{5}_{\Lambda}\mathrm{He}$ as a candidate for a hyperon halo. From there we develop a hyperon-attached in-medium no-core shell model framework and present the first ever ab initio calculation of ${}^{41}_{\Lambda}\mathrm{Ca}$.

By addressing these challenges, this work contributes to a deeper understanding of the strong force in hypernuclear systems and lays the groundwork for future investigations.

Zusammenfassung

Ab Initio Theorie und Maschinelles Lernen für Hyperkerne

Hyperkerne und andere Systeme mit Strangeness bieten einen einzigartigen Einblick in die Eigenschaften der starken Wechselwirkung jenseits herkömmlicher Atomkerne. Aufbauend auf vorangegangenen Entwicklungen des No-Core Schalenmodells für Hyperkerne zusammen mit realistischen baryonischen Wechselwirkungen aus der chiralen effektiven Feldtheorie präsentiert diese Arbeit neue Entwicklungen in drei der wichtigsten Gebiete der ab initio Hyperkernstrukturtheorie: die Modellierung der Hyperon-Nukleon-Wechselwirkung, Präzisionsrechnungen mit quantifizierten Unsicherheiten und der Vorstoß zu schwereren Hyperkernen.

Zunächst präsentieren wir eine neuartige Extrapolationsmethode für Rechnungen mit dem No-Core Schalenmodell basierend auf künstlichen neuronalen Netzwerken. Dank ihrer Qualitäten im Bereich der Mustererkennung sind Methoden basierend auf maschinellem Lernen vielversprechende Ansätze um sowohl die Präzison von Vielteilchenmethoden als auch die Quantifizierung von Unsicherheiten zu verbessern. Insbesondere ermöglichen sie die Extrapolation von anderen Observablen als der Grundzustandsenergie und erlauben die Abschätzung realistischer Unsicherheiten. Darüber hinaus zeigen diese Methoden, dass das Konvergenzverhalten von Observablen über verschiedene Kerne und Hyperkerne hinweg universell ist.

Danach wenden wir uns der Hyperon-Nukleon Wechselwirkung zu, deren freie Parameter aufgrund der knappen Verfügbarkeit von experimentellen Streudaten nur unzureichend genau bestimmt werden können. Basierend auf präzisen Vielteilchenrechnungen demonstrieren wir das Potential von grundzustands und spektroskopischen Daten für Hyperkerne in der p-Schale als zusätzliche Randbedingung für die Hyperon-Nukleon-Wechselwirkung und präsentieren eine optimierte Wechselwirkung in erster chiraler Ordnung, die anschließend für die Untersuchung einer Auswahl an leichten Hyperkernen verwendet wird. Mit nur kleinen Anpassungen der Niedrigenergiekonstanten erreichen wir so eine Reduktion der vorherigen Überbindung des Hyperons, was zu einer deutlich besseren Übereinstimmung mit den experimentellen Daten insbesondere in leichten Hyperkernen führt.

Abschließend präsentieren wir wegweisende Entwicklungen hin zu einer ab initio Beschreibung mittelschwerer Hyperkerne. Dafür übertragen wir zunächst das Konzept der natürlichen Orbitale als optimierte Einteilchenbasis auf Hyperkerne. Die Wellenfunktionen der natürlichen Orbitale eigenen sich hervorragend als Diagnosewerkzeug für strukturelle Eigenschaften von Hyperkernen, was uns ermöglicht ${}^{5}_{\Lambda}\mathrm{He}$ als potentiellen Kandidaten für einen Hyperon-Halo zu identifizieren. Darauf aufbauend entwickeln wir ein Framework für das Hyperon-Attached In-Medium No-Core Schalenmodell und präsentieren erstmals ab initio Berechnungen von ${}^{41}_{\Lambda}\mathrm{Ca}$.

Durch diese Entwicklungen trägt diese Arbeit zu einem tieferen Verständnis der starken Wechselwirkung in Hyperkernen bei und bereitet den Weg für zukünftige Entwicklungen.

Contents

Introduction 1

I The Nuclear Many-Body Problem 5

1 The Nuclear Hamiltonian 7
- 1.1 Chiral Effective Field Theory 8
- 1.2 Similarity Renormalization Group 10

2 The No-Core Shell Model 13
- 2.1 Model Space and Symmetries 14
- 2.2 Importance Truncation 17

3 Model-Space Extrapolation and Uncertainty Quantification 19
- 3.1 Interaction Uncertainties 19
- 3.2 Many-Body Uncertainties 22

II Artificial Neural Networks for the No-Core Shell Model 25

4 Machine Learning 27
- 4.1 Basic Concepts 27
- 4.2 Artificial Neural Networks 29

5 Artificial Neural Networks for the NCSM 33
- 5.1 Network Design 34
- 5.2 Training Data 38
- 5.3 Training 42
- 5.4 Statistical Evaluation 45

6 Predictions of Energies and Spectra 47
- 6.1 Benchmark in Few-Body Systems 47
- 6.2 Application to p-Shell Nuclei 50
- 6.3 Extension to Excitation Energies 54

7 Predictions of Radii 57
7.1 Benchmark in Few-Body Systems 57
7.2 Application to p-Shell Nuclei 59

III Ab Initio Theory for p-Shell Hypernuclei 61

8 The Hypernuclear Hamiltonian 63
8.1 Baryonic Interactions 64

9 Hypernuclei in the No-Core Shell Model 67
9.1 Revisiting Artificial Neural Networks 69

10 Hyperon-Nucleon Interaction Constrained on p-Shell Hypernuclei 73
10.1 The Leading-Order YN Interaction 74
10.2 LEC Sensitivity Analysis 76
10.3 Optimization Procedure 77
10.4 Results for p-Shell Hypernuclei 81

11 Natural Orbitals for Hypernuclei 87
11.1 Conceptual Basics 87
11.2 The Hartree-Fock Method 89
11.3 Many-Body Perturbation Theory 93
11.4 Hypernuclear Natural Orbitals 95
11.5 Natural Orbitals in the Hypernuclear NCSM 103

IV Towards Ab Initio Theory for Medium-Mass Hypernuclei 107

12 In-Medium No-Core Shell Model 109
12.1 Normal Ordering 109
12.2 In-Medium Similarity Renormalization Group 111
12.3 In-Medium No-Core Shell Model 115
12.4 Particle-Attached/-Removed Extension 116

13 Hypernuclei in the IM-NCSM 117
13.1 Hyperon-Attached IM-NCSM 117
13.2 Application to Medium-Mass Hypernuclei 121

Summary and Outlook 123

Appendices **127**

A Numerical Values for ANN Predictions 129
A.1 Ground-State Energies and Spectra 129
A.2 Radii 130

B ANN Predictions for Hypernuclear Observables with YN_{opt} 131
B.1 Predictions for Hyperon Separation Energies 131
B.2 Numerical Values 134

Bibliography 139

Introduction

Theoretical nuclear physics is the branch of physics that aims for an understanding of atomic nuclei. In particular, nuclear structure theory is concerned with the description of the properties of such nuclei, which are self-bound systems that emerge from the strong interaction. Being one of the three fundamental forces governed by the standard model of particle physics, the strong interaction itself is captured in the theory of quantum chromodynamics (QCD). It describes the interaction between quarks and gluons as fundamental particles that make up baryons and mesons including nucleons as the constituents of atomic nuclei. However, these particles exhibit confinement giving QCD its non-perturbative character in the low-energy regime. This is what what poses a major challenge for direct calculations of nuclei or the interaction between their constituents from QCD.

Throughout the past decades, chiral effective field theory (EFT) has become the most successful approach for bridging the gap between QCD and nuclear theory [1, 2]. As the name suggests, it allows to construct an effective interaction on a low-energy scale featuring nucleons and pions as effective degrees of freedom, while being rooted in the fundamental symmetries of QCD. It resembles a low-momentum expansion for the nuclear interaction that comes with free parameters, which need to be fit to experimental scattering data.

In contrast to e.g. the electromagnetic interaction, the strong interaction features a strong repulsive short-range behavior along with strong tensor interactions, making calculations of systems that are bound by the strong force extremely challenging. Here, unitary transformations of the interaction within the similarity renormalization group (SRG) have proven to be a powerful asset that allows to soften the interaction [3–5]. Therefore, they have become a standard tool complementing interactions from chiral EFT.

Naturally, nuclear structure theory has mostly been concerned with nucleonic systems as they make up the matter around us. However, the standard model, and with it QCD, features a total of six quark flavors that give rise to the formation of baryons and mesons beyond nucleons and pions, which consist of up and down quarks only. Substituting one of the quarks in nucleons with a strange quark yields a class of particles called hyperons, in particular, the Λ and Σ baryons. According to their quark content they are assigned with a, historically negative, strangeness quantum number $\mathcal{S} = -1$. Introducing a second strange quark results in the Ξ baryons with $\mathcal{S} = -2$ respectively. Correspondingly, kaons arise as the strange counterparts of the pions. Due to the short lifetime of hyperons in the order of 10^{-10} s they rapidly decay into nucleons and pions. However, if created in the presence of other fermions such as nucleons they can form bound systems known as hypernuclei. Since electromagnetic and strong processes occur on even shorter timescales, hypernuclei are well-defined and measurable systems, yielding great potential to deepen our understanding of the strong force beyond purely nucleonic systems.

Experimentally, hypernuclei have first been observed in the early 1950s created by cosmic rays in an emulsion stack [6]. Their existence has later been confirmed in other emulsion

experiments [7, 8] and, with the rise of particle accelerators, hypernuclei have been produced with kaon, pion, or electron beams, where they were measured with, e.g., high resolution spectrometers [9–12]. Still, measuring hypernuclei with good precision remains challenging due to their short lifetime. Nevertheless, there has been an increasing interest in hypernuclei very recently, which has led to a wave of proposals for new hypernuclear experiments in the near future [13–15].

Besides nuclear and particle physics this interest is also driven by the field of nuclear astrophysics, where the potential existence of hyperons in the extreme conditions inside neutron stars has led to the so-called hyperon puzzle [16–18] whose solution requires a more general understanding of the strong interaction. From a theoretical point of view this calls for models for the interaction between hyperons and nucleons or amongst hyperons themselves. The framework of chiral EFT has successfully been extended to include hyperons [19–22]. Though, the determination of the free parameters in the hyperonic channels is far more difficult due to a lack of hyperon scattering data. Therefore, hypernuclei are the most promising candidates for providing a deeper insight into the strange aspects of the strong interaction.

This leads us back to nuclear structure theory or, more precisely, hypernuclear structure theory. In the nucleonic regime, a variety of ab initio methods such as the no-core shell model (NCSM) [23–26], the coupled-cluster (CC) method [27], the self-consistent Green's function (SCGF) approach [28], the in-medium similarity renormalization group (IM-SRG) [29, 30], and quantum Monte Carlo (QMC) methods [31], as well as hybrid methods like the in-medium no-core shell model (IM-NCSM) [32] have become state-of-the-art and successfully operate in different different areas of the nuclear chart. They all share essential properties encapsulated in the term "ab initio", in particular controlled and quantified uncertainties and systematic improvability. Together with increasingly sophisticated uncertainty estimates based on Bayesian statistics [33–35], nuclear structure physics has entered a precision era.

The picture in the hypernuclear regime, however, is somewhat different as the inclusion of hyperons requires major extensions of the established many-body methods. Until recently, ab initio calculations for hypernuclei were only feasible for few-body systems via methods like Faddeev-Yakubovski (FY) [36, 37], Gaussian expansion methods (GEM) [38], or QMC approaches [39, 40]. But lately, generalized NCSM implementations and nuclear lattice EFT methods have been developed, that give access to observables in (singly-strange) p-shell hypernuclei [41–45]. These observables, in turn, foster our understanding of the strong interaction. However, besides these exciting advances, ab initio hypernuclear structure theory still exhibits deficiencies and limitations that essentially yield three major challenges that we will explore in this work:

The first and most difficult challenge lies with the construction of the interaction. The hypernuclear Hamiltonian features terms that account for interactions of hyperons with nucleons and other hyperons. For the hyperon-nucleon interaction, these terms have been worked out up to the fourth order of the chiral expansion [22], however, the emerging free parameters are poorly constrained due to a lack of hyperon-nucleon scattering data. This contributes to deficiencies like a systematic overbinding of the hyperons in hypernuclei [42]. Moreover, all these interactions have so far been limited to two-body forces, though different studies indicate that the inclusion of three-body forces might play an important role in the description of hypernuclei and neutron stars [18]. Obtaining additional scattering data is challenging due to the short lifetime of hyperons. Hence, other means have to be developed to better constrain hypernuclear interactions.

The strive for precision calculations of hypernuclei poses a second challenge that is tightly connected to uncertainty quantification. While ab initio methods feature controlled uncertainties by construction, accurately quantifying them remains a major problem in, both, nuclear and hypernuclear theory. Consequently, uncertainty quantification has been the focus of various developments in the past decade [34, 46–49]. For most ab initio many-body methods, in particular configuration interaction (CI) methods like the NCSM, the main uncertainty arises from the limited model-space size that can be handled. Even with high-performance computing (HPC) facilities, which have enabled the development of large-scale numerical ab initio methods in the first place, exact results can only be obtained for the smallest of nuclei. This has opened up the question on how to extract meaningful predictions of observables along with reasonable uncertainties from calculations in finite model spaces. Such postprocessing procedures are mostly based on phenomenological and semi-empirical schemes [50–53]. This is starting to change as the rise of machine learning (ML) has also impacted nuclear theory [54, 55]. Being an ideal and very adaptable tool for data processing it yields promising Ansätze for procedures supplemental to ab initio many-body calculations [56, 57]. Hence, ML methods carry the potential to significantly improve the precision of (hyper)nuclear many-body calculations, which is not only essential for a detailed understanding of the strong interaction but also plays a crucial role in the design of future experiments.

The third and final challenge is the everlasting push to heavier systems. As of now, ab initio methods for hypernuclei can only reach up to p-shell systems, after which the capabilities of the NCSM are exhausted. While calculations of medium-mass hypernuclei have been performed with mean-field and perturbative approaches like many-body perturbation theory (MBPT) and Brueckner-Hartree-Fock (BHF) [42, 58], ab initio calculations in this mass regime are highly desirable. Since the ab initio description of nuclei has made tremendous progress in recent years reaching all the way up to $^{208}\mathrm{Pb}$ and pushing to exotic and deformed systems [59–63], these methods provide an ideal starting point for extensions to hypernuclei.

The goal of this work is to address these challenges one by one, to investigate and discuss them in detail, and to present advances on all three frontiers.

This thesis is structured as follows: The first part is focused on the basic premise of the nuclear many-body problem, including the modeling of the Hamiltonian and the NCSM as our many-body method of choice. Next, we introduce the basic concepts of ML and artificial neural networks in part two and discuss their potential as a supplemental tool for the NCSM. Part three is then focused on ab initio theory for p-shell hypernuclei, where we explicitly address the problems of the chiral hyperon-nucleon interaction and demonstrate a way to improve it based on p-shell hypernuclear structure data. Finally, in part four we discuss how ab initio theory for hypernuclei can be extended to the medium-mass regime and present the first ab initio calculations of medium-mass hypernuclei in the IM-NCSM framework.

Part I

The Nuclear Many-Body Problem

1

The Nuclear Hamiltonian

At the core of nuclear structure theory is the non-relativistic nuclear many-body problem given by the stationary Schrödinger equation

$$\hat{H}|\Psi_n\rangle = E_n|\Psi_n\rangle \tag{1.1}$$

with the nuclear Hamiltonian $\hat{H}$, its eigenvalues E_n, and corresponding eigenstates $|\Psi_n\rangle$. The difficulty of solving this equation strongly depends on the complexity of the employed Hamiltonian. Since nuclei are self-bound systems we consider a translationally invariant intrinsic Hamiltonian

$$\hat{H}_{\text{int}} = \hat{T} - \hat{T}_{\text{cm}} + \hat{V} = \hat{T}_{\text{int}} + \hat{V} \tag{1.2}$$

which, in general, consists of one term for the intrinsic kinetic energy $\hat{T}_{\text{int}} = \hat{T} - \hat{T}_{\text{cm}}$ and one for the potential $\hat{V}$, where the latter takes account for the interactions among nucleons. The operator for the intrinsic kinetic energy of an A-body system can be written down explicitly as

$$\hat{T}_{\text{int}} = \hat{T} - \hat{T}_{\text{cm}} = \sum_{i=1}^{A} \frac{\hat{p}_i^2}{2\hat{m}_i} - \frac{\hat{P}_{\text{cm}}^2}{2\hat{M}}, \tag{1.3}$$

where $\hat{m}_i$ and $\hat{\mathbf{p}}_i$ are the mass and momentum operators for the i-th particle and $\hat{M}$ and $\hat{\mathbf{P}}_{\text{cm}}$ are the total mass and the center-of-mass momentum which can again be expressed in terms of single-particle operators

$$\hat{M} = \sum_{i=1}^{A} \hat{m}_i\,, \quad \hat{\mathbf{P}}_{\text{cm}} = \sum_{i=1}^{A} \hat{\mathbf{p}}_i. \tag{1.4}$$

Note that we intentionally keep track of the masses of the individual particles via the introduced mass operators. While this is often disregarded for nucleonic systems under the assumption of equal nucleon masses $m_{\text{N}} \equiv m_p = m_n$, we choose a more general formulation here in foresight of the extension to hypernuclei in part III, where we need to account for the different masses of nucleons and hyperons explicitly.

Conveniently, the kinetic energy can be rewritten in terms of relative momenta between two particles $\hat{\mathbf{q}}_{ij} = \hat{\mu}_{ij}(\hat{\mathbf{p}}_j/\hat{m}_j - \hat{\mathbf{p}}_i/\hat{m}_i)$ and the reduced mass $\hat{\mu}_{ij} = (\hat{m}_i\hat{m}_j)/(\hat{m}_i + \hat{m}_j)$

yielding

$$\begin{aligned}\hat{T}_{\text{int}} &= \sum_{i=1}^{A} \frac{\hat{p}_i^2}{2\hat{m}_i} - \frac{1}{2\hat{M}} \sum_{i,j=1}^{A} \hat{m}_i \hat{m}_j \frac{\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j}{\hat{m}_i \hat{m}_j} \\ &= \sum_{i=1}^{A} \frac{\hat{p}_i^2}{2\hat{m}_i} - \frac{1}{2} \sum_{i=1}^{A} \left(\frac{\hat{p}_i^2}{2\hat{m}_i} + \frac{\hat{p}_i^2}{2\hat{m}_i} \right) + \frac{1}{2\hat{M}} \sum_{i,j=1}^{A} \hat{m}_i \hat{m}_j \frac{\hat{q}_{ij}^2}{2\hat{\mu}_{ij}} \\ &= \sum_{i<j}^{A} \frac{\hat{m}_i + \hat{m}_j}{\hat{M}} \frac{\hat{q}_{ij}^2}{2\hat{\mu}_{ij}},\end{aligned} \tag{1.5}$$

where we have used

$$\frac{\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j}{\hat{m}_i \hat{m}_j} = \frac{\hat{p}_i^2}{2\hat{m}_i^2} + \frac{\hat{p}_j^2}{2\hat{m}_j^2} - \frac{1}{2} \left(\frac{\hat{p}_j^2}{\hat{m}_j} - \frac{\hat{p}_i^2}{\hat{m}_i} \right)^2 = \frac{\hat{p}_i^2}{2\hat{m}_i^2} + \frac{\hat{p}_j^2}{2\hat{m}_j^2} - \frac{\hat{q}_{ij}^2}{2\hat{\mu}_{ij}^2}. \tag{1.6}$$

For nucleonic systems the assumption of equal nucleon masses holds and the kinetic energy can further be simplified to

$$\hat{T}_{\text{int}} = \frac{2}{A} \sum_{i<j}^{A} \frac{\hat{q}_{ij}^2}{m_{\text{N}}}. \tag{1.7}$$

Constructing the nuclear potential $\hat{V}$ poses a far more complicated challenge as the strong force, which describes the interaction between nucleons, is not yet fully understood. In order to allow for a more conceptual approach it is commonly decomposed into terms sorted by particle rank

$$\hat{V} = \hat{V}_{\text{NN}} + \hat{V}_{\text{3N}} + \ldots + \hat{V}_{\text{AN}}. \tag{1.8}$$

Due to computational limitations and limited availability of interactions it is commonly truncated at the three-body level. However, contributions from higher particle ranks are expected to be rather small [64] [1]. While there have been successful phenomenological parametrizations of nuclear interactions based on symmetry considerations [67–69], most ab initio methods nowadays resort to interactions constructed from chiral EFT.

1.1 Chiral Effective Field Theory

Nucleons or, more general, baryons consist of quarks and gluons. Hence, their interaction is governed by QCD. In particular, it is the low-energy regime of QCD that is of interest for nuclear structure theory. In this regime quarks and gluons exhibit confinement, thus, preventing a perturbative description of the interaction in terms of its fundamental particles.

Chiral EFT presents an approach to construct baryon-baryon interactions from the most general Lagrangian that obeys the symmetries of QCD. It takes advantage of the comparably

[1] Note that the exclusion of higher-body forces results in major deficits in the description of some specific cases such as the Hoyle state in ^{12}C, which is understood as an α-clustering phenomenon and, therefore, dominated by many-body effects beyond the three-body level [65, 66].

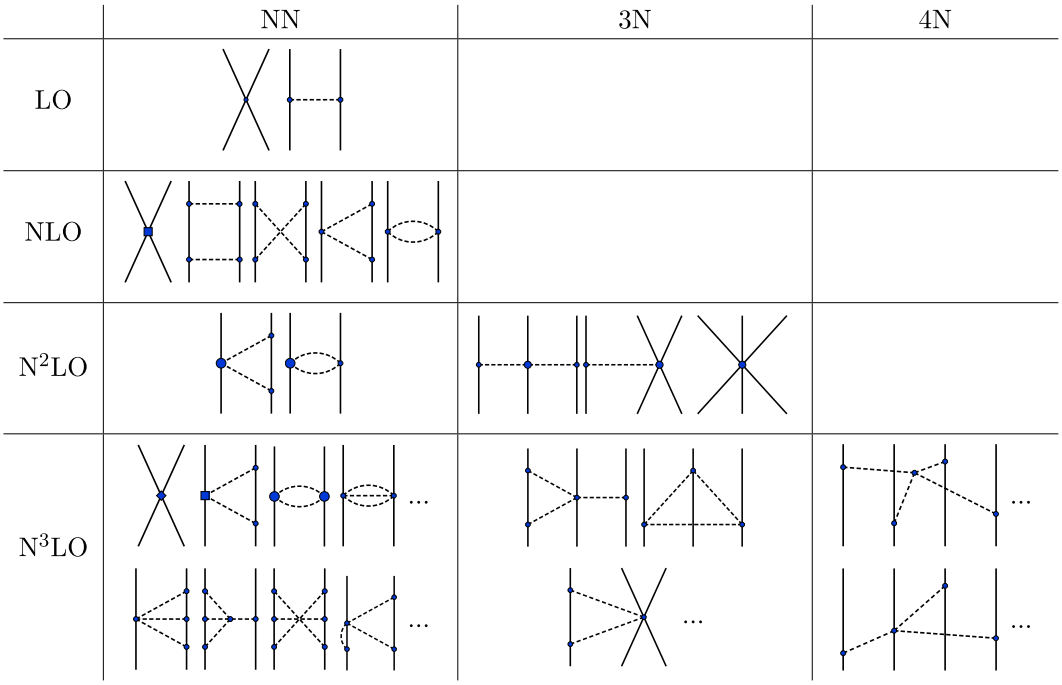

Figure 1.1: Diagrammatic representation of the Weinberg power counting in chiral EFT with multi-nucleon forces emerging at different chiral orders.

large gap between the pion mass and heavier mesons, featuring only nucleons and pions as active degrees of freedom. Since a detailed discussion of chiral EFT would go far beyond the scope of this work and has been extensively provided in other work, we refer to [1, 2, 70, 71] for further reading and focus on the conceptual characteristics of chiral EFT that are of major importance for nuclear structure calculations.

First and foremost, chiral EFT features an expansion around a small parameter $Q = P/\Lambda_\chi$ where P is the typical momentum scale of a nucleus and Λ_χ is the so-called breakdown scale that ensures the separation of the aforementioned mass regimes. The individual terms of the interaction can then be organized in powers of Q giving rise to a power counting scheme as depicted in Fig. 1.1, which was first introduced by Weinberg [72–75]. This power counting not only leads to a gradual emergence of multi-nucleon forces, which aligns with the previously mentioned decomposition of the nuclear potential, but, more importantly, allows for systematic order-by-order improvement of the interaction – a key aspect of ab initio theory. The latter further enables the estimation of uncertainties that arise from the truncation of the chiral expansion, which we will come back to later. Looking at the individual orders we find that the leading order (LO) and next-to-leading order (NLO) contributions consist of nucleon-nucleon (NN) diagrams only while three-nucleon (3N) and 4N forces enter at next-to-next-to-leading order (N^2LO) and next-to-next-to-next-to-leading order (N^3LO) respectively.

When working with these potentials one encounters infinities. These can be avoided through different regularization schemes, most commonly through cutoff functions in momentum space, coordinate space, or both. These functions additionally require the definition of a cutoff scale $\Lambda_C < \Lambda_\chi$. Finally, being an effective theory, free parameters called low-energy constants (LECs) enter at every chiral order. These parameters, which correspond to the occurring contact terms, have to be constrained through fits to experimental data or by other means.

Apparently, the construction of an interaction within the chiral EFT framework is certainly not unique. Even if we limit ourselves to a Weinberg power counting with a fixed breakdown scale, the choice of regularization scheme gives rise to what is considered a family of interactions. Within these families interactions can still differ by the chiral order, regularization cutoff, and the fitting strategy for the LECs. Moreover, one could raise aspects such as the inclusion of Δ particles as additional degrees of freedom. The works in [52, 76–84] resemble an incomplete list of the widely accepted families used in modern calculations.

As of today, the chiral expansion in a nucleonic setting has been worked out up to N^{4}LO and beyond in the NN sector [85] and up to N^{3}LO for 3N and 4N forces [78]. However, this is not the case for all families of interactions.

1.2 Similarity Renormalization Group

A pronounced feature of realistic nuclear Hamiltonians are a strong short-range repulsion and strong tensor forces. On the level of matrix elements, these appear as strong couplings of high- and low-momentum states. When solving the nuclear many-body problem for such a Hamiltonian these features become the bottleneck for most many-body methods as they induce correlations in the many-body states, which are computationally costly as very large model spaces are required to capture them. Converged calculations are, therefore, not feasible for systems beyond the few-body sector.

In order to address this issue, the Hamiltonian can be preconditioned such that the convergence of observables w.r.t. model-space size is accelerated, while their values remain unchanged. This can be achieved through unitary transformations, which guarantee the invariance of the observables. Thus, for most modern nuclear structure calculations, the Hamiltonian is pre-processed via a SRG transformation [86–89], i.e., a continuous unitary transformation

$$\hat{H}(\alpha) = \hat{U}^\dagger(\alpha)\hat{H}\hat{U}(\alpha) \tag{1.9}$$

parametrized through a so-called flow parameter α with $\hat{U}(0) = \hat{1}$. Thanks to the unitarity $\hat{U}(\alpha)\hat{U}^\dagger(\alpha) = \hat{1}$ of the transformation it can easily be shown that the eigenvalues of the Hamiltonian and, thus, of the many-body system of interest remain unchanged

$$\begin{aligned} & \hat{H}|\Psi_n\rangle = E_n|\Psi_n\rangle \\ \Leftrightarrow \quad & \hat{U}^\dagger(\alpha)\hat{H}\hat{U}(\alpha)\hat{U}^\dagger(\alpha)|\Psi_n\rangle = \hat{U}^\dagger(\alpha)E_n|\Psi_n\rangle \\ \Leftrightarrow \quad & \hat{H}(\alpha)|\tilde{\Psi}_n\rangle = E_n|\tilde{\Psi}_n\rangle \end{aligned} \tag{1.10}$$

It is important to note that the corresponding eigenstates do change under the transformation $|\Psi_n\rangle \to |\tilde{\Psi}_n\rangle \equiv \hat{U}(\alpha)|\Psi_n\rangle$. As a consequence, operators for observables other than energy need to be transformed consistently

$$\hat{O}(\alpha) = \hat{U}^\dagger(\alpha)\hat{O}\hat{U}(\alpha) \tag{1.11}$$

in order to recover the correct expectation values

$$O_n = \frac{\langle \Psi_n | \hat{O} | \Psi_n \rangle}{\langle \Psi_n | \Psi_n \rangle} = \frac{\langle \Psi_n | \hat{U}(\alpha) \hat{U}^\dagger(\alpha) \hat{O} \hat{U}(\alpha) \hat{U}^\dagger(\alpha) | \Psi_n \rangle}{\langle \Psi_n | \hat{U}(\alpha) \hat{U}^\dagger(\alpha) | \Psi_n \rangle} = \frac{\langle \tilde{\Psi}_n | \hat{O}(\alpha) | \tilde{\Psi}_n \rangle}{\langle \tilde{\Psi}_n | \tilde{\Psi}_n \rangle}. \tag{1.12}$$

So far we have not explicitly specified the transformation operator $\hat{U}$. Instead, we define it implicitly by looking at the derivation of Eq. (1.9) w.r.t. α

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{H}(\alpha) &= \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}^\dagger(\alpha) \right) \hat{H} \hat{U}(\alpha) + \hat{U}^\dagger(\alpha) \hat{H} \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}(\alpha) \right) \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}^\dagger(\alpha) \right) \hat{U}(\alpha) \hat{H}(\alpha) + \hat{H}(\alpha) \hat{U}^\dagger(\alpha) \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}(\alpha) \right) \\ &= \left[\hat{\eta}(\alpha), \hat{H}(\alpha) \right], \end{aligned} \tag{1.13}$$

where we have introduced the so-called generator of the SRG transformation

$$\hat{\eta}(\alpha) \equiv -\hat{U}^\dagger(\alpha) \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}(\alpha) \right) \tag{1.14}$$

and made use of its anti-hermiticity, which can be derived from the derivative of the unitarity condition $\hat{U}^\dagger(\alpha)\hat{U}(\alpha) = \hat{1}$ w.r.t. α

$$\begin{aligned} & \frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}^\dagger(\alpha) \hat{U}(\alpha) = 0 \\ \Leftrightarrow \quad & -\hat{U}^\dagger(\alpha) \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}(\alpha) \right) = \left(\frac{\mathrm{d}}{\mathrm{d}\alpha} \hat{U}^\dagger(\alpha) \right) \hat{U}(\alpha) \\ \Leftrightarrow \quad & \hat{\eta}(\alpha) = -\hat{\eta}^\dagger(\alpha). \end{aligned} \tag{1.15}$$

The relation derived in Eq. (1.13) is known as the flow equation of the SRG and describes how the Hamiltonian is transformed based on a given generator. Hence, instead of constructing a unitary operator $\hat{U}(\alpha)$ we now need to find an expression for $\hat{\eta}(\alpha)$ such that the evaluation of Eq. (1.13) yields an evolved Hamiltonian with the properties we are aiming for. Up to this point the derivations are completely general and different generators are chosen depending on the application.

The simplest ansatz to construct a generator as originally introduced by Wegner [4] is given by

$$\hat{\eta}(\alpha) = \left[\hat{H}^{\mathrm{d}}(\alpha), \hat{H}^{\mathrm{od}}(\alpha) \right], \tag{1.16}$$

where $\hat{H}^{\mathrm{d}}(\alpha)$ and $\hat{H}^{\mathrm{od}}(\alpha) = \hat{H}(\alpha) - \hat{H}^{\mathrm{d}}(\alpha)$ represent the diagonal and off-diagonal parts of the Hamiltonian w.r.t. a specific basis representation. The flow equation Eq. (1.13) will then suppress the off-diagonal part as it approaches its fix point at $\hat{H}^{\mathrm{od}} = 0$. This allows for a very flexible construction of generators as the choice of basis and the partitioning into diagonal and off-diagonal parts can easily be adapted to different kinds of decoupling.

While we will look at other applications of SRG transformations in Sec. 12.2, let us recall that, for now, the goal is to pre-process or, as it is often referred to, soften the Hamiltonian in a manner that reduces the contributions that couple high and low-momentum states in order to improve the convergence rate of observables. When illustrated in momentum space

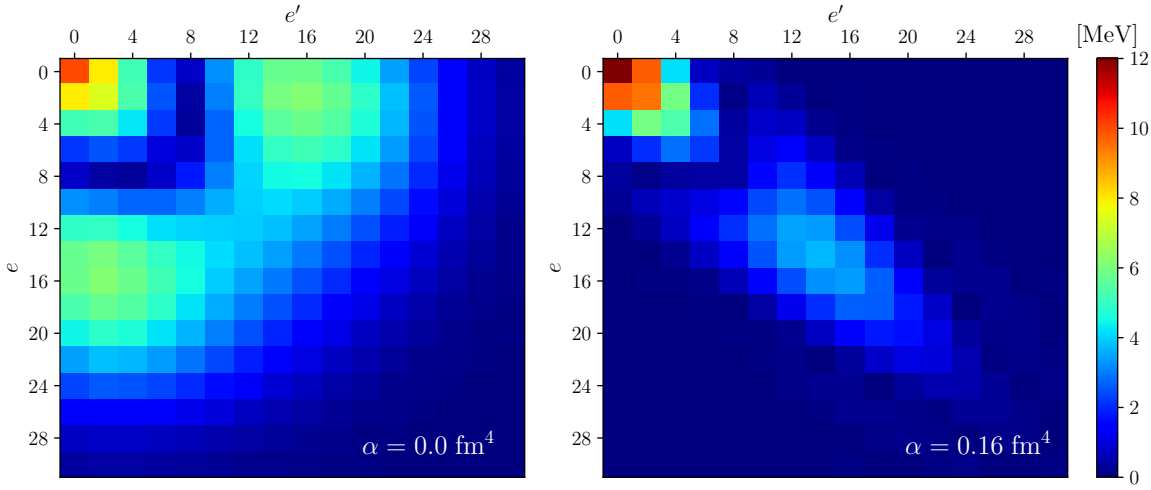

Figure 1.2: Matrix representation of the $^3\mathrm{S}_1$ channel of the NN interaction by Entem, Machleidt and Nosyk [85] at N^{3}LO with cutoff $\Lambda_\mathrm{C} = 500$ MeV in HO representation. The matrix elements are presented bare (left) and SRG evolved to flow parameter $\alpha = 0.16$ fm^4 (right).

or in harmonic oscillator (HO) representation, as done in Fig. 1.2, this becomes equivalent to suppressing the off-diagonal parts of the Hamiltonian.

In order to achieve this, we employ what is called free-space SRG, which is independent of the chosen matrix representation, with the goal of suppressing off-diagonal matrix elements in momentum space. Hence, we need to choose a $\hat{H}^\mathrm{d}(\alpha)$ accordingly. The simplest choice of a generator in free-space SRG is

$$\hat{\eta}(\alpha) = \frac{m^2}{\hbar^4}\left[\hat{T}_\mathrm{int}, \hat{H}(\alpha)\right], \tag{1.17}$$

as $\hat{T}_\mathrm{int}$ is diagonal in momentum space. This will drive the Hamiltonian to a band-diagonal form and can, therefore, also be understood as a pre-diagonalization.

Note that the evaluation of the flow equation and the occurring commutators lead to so-called induced many-body forces, which are terms that contribute at higher particle ranks up to the A-body level. While the exact observables are in theory recovered when including all of these terms, in practice they need to be truncated, commonly at the three-body level. Hence, the eigenvalues of the evolved Hamiltonian will slightly differ from the ones of the unevolved or bare Hamiltonian.

The big advantage of the free-space SRG is its flexibility. Since it is independent of a many-body basis representation, any evolved operators including the Hamiltonian can be used in a variety of many-body methods. Hence, all Hamiltonians employed in this work are SRG evolved to $\alpha = 0.08$ fm^4.

2

The No-Core Shell Model

The NCSM [23, 24, 26] is a conceptually simple yet powerful ab initio method that is capable of delivering a numerical solution to the nuclear many-body problem. It is a member of the class of so-called CI methods, which share the idea of expanding the eigenstates of the Hamiltonian in a complete many-body basis $\{|\Phi_j\rangle\}$

$$|\Psi_n\rangle = \sum_j |\Phi_j\rangle\langle\Phi_j|\Psi_n\rangle. \tag{2.1}$$

In this way the Schrödinger equation Eq. (1.1) can be transformed into a matrix eigenvalue problem

$$\sum_j \langle\Phi_i|\hat{H}|\Phi_j\rangle\langle\Phi_j|\Psi_n\rangle = E_n\langle\Phi_i|\Psi_n\rangle \quad \forall i. \tag{2.2}$$

Since we are trying to solve this for a system of A nucleons, which are fermions, our problem is embedded into an A-body antisymmetric Hilbert space $\mathcal{H}_A$. Hence, antisymmetrized A-body product states, also known as Slater determinants, provide a simple way of constructing a many-body basis that spans $\mathcal{H}_A$. The basis states are constructed from a complete single-particle basis $\{|a_i\rangle\}$

$$|\Phi_j\rangle \equiv |\{a_1, a_2, \ldots, a_A\}_j\rangle_a = \frac{1}{\sqrt{A!}} \sum_\pi \mathrm{sgn}(\pi)\hat{P}_\pi \left\{|a_1\rangle \otimes |a_2\rangle \otimes \ldots \otimes |a_A\rangle\right\}_j \tag{2.3}$$

with permutation operator $\hat{P}_\pi$ and the signum $\mathrm{sgn}(\pi)$ of a given permutation π. While the single-particle basis can be chosen freely, the standard choice is a spherical HO basis, because the wavefunctions are known analytically and the basis has multiple beneficial properties that will enter in the following considerations. Note that the HO basis is not unique either as the wavefunctions depend on the frequency $\hbar\Omega$ of the underlying HO potential. We will take advantage of this dependence later in this work.

While the choice of $\hbar\Omega$ affects the whole basis, an individual HO single-particle state

$$|a_i\rangle = \left\{|n(ls)jm_j\rangle \otimes |tm_t\rangle\right\}_i \tag{2.4}$$

is characterized by radial n, orbital angular momentum l, spin s, total angular momentum j, and total angular momentum projection m_j quantum numbers as well as isospin t with its

projection m_t that indicate the particle species. The corresponding single-particle energy of such a state is given by $\epsilon = (2n + l + \frac{3}{2})\hbar\Omega$.

A key feature of the HO basis is the separation of intrinsic and center-of-mass degrees of freedom such that the eigenstates of the Hamiltonian factorize

$$|\Psi_n\rangle = \{|\Psi_{\text{int}}\rangle \otimes |\Psi_{\text{cm}}\rangle\}_n. \tag{2.5}$$

This allows for the removal of spurious center-of-mass excitations from the low-lying excitation spectrum through a suitable modification of the Hamiltonian in the manner

$$\hat{H} = \hat{H}_{\text{int}} + \lambda_{\text{cm}}\hat{H}_{\text{cm}}, \tag{2.6}$$

where λ_{cm} controls the strength of the added HO center-of-mass Hamiltonian, which is given by

$$\hat{H}_{\text{cm}} = \frac{\hat{\mathbf{P}}_{\text{cm}}^2}{2\hat{M}} + \frac{1}{2}\hat{M}\Omega^2\hat{\mathbf{R}}_{\text{cm}}^2 - \frac{3}{2}\Omega \quad , \quad \hat{\mathbf{R}}_{\text{cm}} = \frac{1}{\hat{M}}\sum_{i=1}^{A}\hat{m}_i\hat{\mathbf{r}}_i. \tag{2.7}$$

Here, $\hat{\mathbf{R}}_{\text{cm}}$ denotes the center-of-mass coordinate obtained from the single-particle coordinates $\hat{\mathbf{r}}_i$ and frequency Ω is the same as for the employed HO basis.

The solution of the eigenvalue problem further gives access to other observables such as radii or electromagnetic properties through the computed eigenstates. For any operator $\hat{O}$ the expectation value w.r.t. a given NCSM eigenstate can be calculated as

$$\langle\Psi_n|\hat{O}|\Psi_n\rangle = \sum_{ij}\langle\Psi_n|\Phi_i\rangle\langle\Phi_i|\hat{O}|\Phi_j\rangle\langle\Phi_j|\Psi_n\rangle. \tag{2.8}$$

In the context of this work, mean-square mass radii are of particular interest for which the corresponding operator is given by

$$\hat{R}_{\text{ms}}^2 = \frac{1}{A}\sum_{i=1}^{A}(\hat{\mathbf{r}}_i - \hat{\mathbf{R}}_{\text{cm}})^2 = \frac{1}{\hat{M}}\sum_{i<j}^{A}\left(\frac{\hat{m}_i + \hat{m}_j}{A} - \frac{\hat{m}_i\hat{m}_j}{\hat{M}^2}\right)\hat{\mathbf{r}}_{ij}^2 \tag{2.9}$$

with relative coordinate $\hat{\mathbf{r}}_{ij} = \hat{\mathbf{r}}_j - \hat{\mathbf{r}}_i$. Under the assumption of equal nucleon masses this can further be simplified to

$$\hat{R}_{\text{ms}}^2 = \frac{1}{A^2}\sum_{i<j}^{A}\hat{\mathbf{r}}_{ij}^2. \tag{2.10}$$

2.1 Model Space and Symmetries

With the matrix eigenvalue problem set up, we can address the issue of the dimensionality of the matrix being infinite. In order to render the problem finite, and computationally feasible, a truncation needs to be employed. It is the choice of truncation that distinguishes the NCSM from other CI methods. We employ a truncation on the many-body level by limiting the sum over all single-particle HO excitation quanta w.r.t. the lowest Pauli-allowed state $|\Phi_0\rangle$ to $N_{\max}$

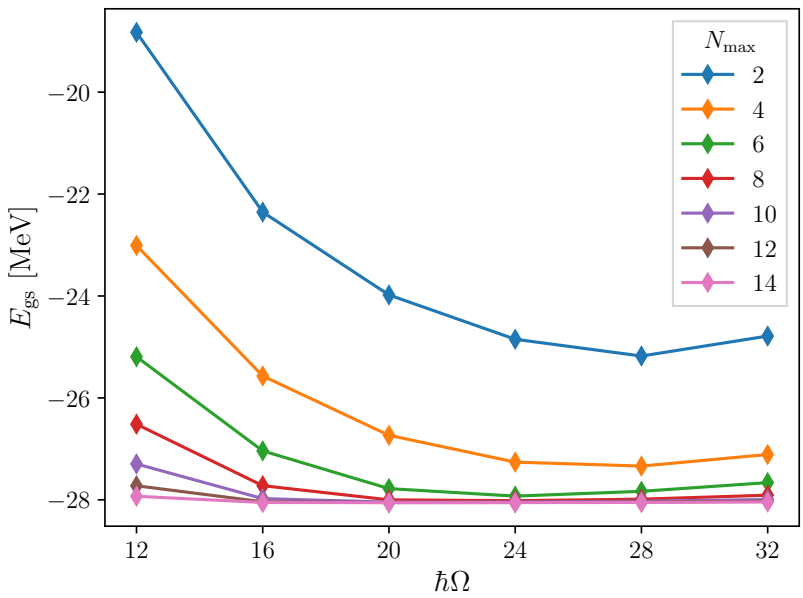

Figure 2.1: NCSM calculations for the ground-state energy of ^{4}He with the NN+3N interaction by Hüther et al. [78] that has been SRG evolved to $\alpha = 0.08$ fm^4.

such that

$$N_{\max} \geq N_{|\Phi\rangle} = \sum_{i=1}^{A} e_{i,|\Phi\rangle} - E_0. \tag{2.11}$$

Here, $e_{i,|\Phi\rangle}$ is the principle quantum number $e = 2n + l$ of the i-th HO state in $|\Phi\rangle$ and E_0 resembles the sum over all e in $|\Phi_0\rangle$, which is chosen to be a Slater determinant with the lowest sum of single-particle energies. Hence, the model space contains all Slater determinants that obey this truncation

$$\mathcal{M}_{\text{NCSM}} = \left\{\text{all } |\Phi\rangle \text{ for which } N_{|\Phi\rangle} \leq N_{\max}\right\}. \tag{2.12}$$

Since the model space for $N_{\max} \to \infty$ corresponds to the full Hilbert space the energies converge towards the exact solution with increasing $N_{\max}$. Moreover, the calculation obeys a variational principle and we, therefore, find a monotonously decreasing convergence behavior as illustrated in Fig. 2.1. We, further, observe that the rate of convergence depends on the chosen $\hbar\Omega$ and we can find an optimum that is often referred to as the variational minimum. While the $N_{\max}$ truncation yields systematics that allow for a controlled convergence towards the exact solution, model spaces still grow rapidly with $N_{\max}$ and factorially w.r.t. particle number A. Thus, achieving convergence is, in practice, impossible for all but the lightest nuclei. One can think of two ways to further reduce the model-space size.

First, symmetries of the system can be exploited. Charge conservation yields that the number of neutrons and protons and, thus, the projection of the total isospin $M_T = \sum_{i=1}^{A} m_{t_i}$ is conserved. Hence, only Slater determinants with the correct M_T enter the model space. We can further make use of the rotational invariance of our system. This implies that the projection of the total angular momentum $M_J = \sum_{i=1}^{A} m_{j_i}$ is conserved and the energy spectrum is degenerate with respect to it. Therefore, we can fix M_J to a certain value und

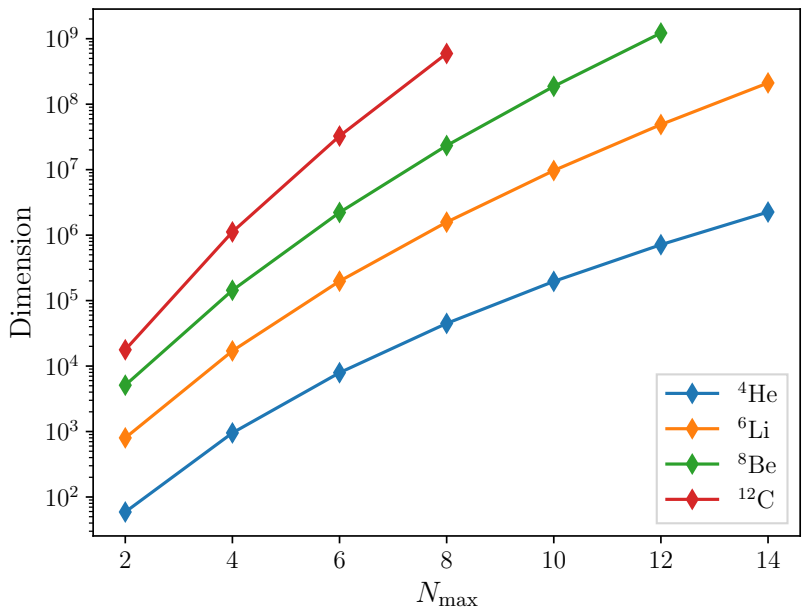

Figure 2.2: Growth of NCSM model-space size with $N_{\max}$ for different nuclei with increasing particle number A.

exclude all Slater determinants with other M_J from the model space. Also, the Hamiltonian preserves parity, which again transfers to the Slater determinants, yielding another restriction on the size of the model space.

Second, additional truncations, e.g., on the single-particle level can be introduced. The simplest truncation is an upper limit to the number of excitations of a single particle, where the principle quantum number is limited to $e_{\max}$. Analogously, other quantum numbers such as the orbital angular momentum can be truncated individually.

Unfortunately, even with all these truncations, the model space dimensions easily grow to 10^7 and above (see Fig. 2.2) if we want to get anywhere close to convergence. Hence, high-performance computing and highly sophisticated numerical algorithms, like the Lanczos algorithm [90], are needed to make the extraction of the lowest-lying eigenstates computationally feasible.

The Jacobi No-Core Shell Model. Instead of using Slater determinants one can alternatively construct a many-body basis in terms of generalized relative and center-of-mass coordinates, so-called Jacobi coordinates, giving rise to the Jacobi NCSM [91]. This makes extended use of the separation of relative and center-of-mass degrees of freedom in the HO basis, where the latter can be eliminated completely, thus, effectively reducing the size of the Hilbert space by the equivalence of a single particle. In addition, the Jacobi many-body states are fully coupled in terms of angular momenta, carrying a good total A-body angular momentum J. Hence, the model-space size can further be reduced to a single choice of J.

Note that none of these preselections affect the model-space truncation scheme and, therefore, NCSM and Jacobi NCSM calculations at the same $N_{\max}$ are equivalent. Consequently, the Jacobi NCSM allows for very efficient computation of small systems up to $A \lesssim 6$. Beyond that the construction of the coupled basis becomes increasingly complicated and time consuming.

2.2 Importance Truncation

So far we have discussed truncations that act as upper limits to single-particle quantum numbers or sums of those. While they provide an efficient framework which allows for systematic improvability, model spaces still grow intractably large, especially for increasing particle number, and the truncations will eventually cut off too much of the relevant physics.

Here, importance truncation (IT) [25] provides an alternative or rather complementary truncation scheme, based on a physics-informed estimate of the relevance of a single Slater determinate. It allows for a selective removal of insignificant many-body states from the model-space. In order to provide such an estimate we need to construct an importance measure that yields information for the individual basis states. To ensure that this information is physically meaningful, we start from a set of reference states $|\Psi_m^{\text{ref}}\rangle$ that resemble a good approximation to the states we aim to describe. A good candidate for such a set of reference states is the solution of a CI calculation in a smaller model space also called reference space $\mathcal{M}_{\text{ref}}$. These states can then also be expressed through an expansion in Slater determinants

$$|\Psi_m^{\text{ref}}\rangle = \sum_{\mu\in\mathcal{M}_{\text{ref}}} |\Phi_\mu\rangle\langle\Phi_\mu|\Psi_m^{\text{ref}}\rangle. \tag{2.13}$$

With this we can make use of multi-configurational perturbation theory for which the first-order correction κ_ν for a given state $|\Phi_\nu\rangle$ reads

$$\kappa_\nu^m = -\frac{\langle\Phi_\nu|\hat{H}|\Psi_m^{\text{ref}}\rangle}{\Delta E_\nu} = -\sum_{\mu\in\mathcal{M}_{\text{ref}}} \frac{\langle\Phi_\nu|\hat{H}|\Phi_\mu\rangle\langle\Phi_\mu|\Psi_m^{\text{ref}}\rangle}{\Delta E_\nu} \quad \text{for} \quad |\Phi_\nu\rangle \notin \mathcal{M}_{\text{ref}}, \tag{2.14}$$

with ΔE_ν being the HO excitation energy of the corresponding state. This quantity κ_ν^m yields enough information about the relevance of a state since higher values imply larger contributions to the energy of the system. Thus, it can directly be used as the importance measure we are looking for. The truncation is employed by introducing a lower limit κ_{min}, below which the states are being discarded. The resulting model space for a so-called importance-truncated no-core shell model (IT-NCSM) calculation is then given by

$$\mathcal{M}_{\text{IT-NCSM}} = \mathcal{M}_{\text{ref}} \cup \{\text{all } |\Phi_\nu\rangle \notin \mathcal{M}_{\text{ref}} \text{ for which } |\kappa_\nu^m| \geq \kappa_{\text{min}} \text{ for any } m\}, \tag{2.15}$$

where the full NCSM model space is recovered for $\kappa_{\text{min}} = 0$. As with every truncation we need to be aware that an additional error is introduced which needs to be taken into account. Typically, IT-NCSM calculations are performed for a sequence of different κ_{min}, which can then be extrapolated to the limit of vanishing threshold. The extrapolation is generally performed through fits of second- or third-order polynomial functions, where polynomials of neighboring orders serve as uncertainty estimates. The extrapolation procedure is illustrated in Fig. 2.3. In most applications the IT errors are much smaller than errors induced from other truncations such as the N_{max} truncation and are, therefore, being neglected.

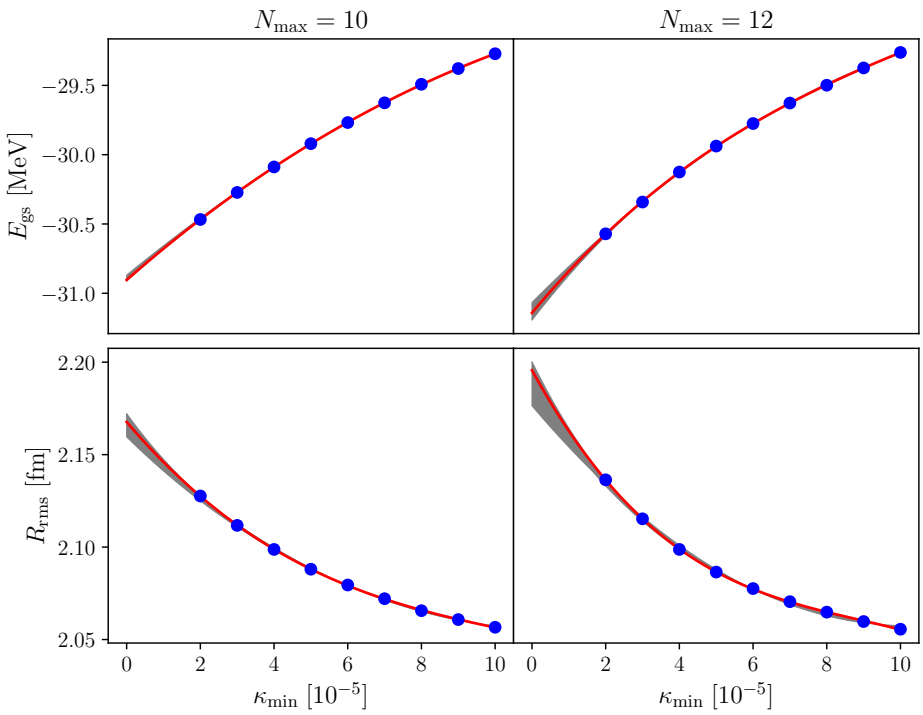

Figure 2.3: IT extrapolations based on $\kappa_{\min}$ sequences for ground-state energy and rms radius of ^{6}Li at $N_{\max}$ =10,12.

3

Model-Space Extrapolation and Uncertainty Quantification

In the previous chapters we have discussed various truncation schemes that are required to convert the nuclear many-body problem into a computable form. However, the benefits of these truncations come at the cost of induced errors as we are not able to retain the exact solution of the problem. This naturally raises questions on how to estimate converged results or, at least, provide realistic uncertainties for the obtained results. But, a correct quantification of all uncertainties is extremely challenging. Fortunately, ab initio theory and its inherent systematics allow for promising conceptual approaches.

In principle we need to address uncertainties from multiple sources:

- the choice of the parameters in chiral EFT,
- the truncation of the chiral expansion,
- the $N_{\max}$ truncation of the model space,
- and other model-space truncations such as IT,

which are usually separated into two major groups: interaction uncertainties that are comprised of all aspects regarding the construction of an interaction from chiral EFT, and many-body uncertainties that include all errors and truncations related to the solution of the many-body problem. While the assessment of individual sources of uncertainty can be very challenging depending on the truncation scheme, combining different uncertainties that are not completely unrelated is even more so. We will, therefore, focus on the individual uncertainties and will not attempt a combined uncertainty quantification.

3.1 Interaction Uncertainties

Let us first consider interaction uncertainties. There a two different methods which are commonly employed in order to construct uncertainty estimates at an interaction level. Both are built on different sources of errors, on one hand the truncation of the chiral expansion and on the other hand the choice of cutoff and LECs.

Truncation Error of the Chiral Expansion. We start with the former as it is considered to yield the largest uncertainties. This is due to the fact that chiral interactions, especially with 3N forces included, are only available up to N^{2}LO or N^{3}LO depending on the family. Hence, the uncertainty needs to be constructed from a maximum of four data points. However, one can exploit the systematics of the power counting as proposed in [92]. It resembles an expansion in a small parameter Q, which roughly equals $1/3$ for most applications. As a consequence, one would expect the contribution for a given observable at a given chiral order to also be about a third of the previous order, corresponding to the associated power of Q. Under this assumption the chiral expansion of an observable X at n-th order $X^{(n)}$ can be written as

$$X^{(n)} = X^{(0)} + \Delta X^{(2)} + \ldots + \Delta X^{(n)}, \tag{3.1}$$

where $\Delta X^{(i)} = X^{(i)} - X^{(i-1)}$ with the exception of $\Delta X^{(2)} = X^{(2)} - X^{(0)}$ since there is no contribution at Q^1. Following the previous reasoning, each correction at order i should then approximately equal the correction at order j scaled by the correct power of Q yielding

$$\Delta X^{(i)} = \mathcal{O}(Q^{i-j}\Delta X^{(j)}). \tag{3.2}$$

The uncertainty $\delta X^{(i)}$ for a result at a given order i is then conservatively estimated to be

$$\delta X^{(i)} = \max_{2\leq j\leq i}\left(Q^{i+1}\left|X^{(0)}\right|, Q^{i-j+1}\left|\Delta X^{(j)}\right|\right). \tag{3.3}$$

While this algorithm is straightforward and easy to implement due to its simplicity, there also is a more sophisticated estimation of the truncation uncertainty developed by the BUQEYE collaboration [34]. It is based on Bayesian statistics with the goal to find a probability distribution for the exact observable based on calculations at the first orders. The ansatz differs from the previous one by expressing Eq. (3.1) as a power-series

$$X^{(n)} = X_{\text{ref}} \sum_{i=0}^{n} c_i Q^i \tag{3.4}$$

in terms of coefficients c_i, which are to be understood relative to a reference scale X_{ref}. Note that we can account for the missing contribution at Q^1 by requiring $c_1 = 0$. This reference scale is typically chosen to be either the experimental value or any other reasonable estimate. A strong dependence on it is not expected as long as the resulting coefficients are of natural size, i.e., $\mathcal{O}(1)$. We now aim for a prediction of the c_i for higher orders based on the known coefficients. In a Bayesian framework the first step is choosing priors. The choice of priors only has a minor impact on the final result and can, therefore, be chosen to have beneficial analytical properties. Hence, we assume that the conditional probability distribution pr for the coefficients c_i is given by a normal distribution $\mathcal{N}$ with marginal variance $\bar{c}^2$

$$\mathrm{pr}(c_i|\bar{c}^2) \propto \mathcal{N}(0, \bar{c}^2) \quad , \quad \mathrm{pr}(\bar{c}^2) \propto \chi^{-2}(\nu_0, \tau_0^2), \tag{3.5}$$

where the latter has been chosen to be from an inverse χ^2 distribution with the number of degrees of freedom ν_0 and the scale τ_0 for the prior. The prior belief about the uncertainty of

an observable at order n then takes the form

$$\mathrm{pr}(\delta X^{(n)}|\bar{c}^2, Q) \propto \mathcal{N}\left(0, X_{\mathrm{ref}}^2 \frac{Q^{2(n+1)}}{1-Q^2}\bar{c}^2\right). \tag{3.6}$$

In order to update this with the information we have about the known coefficients $\mathbf{c}_j = \{c_0, c_1, \ldots, c_j\}$ we first need to update $\bar{c}^2$ through

$$\mathrm{pr}(\bar{c}^2|\mathbf{c}_j) \propto \mathrm{pr}(\mathbf{c}_j|\bar{c}^2)\mathrm{pr}(\bar{c}^2) = \mathrm{pr}(\bar{c}^2)\prod_{i\leq j}\mathrm{pr}(c_i|\bar{c}^2) \tag{3.7}$$

making use of Bayes' theorem. This, again, results in a scaled inverse χ^2 distribution with updated parameters

$$\mathrm{pr}(\bar{c}^2|\mathbf{c}_j) \propto \chi^{-2}(\nu, \tau^2) \quad \text{with} \quad \nu = \nu_0 + n_c,\ \tau^2 = (\nu_0\tau_0^2 + \mathbf{c}_j^2)/\nu, \tag{3.8}$$

where n_c is the number of non-zero coefficients in $\mathbf{c}_j$. By marginalizing over $\bar{c}^2$ we can obtain the posterior distribution for the truncation error given by a Student-t distribution

$$\mathrm{pr}(\delta X^{(n)}|\bar{c}^2, Q) \propto t_\nu\left(0, X_{\mathrm{ref}}^2 \frac{Q^{2(n+1)}}{1-Q^2}\tau^2\right), \tag{3.9}$$

from which we can extract an uncertainty by integrating over it up to a certain degree of belief, which is usually chosen to be 68% or 95%. For more details on the concept and derivation we refer to [33, 34, 47, 93].

Choice of Parameters in Chiral EFT. Let us further discuss how to estimate an interaction uncertainty from the parameter dependencies, in particular cutoff Λ_C and LECs. As already mentioned, LECs occur at different orders of the chiral interaction and correspond to different diagrams in the chiral expansion. In the purely nucleonic sector there is an abundance of NN scattering data that can be used to very tightly constrain all LECs associated with two-body terms. However, at the currently available N^{2}LO and N^{3}LO interactions one is left with two three-body LECs called c_D and c_E. There are various strategies on how to fit these to 3N or many-body observables resulting in different predictions for observables [76, 78, 84, 94].

In order to estimate an uncertainty from the choice of such parameters, the straightforward method is to sample the space of all predictions for a given observable $p(\Lambda_C, \mathrm{LECs})$ multiple times and turn to statistical methods to quantify the spread of the resulting distribution and from that an uncertainty [35]. This can for example be done by performing calculations for Hamiltonians obtained with various combinations of parameters as discussed in [95]. While this is very costly with conventional many-body methods, the recent rise of emulators, in particular eigenvector continuation (EC), has given access to millions of samples with minimal computational effort [96, 97].

One should also consider looking at the dependence on the regulator scheme, however, it has only recently, with sufficient precision of the many-body calculations, become apparent, that calculations with interactions based on different regulator schemes do not necessarily agree with each other. Therefore, this is a rather new discussion and, while it is of major importance, there are no established uncertainty estimates we could discuss here.

3.2 Many-Body Uncertainties

With the interaction uncertainties covered, we will now turn to the estimation of many-body uncertainties. While most of these schemes can be applied to different many-body methods, we will focus on the NCSM as our many-body method of choice. Since the NCSM is an ab initio method, we can exploit the systematics of the model-space truncation scheme similar to the truncation error of the chiral expansion. Besides estimating uncertainties for the obtained results, our goal is also to extrapolate results from finite model spaces to the full infinite Hilbert space. In order to do so, we will exploit that the convergence behavior for NCSM calculations is controlled by two parameters, the model space truncation $N_{\max}$ and the HO frequency of the underlying single-particle basis $\hbar\Omega$. As the convergence patterns for different observables differ drastically and can show very different characteristics, as illustrated in Fig. 3.1, we limit ourselves to energies and radii and discuss them separately.

Extrapolation of Energies. Bound-state energies are naturally the first observable to look at, as they are directly obtained as solutions of the Schrödinger equation. Being a variational method, energies in the NCSM exhibit a monotonously decreasing convergence. A typical set of such converging series is shown in the left-hand panel of Fig. 3.1. The different sequences emerge from different choices of $\hbar\Omega$. By construction, they all converge to the same limit. At first glance these sequences seem to resemble an exponential decay. Thus, the convergence behavior is typically modeled with some kind of exponential function [50–52]

$$E^{\hbar\Omega}(N_{\max}) = a \cdot e^{-b\,N_{\max}} + E_{\infty}^{\hbar\Omega}, \tag{3.10}$$

where the converged energy $E_{\infty}^{\hbar\Omega}$ as well as $a, b > 0$ are fit parameters. The superscript $\hbar\Omega$ indicates the dependence on the HO frequency. In the simplest application this fitted limit and its fit uncertainty can directly be interpreted as an extrapolation with estimated error. However, when doing this for different $\hbar\Omega$ it becomes apparent that the different limits can be inconsistent with each other, clearly a deficiency of the extrapolation method. In order to include information from multiple $\hbar\Omega$ one could either simultaneously fit multiple exponential functions that share the same limit or, as more commonly done, construct a more reasonable

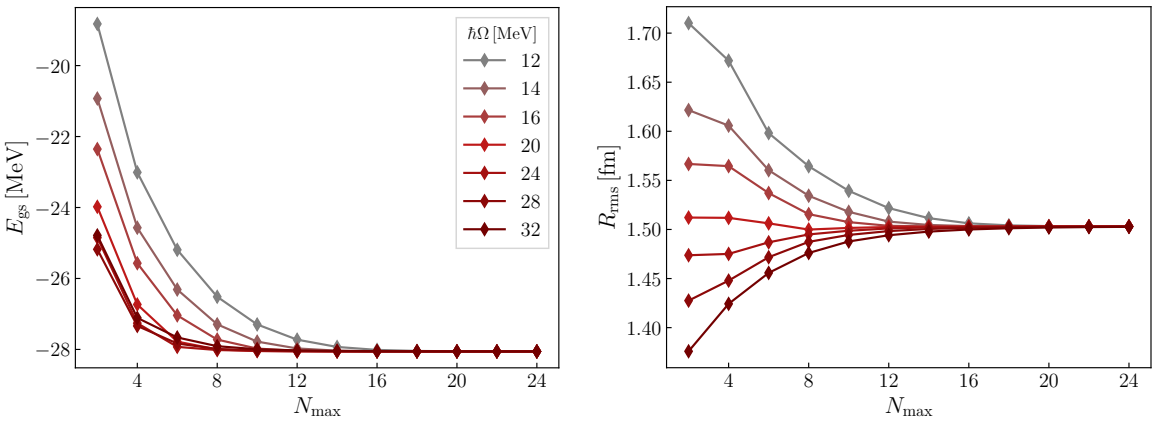

Figure 3.1: NCSM calculations for the ground-state energy (left) and rms radius (right) of ^{4}He.

uncertainty measure from exponential extrapolations of single sequences. For our applications we stick to the procedure described in [52], which we will refer to as classical extrapolation from now on. Here, exponentials of the form Eq. (3.10) are fitted to single sequences for multiple different $\hbar\Omega$ around the variational minimum. The best estimate of the extrapolation is then defined as

$$E_\infty = E_\infty^{\hbar\Omega} \quad \text{for which} \quad |E_\infty^{\hbar\Omega} - E^{\hbar\Omega}(\mathcal{N}_{\max})| \quad \text{is minimal.} \tag{3.11}$$

Here, $\mathcal{N}_{\max}$ indicates the highest $N_{\max}$ available in the fit data. The corresponding uncertainty is assumed to be the maximum of the following three estimates:

- the difference of two extrapolations at consecutive $N_{\max}$ $|E_\infty^{\hbar\Omega}[\mathcal{N}_{\max}] - E_\infty^{\hbar\Omega}[\mathcal{N}_{\max} - 2]|$;
- half the variation in $E_\infty^{\hbar\Omega}$ over a 8-MeV interval in $\hbar\Omega$ around the variational minimum;
- 20% of $|E_\infty^{\hbar\Omega} - E^{\hbar\Omega}(\mathcal{N}_{\max})|$.

While this procedure is practical and phenomenologically motivated, there is another, more elaborate method called infrared (IR) extrapolation, which also builds on an exponential modeling [98–100]. Conceptually, this is an attempt to capture the convergence behavior in dependence on the IR and ultraviolet (UV) cutoffs L and Λ. The UV convergence is understood to be much faster than the IR convergence in suitable model spaces, i.e., for suitable choices of $\hbar\Omega$. It further has been shown, that the IR convergence is exponential in L, however, the exact determination of L remains challenging [49, 53]. Hence, the resulting fit function is

$$E_{\text{IR}}(L) = a \cdot e^{-2\kappa_\infty L} + E_\infty, \tag{3.12}$$

where a, κ_∞ and E_∞ are fit parameters. Additional information on the nucleus under consideration enters in $L \equiv L(N_{\max}, \hbar\Omega)$. Uncertainties can, again, be obtained from the uncertainties of the fit or statistically from fits to multiple subsets of the data. For a more detailed discussion of the method we refer to [49]. The downside of this approach is, that it imposes additional requirements and constraints on the many-body calculations. This results in rather exotic choices of $\hbar\Omega$. While this is unproblematic for the standard formulation of the NCSM, the cost for IT-NCSM calculations varies strongly with $\hbar\Omega$ and and increases rapidly for choices of $\hbar\Omega$ far from the variational minimum. For most realistic Hamiltonians it further requires sufficiently large $N_{\max}$ that tend to be outside the current reach of many-body methods especially if 3N forces are taken into account. It is, therefore, not feasible for our applications.

Extrapolation of Radii. Contrary to energies, radii do not obey any kind of principle that would constrain the convergence behavior, which, therefore, takes a much more complicated form as shown in the right-hand panel of Fig. 3.1. In practice, some extrapolation methods that work in single cases are employed manually but, to our knowledge, there exists no generalized extrapolation scheme that would allow a straightforward extrapolation of radii. One advantage of radii over energies is the strong correlation and, therefore, dependence on $\hbar\Omega$. Hence, fine-tuning the single-particle basis can lead to rapid convergence, but it is not guaranteed. We will, therefore, refrain from any kind of classical extrapolation for radii.

It should further be mentioned that the IR extrapolation scheme can also be extended to radii [53, 98]. However, this comes with the same limitations as for energies.

Part II

Artificial Neural Networks for the No-Core Shell Model

4

Machine Learning

Throughout the past two decades, ML algorithms have experienced major computational and methodological advances leading to a major impact in all areas related to data science and artificial intelligence. With their exceptional performance in a variety of computational tasks, such as optimization, interpolation, classification, or pattern recognition, they have quickly entered not only our everyday life but also a wide range of scientific fields including nuclear physics. Neural-network quantum states [101, 102], Bayesian ML [103–105], extrapolation tools [56, 57], and other data-driven approaches [106, 107] are only a few examples of ML applications in nuclear structure theory. The interested reader can find a comprehensive overview of most ongoing developments in [54, 55].

In the context of this work we will focus on ML applications that are supplemental to nuclear many-body methods and the NCSM in particular. The main bottleneck of all of these methods are rapidly growing model spaces. We have already discussed some schemes that attempt to quantify, or at least approximate, the model-space dependence of observables within these methods but, besides some successes with bound-state energies, the exact functional dependence remains a mystery. Fortunately, ML applications have proven very powerful in situations where this is the case as they are designed to emulate complex correlations without requiring initial knowledge of the problem under consideration. Hence, we will discuss how ML can extend the reach of nuclear structure calculations and build a tool that is capable of performing model-space extrapolations for observables beyond bound-state energies.

But, before we turn to specific applications let us first discuss some of the most important basics of ML.

4.1 Basic Concepts

The main idea behind ML is to construct generalized algorithms that are able to (iteratively) adapt to data in order to capture any relationship between variables. They can also be understood as universal function approximators [108]. It is the generalization that makes ML such a useful tool, as it can adapt to various problems that would either take too long to compute or where insufficient knowledge about the problem is available. The most common categories of ML algorithms are

- **supervised learning:** the algorithm is informed with labeled data, i.e., the desired output for every input is known;
- **unsupervised learning:** the algorithm is informed with unlabeled data, i.e., it is constructed to find unknown patterns in the data;
- **reinforcement learning:** the algorithm is build to solve a problem by maximizing a defined measure also known as reward.

All ML applications discussed in this work rely on supervised learning. In order to get a basic understanding of supervised learning let us first look at linear regression, which is one of the most basic ML algorithms that still contains all relevant aspects of ML we need to introduce. In its simplest form linear regression can be understood as the best fit of a straight line to a set of data points, the so called training data. For simplicity we will limit these considerations to two dimensions. The training data then consists of labeled training samples $\mathcal{S}_i = (x_i, y_i^{\mathrm{t}})$, meaning that we have assigned a target value y_i^{t} to every value x_i. Hence, the function we want to fit to this data is a line

$$y(x) = wx + b. \tag{4.1}$$

In foresight of the subsequent discussion of more complex ML models we need to introduce some terminology. Thus, we will call y the output for a given input x and w, b are free parameters where we make a distinction between weight w and bias b. We now want an algorithm that automatically adjusts the free parameters in such a way that we end up with an optimal description of the training data, i.e., the best fit. In order to define what optimal means in this context, we further need a measure that accounts for the difference of the output y and the target output y^{t} across the whole training set. A common choice is the mean-square error (MSE)

$$\mathcal{L}_{\mathrm{MSE}}(y_i, y_i^{\mathrm{t}}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^{\mathrm{t}})^2, \tag{4.2}$$

where N is the number of outputs. This measure is referred to as loss or loss function respectively. We are now left with a minimization problem w.r.t. this loss. Linear regression is special in the sense that this minimization problem can be solved analytically, however, we will instead discuss a numerical algorithm called gradient descend as it can also be employed in ML algorithms beyond linear regression.

Gradient descend is an iterative algorithm used to find a minimum of a function. It requires an initial choice of the free parameters $w = w_0, b = b_0$, which are updated iteratively in the direction of the steepest descend or, in mathematical terms, the gradient

$$w_{j+1} = w_j - \eta\frac{\partial\mathcal{L}}{\partial w} \quad , \quad b_{j+1} = b_j - \eta\frac{\partial\mathcal{L}}{\partial b}. \tag{4.3}$$

Here we have introduced the so-called learning rate η that controls the step size of the iteration. The optimization step can then be repeated until the loss levels out and the minimum is reached. Note that the choice of learning rate is of particular importance for more complicated loss functions with multiple minima since the step size controls which minimum the optimization will end up in. Overall, this example resembles a simple but complete ML algorithm that we can now adapt to more complicated applications.

4.2 Artificial Neural Networks

The ML model we employ in this work is an artificial neural network (ANN), which is at the core of the ML discipline called deep learning [109–111]. As the name suggests it is loosely modeled after biological neural networks and does, therefore, consist of artificial neurons. An illustration of such a neuron is given in Fig. 4.1a. Similar to our previous example these neurons take an input $\mathbf{x}$ from which they generate an output y based on other parameters again called weights $\mathbf{w}$ and bias b such that

$$y = \sigma\left(\sum_j x_j w_j + b\right). \tag{4.4}$$

The major difference to the linear case is the activation function σ, which is also known as nonlinearity as it usually is some kind of non-linear function. These neurons can then be combined to arbitrarily complicated structures, the ANNs. In a standard fully-connected feed-forward (FCFF) network the neurons are organized in layers (see Fig. 4.1b). Every neuron takes the outputs of all neurons in the preceding layer as input and passes its output on to all neurons in the subsequent layer. Hence we can modify the previous equation such that the output of the i-th neuron in the l-th layer is given by

$$y_i^{[l]} = \sigma\left(\sum_j x_j^{[l-1]} w_{ji}^{[l]} + b_i^{[l]}\right). \tag{4.5}$$

Here, the weights can be understood as the strengths of the respective connection between neurons and the biases resemble internal thresholds.

Depending on the problem, ANNs occur in different shapes and sizes and their modularity allows, and also requires, thorough adaption to the given task. Most commonly employed architectures besides FCFF networks are convolutional neural networks (CNN) for image

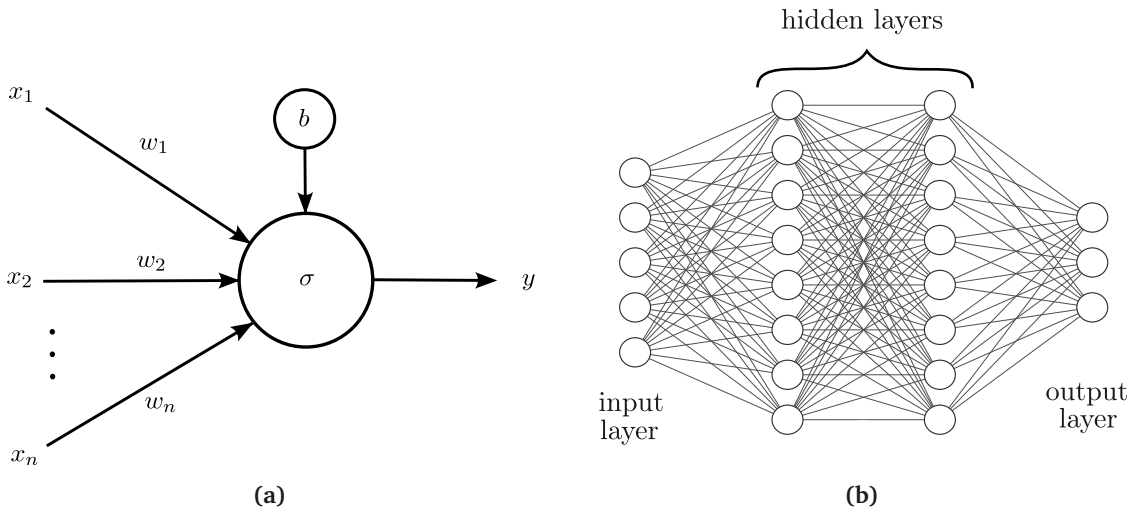

Figure 4.1: Schematic representation of an artificial neuron (a) and a feed-forward ANN constructed from such neurons (b).

processing [112, 113], recurrent neural networks (RNN) for time-series prediction such as natural language processing [114], or even more sophisticated generic algorithms such as neuroevolution of augmented topologies (NEAT) that allow for automatic iterative adaption of the network structure itself [115]. What all of these architectures have in common is that they resemble high-dimensional functions with up to hundreds of billions of free parameters [116] that need to be determined from training data.

Mathematically the ANN resembles a set of nested functions and matrix multiplications. Assume that we have L layers in our network and that

$$\begin{aligned} f^{[l]} : \mathbb{R}^{\dim(l-1)} &\longmapsto \mathbb{R}^{\dim(l)} \\ \mathbf{y}^{[l-1]} &\longmapsto \mathbf{y}^{[l]} = \sigma\left(\mathbf{w}^{[l]}\mathbf{y}^{[l-1]} + \mathbf{b}^{[l]}\right) \end{aligned} \tag{4.6}$$

describes how the the outputs of the l-th layer $\mathbf{y}^{[l]}$ are obtained from those in the $(l-1)$-th layer $\mathbf{y}^{[l-1]}$ respectively. Here, $\dim(l)$, $\mathbf{b}^{[l]}$, and $\mathbf{w}^{[l]}$ are the size of layer l, the vector of biases in layer l, and the matrix of weights connecting layers l and $l-1$. The corresponding function representing the whole network can then be denoted as

$$\begin{aligned} \mathcal{F} : \mathbb{R}^{\dim(1)} &\longmapsto \mathbb{R}^{\dim(L)} \\ \mathbf{x} &\longmapsto \mathbf{y}^{[L]} = \left(f^{[L]} \circ \ldots \circ f^{[1]}\right)(\mathbf{x}). \end{aligned} \tag{4.7}$$

Training an ANN basically resembles a high-dimensional fit of $\mathcal{F}$ to a set of training data. In order to do so, we start by taking one of the training samples $\mathcal{S}_i = (\mathbf{x}_i, \mathbf{y}_i^{\mathrm{t}})$ and pass its input component $\mathbf{x}_i$ through the ANN. This is known as forward pass. We can then assess the deviation of the prediction from the target value via the loss function. Sticking to our previous example of MSE this is given by

$$\mathcal{L}_{\mathrm{MSE}}(\mathbf{x}_i, \mathbf{y}_i^{\mathrm{t}}) = \frac{1}{\dim(L)}(\mathcal{F}(\mathbf{x}_i) - \mathbf{y}_i^{\mathrm{t}})^2. \tag{4.8}$$

Now, all free parameters in the ANN, i.e., all weights and biases are updated based on the obtained loss, starting with the last layer all the way towards the first layer. This process known as backpropagation [117] is key to large-scale ML applications as it allows for a systematic optimization of all parameters with the goal of minimizing the loss. It requires the partial derivative of the loss function with respect to every single free parameter in the network which is then iteratively updated according to Eq. (4.3). Fortunately, the nested structure of FCFF networks allows for computationally efficient layer-by-layer calculations using the chain rule. The partial derivative of $\mathcal{L}$ with respect to the weights $\mathbf{w}^{[l]}$ or biases $\mathbf{b}^{[l]}$ in layer l is given by

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}^{[l]}} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{[l]}} \cdot \left[\mathbf{y}^{[l-1]}\right]^{\mathrm{T}} \quad , \quad \frac{\partial \mathcal{L}}{\partial \mathbf{b}^{[l]}} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{[l]}} \tag{4.9}$$

$$\text{with} \quad \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{[l]}} = \left[\mathbf{w}^{[l+1]}\right]^{\mathrm{T}} \cdot \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{[l+1]}}\, \sigma^{[l]\prime}(\mathbf{z}^{[l]}), \tag{4.10}$$

where $\mathbf{z}^{[l]} = \mathbf{w}^{[l]}\mathbf{y}^{[l-1]} + \mathbf{b}^{[l]}$. While we have limited our discussion here to the simplest case we want to emphasize that there are modern, more sophisticated backpropagation algorithms that

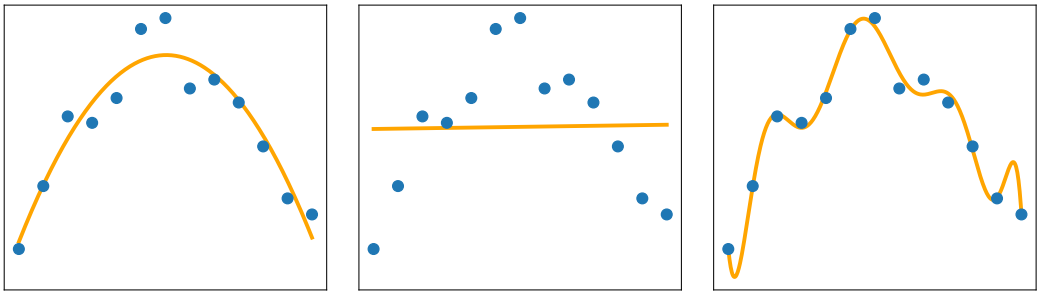

Figure 4.2: Schematic depiction of an ANN describing data well (left), one that suffers from underfitting (center), and one that clearly overfits the data (right).

introduce modifications of the above equations [118–121]. We further note that in practice backpropagation is not carried out with individual samples but in a stochastic manner with batches of samples. This improves the training process significantly and the above derivation can easily be adapted to batches. A full iteration through the whole training set is called an epoch and successful training usually requires multiple epochs. The training of an ANN is finished when the loss does not decrease anymore.

While at first glance this seems like a straight-forward procedure to fit any function, there are some potential pitfalls and bottlenecks that should be mentioned. First of all, fixing many free parameters requires a lot of data. Hence, a sufficiently large set of training data is required and a rule of thumb is that more training data is always better. However, the data needs to be of good quality, which brings us to the second problem: any biases in the training data will affect the performance of the ANN. The only way to circumvent this is to train on a well balanced set of data, which resembles the unseen data, for which predictions should be made, in every aspect. In this regard preprocessing the data, e.g., via normalization can facilitate the training process and improve predictions.

Another potential issue that is related to the number of free parameters is under- or overfitting. Ideally, if the size of the ANN is selected accordingly, the overall functional dependence of the data is described properly as depicted in the left-hand panel of Fig. 4.2 and is expected to yield accurate predictions on unseen data. If the ANN is too small, i.e., the number of free parameters is insufficient to describe the data, one ends up with a network that is unable to capture any correlations in the data (center panel). On the other hand, if the size of the ANN greatly exceeds the complexity required to describe the functional dependence of the data one might obtain an ANN that perfectly describes the training data but is unable to make reliable predictions for unseen data as it has not captured the overall functional dependence (right-hand panel). Reasoning along the same lines, it is not guaranteed that the ANN will provide reasonable predictions for data outside of the range of training data, even if the network structure was chosen properly. In other terms, ANNs tend to excel in interpolation tasks but extrapolation remains very challenging.

Finally, one is not guaranteed to find the global minimum of the loss function during training. In practice the loss function can become arbitrarily complicated with plenty of local minima or even multiple global minima. If the learning rate is comparably small the training is likely to end up in a local minimum, while a rather large learning rate might not allow to find a proper minimum at all. In early stages of the training process it can be advantageous to start with a rather high learning rate. This will prevent getting stuck in a small local minimum.

At later times one wants to reduce the step size of the optimization in order to achieve better accuracy. Hence, algorithms called learning rate schedulers can be employed that adapt the learning rate under specific circumstances.

Apparently there is a lot more to training an ANN than simply fitting a function. All of the additional adjustments we discussed come with their own parameters. These so-called hyperparameters need to be tuned carefully in order to achieve high quality results.

5

Artificial Neural Networks for the NCSM

As already mentioned, our goal is to construct an ANN that can extend the reach of modern many-body methods in particular the NCSM. We want to emphasize that, if employed in a "blackbox manner" as done here, extracting information from an ANN about the underlying physics is extremely challenging if not impossible. Therefore, all applications discussed here are solely focused on pushing computational boundaries and are to be understood as supplemental to actual physics-driven models. The particular limitation we want to tackle is the model-space truncation required to render the problem computationally feasible and the resulting error induced by it. Ideally, we would like to construct a tool that provides reliable predictions for the solution in the full Hilbert space based on calculations in small model spaces. In Sec. 3.2 we have discussed some established methods and their deficiencies including the lack of extrapolation methods beyond bound-state energies. We, therefore, aim to build an ANN that replaces phenomenological extrapolation functions and is capable of predicting converged values from the first few steps of a converging series. In particular, it should be constructed in such a way that it can be generalized to arbitrary observables.

In recent years there have been first applications of ANNs in this regard, which have demonstrated the potential of ANNs for this task and also some shortcomings [56, 57]. But, before we go into details, let us first take a closer look at the problem and motivate the use of ANNs.

Figure 5.1 shows a set of NCSM calculations for the ground-state energy of ^{3}H. For every $\hbar\Omega$ one finds a sequence of results obtained in successively increasing model-space truncations $N_{\max}$. By construction, all of these sequences converge towards the same limit for $N_{\max} \to \infty$. We emphasize that the results and, therefore, the convergence patterns depend on, both, the truncation parameter $N_{\max}$ and the HO frequency $\hbar\Omega$. The left-hand panel of the figure further shows classical extrapolations at $\mathcal{N}_{\max} = 8, 10$ and 12 using the procedure described in Sec. 3.2. Here, the model-space dependence is modeled by an exponential function Eq. (3.10) which is fitted for each $\hbar\Omega$ separately. Clearly, the actual convergence pattern is not exactly given by an exponential decay. This results in inconsistent predictions at different $\mathcal{N}_{\max}$ and also for different $\hbar\Omega$. It further indicates that we are looking for a more complicated function to fit to the data. However, the structure of this function is unknown. Hence, ANNs seem to be the ideal tool to use as they can approximate any function without prior information on

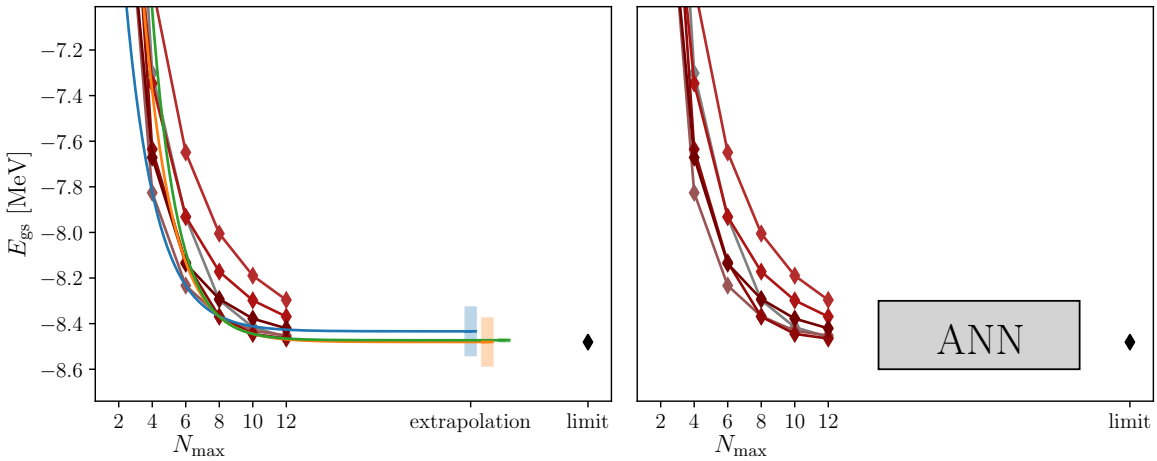

Figure 5.1: Typical set of sequences of NCSM calculations for the ground-state energy of ^{3}H at different $\hbar\Omega = 14, 16, 20, 24, 28, 32$ MeV. In the left-hand panel classical extrapolations are added including data up to $N_{max} = 8, 10$ and 12 respectively. These extrapolations are going to be replaced by an ANN (right-hand panel).

its form, as long as sufficient amounts of training data are available. The conceptual idea is sketched in the right-hand panel of Fig. 5.1.

We, finally, mention some general aspects of the problem, which will become important for the subsequent discussion. The model spaces grow factorially with particle number A, this is of particular importance when it comes to availability of potential training data for heavier nuclei for which calculations are rarely accessible above $N_{max} = 10$ in the case of realistic NN+3N Hamiltonians. Moreover, while sequences for any $\hbar\Omega$ converge to the same limit, they do so at drastically different rates. For all considerations here, we will assume that we deal with values of $\hbar\Omega$ that are in the vicinity of, or ideally centered around, the variational minimum for the ground-state energy of the system.

The work presented in this chapter and the following one has led to these publications [122, 123].

5.1 Network Design

To begin with, it should be noted that there is no unique way to approach the model-space extrapolation for NCSM calculations. Different approaches are conceptually different and require ANNs that are adapted to the specific features or requirements of the respective approach.

When it comes to designing an ANN the first thing that needs to be addressed is the network topology along with the structure of the data samples, as they define the size of the input layer and the output layer. The model-space extrapolation with ANNs has first been proposed by Negoita et al. [56], where they have used an ANN to approximate the function $O(\hbar\Omega, N_{max})$ for any observable O, hence they have emulated a mapping that attempts to predict the result of an NCSM calculation for a given input of $\hbar\Omega$ and N_{max}. This can already be achieved with a fairly small ANN consisting of an input layer with two neurons, one for $\hbar\Omega$ and one for N_{max}, one hidden layer with eight neurons and one output neuron yielding

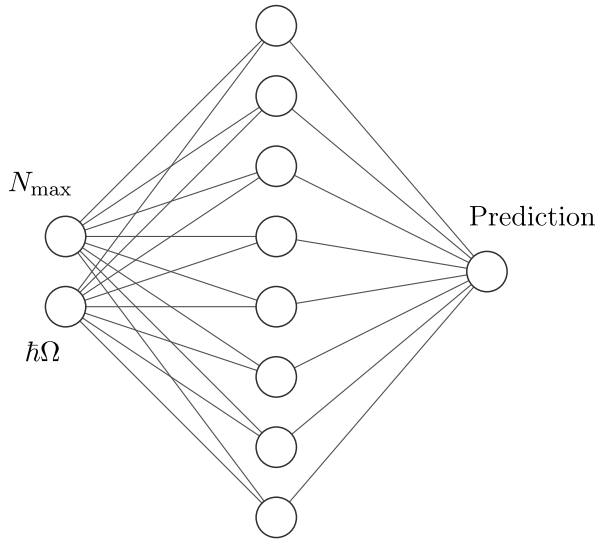

Figure 5.2: Topology of an ANN as employed in [56].

the prediction for the observable under investigation (see Fig. 5.2). Together with a Sigmoid activation function it can be proven that a single hidden layer is sufficient to approximate any continuous function [124]. Such a network can directly be trained on available NCSM data without requiring any preprocessing. A prediction for the converged value can then be obtained from predictions of the ANN for very large values of $N_{\max}$. By training multiple ANNs one can extract a statistical uncertainty from the multitude of predictions. They were able to obtain good results for ground-state energies and proton root-mean-square (rms) radii for ^{6}Li based on NCSM data up to $N_{\max} = 18$ with the predictions at different $\mathcal{N}_{\max}$ being approximately in agreement with each other. However, some deficits and potential bottlenecks of this ANN topology are forseeable. On one hand, the ANNs can only be trained for one specific combination of nucleus, interaction, and state and have to be retrained for every new case. Especially in heavier nuclei this can easily result in a shortage of training data due to the aforementioned scaling of model-space size with particle number. In addition, this approach seems to require calculations in rather large model spaces in order to provide precise and consistent predictions, which become computationally unfeasible as soon as a realistic Hamiltonian with 3N force is employed. On the other hand, the distribution of predictions exhibits a multimodal structure in some cases resulting in multiple predictions or unclear uncertainties (see Fig. 5 in [57]). While the latter can be remedied through a balancing of the training data and the inclusion of additional correlations as discussed in [57], issues regarding insufficient training data and, thus, limited applicability remain. Moreover, this application of ANNs resembles a true extrapolation beyond the training data, a task which is known to be extremely challenging for ML applications [125].

We, therefore, propose a different ansatz published in [122, 123] that essentially converts the extrapolation problem into an interpolation problem, thus facilitating the task for an ANN. The ansatz further exhibits a certain universality as ANNs are not trained for specific cases but can be applied to any combination of p-shell nuclei and realistic interactions. The idea is to directly predict the actual converged value for a given observable from a set of NCSM calculations in different finite model spaces instead of predicting the result for a given model space characterized by $N_{\max}$ and $\hbar\Omega$. This eliminates the need to explicitly model

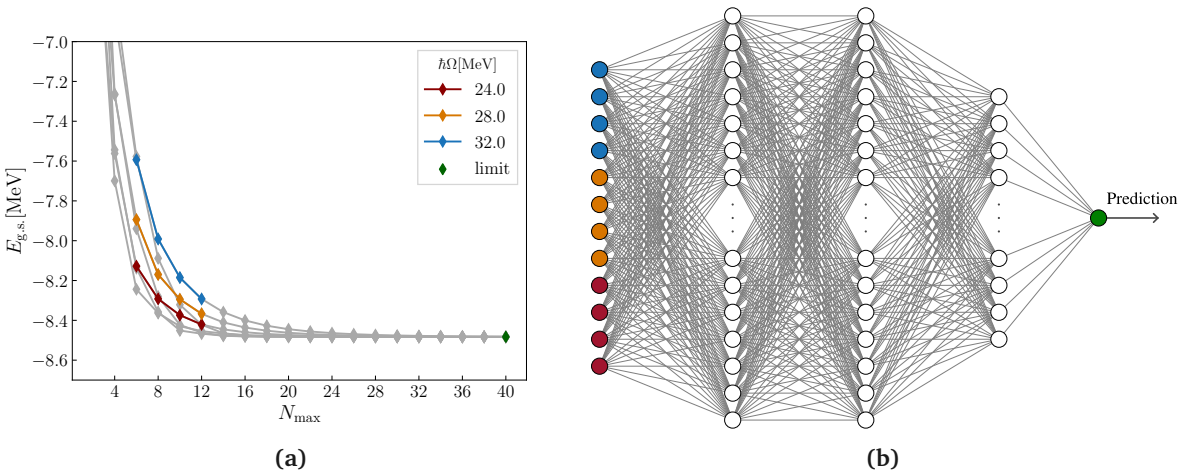

Figure 5.3: Set of NCSM calculations for the ground-state energy of ^{3}H (a) and a visualization of an FCFF ANN topology with three hidden layers (b). The colored data and neurons indicate input and output of the ANN in ABS mode (see text for details).

the convergence behavior and instead allows the ANN to treat the problem as a pattern recognition task, thus, predicting the converged value from the convergence pattern spanned by the available data. The data used as input for a prediction consists of F sequences of NCSM calculations for different $\hbar\Omega$ with length L, i.e., L consecutive values of $N_{\max}$. Note that L significantly limits the applicability to heavier p-shell nuclei as calculations can only be performed for very few values of $N_{\max}$. The topology of the ANN is then given by an input layer that has $F \cdot L$ neurons and an output layer that is a single neuron providing the prediction for the converged value. The number and size of the hidden layers can be chosen arbitrarily and we will discuss different topologies later on. A schematic representation of an ANN with $F = 3$ and $L = 4$ with three hidden layers is shown in Fig. 5.3b. Complementary, Fig. 5.3a indicates how the data is piped into the ANN, as the colored sample corresponds to a possible input to the first layer of the network. In the following, we will discuss different types of data formatting, i.e., different ways to pipe the data into the ANN, which we refer to as input modes.

The ABS Mode. The intuitive choice is feeding the raw NCSM output into the network, which is called ABS mode. Hence, the ANN effectively resembles a mapping

$$M_{\mathrm{ABS}} : \mathcal{S}_{\mathrm{ABS}} \longmapsto O^{\infty}, \tag{5.1}$$

where O^{∞} is the converged value for observable O and $\mathcal{S}_{\mathrm{ABS}}$ is an input sample of the shape

$$\begin{aligned} \mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}} = \Big(& O_{\hbar\Omega_1}^{\mathcal{N}_{\max}-2(L-1)}, \ldots, O_{\hbar\Omega_1}^{\mathcal{N}_{\max}-2}, O_{\hbar\Omega_1}^{\mathcal{N}_{\max}}, \\ & O_{\hbar\Omega_2}^{\mathcal{N}_{\max}-2(L-1)}, \ldots, O_{\hbar\Omega_2}^{\mathcal{N}_{\max}-2}, O_{\hbar\Omega_2}^{\mathcal{N}_{\max}}, \\ & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ & O_{\hbar\Omega_F}^{\mathcal{N}_{\max}-2(L-1)}, \ldots, O_{\hbar\Omega_F}^{\mathcal{N}_{\max}-2}, O_{\hbar\Omega_F}^{\mathcal{N}_{\max}} \Big) \end{aligned} \tag{5.2}$$

with $\mathcal{N}_{\max}$ being the largest $N_{\max}$ in the sample while $O_{\hbar\Omega}^{N_{\max}}$ is the observable obtained for HO frequency $\hbar\Omega$ and model-space truncation $N_{\max}$. This input mode holds the advantage that the data does not require any preprocessing and the ANN directly predicts the converged observable. However, especially binding energies cover a wide range of values which can become problematic for the ANN.

The DIFF Mode. This input mode is designed to reduce the dependence on the absolute value compared to the ABS mode. Instead of looking at the raw NCSM output, the differences between two consecutive $N_{\max}$ steps are used as inputs for the ANN. Hence, we refer to this mode as DIFF mode. Note that the network does not carry information on the absolute value of the observable anymore. Therefore, it cannot predict the converged value directly. Instead, we will have it predict the difference between the mean of all input values at $\mathcal{N}_{\max}$ and the converged value, i.e.,

$$\Delta_O^\infty = O^\infty - \mathrm{mean}(O_{\hbar\Omega}^{\mathcal{N}_{\max}}) \quad , \quad \Delta_E^\infty = E^\infty - \min(E_{\hbar\Omega}^{\mathcal{N}_{\max}}). \tag{5.3}$$

such that the converged value can be recovered from the output by $O^\infty = \mathrm{mean}(O_{\hbar\Omega}^{\mathcal{N}_{\max}}) + \Delta^\infty$. For energies this difference is obtained with the minimum instead of mean as they obey a variational principle. We end up with a mapping

$$M_{\mathrm{DIFF}} : \mathcal{S}_{\mathrm{DIFF}} \longmapsto \Delta_O^\infty \tag{5.4}$$

with input sample

$$\begin{aligned}\mathcal{S}_{\mathrm{DIFF}}^{\mathcal{N}_{\max}} = \Big(&\Delta_{\hbar\Omega_1}^{\mathcal{N}_{\max}-2(L-2)}, \dots, \Delta_{\hbar\Omega_1}^{\mathcal{N}_{\max}-2}, \Delta_{\hbar\Omega_1}^{\mathcal{N}_{\max}}, \\ &\Delta_{\hbar\Omega_2}^{\mathcal{N}_{\max}-2(L-2)}, \dots, \Delta_{\hbar\Omega_2}^{\mathcal{N}_{\max}-2}, \Delta_{\hbar\Omega_2}^{\mathcal{N}_{\max}}, \\ &\quad\vdots \qquad\qquad\quad \vdots \qquad\quad \vdots \\ &\Delta_{\hbar\Omega_F}^{\mathcal{N}_{\max}-2(L-2)}, \dots, \Delta_{\hbar\Omega_F}^{\mathcal{N}_{\max}-2}, \Delta_{\hbar\Omega_F}^{\mathcal{N}_{\max}}\Big),\end{aligned} \tag{5.5}$$

where we have defined $\Delta_{\hbar\Omega}^{N_{\max}} := O_{\hbar\Omega}^{N_{\max}} - O_{\hbar\Omega}^{N_{\max}-2}$. Note that compared to the ABS mode the effective length of the input sequences decreases such that $L_{\mathrm{DIFF}} = L - 1$ and the size of the input layer shrinks accordingly.

The MINMAX Mode. In order to reduce the scale dependence of the data even further and facilitate the training of the ANN, we introduce a third input mode referred to as MINMAX mode. This mode is based on the minmax normalization technique that is commonly employed in ML applications. In general, the input data is scaled to the interval $[a, b]$. We choose this interval such that the data is normalized to $[0, 1]$. A given sample in the ABS mode $\mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}}$ can then be normalized via

$$\mathcal{S}_{\mathrm{MINMAX}}^{\mathcal{N}_{\max}} = \frac{\mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}} - \min\left(\mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}}\right)}{\max\left(\mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}}\right) - \min\left(\mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}}\right)}. \tag{5.6}$$

Similar to the DIFF mode we can recover the converged value O^∞ from the networks output $O^\infty_{\mathrm{MINMAX}}$ through the inverse transformation

$$O^\infty = O^\infty_{\mathrm{MINMAX}}\left(\max\left(\mathcal{S}^{\mathcal{N}_{\max}}_{\mathrm{ABS}}\right) - \min\left(\mathcal{S}^{\mathcal{N}_{\max}}_{\mathrm{ABS}}\right)\right) + \min\left(\mathcal{S}^{\mathcal{N}_{\max}}_{\mathrm{ABS}}\right). \tag{5.7}$$

As the input to the ANN is always in the defined range no degrees of freedom in the network need to be used for an adjustment or scaling of the input data and the ANN will focus on learning the detailed features of the convergence pattern.

5.2 Training Data

Key to any good ML application is high-quality training data. As we aim to directly train on the converged value we rely on fully converged data sets of NCSM calculations. Unfortunately, those are only available for nuclei up to $A = 4$. Hence, our training set will be limited to the few-body systems ^{2}H, ^{3}H, and ^{4}He. One can question whether such a training set is suitable for predictions of other, in particular heavier, p-shell nuclei. Experience has shown that the convergence pattern for a given observable in nuclei with $A > 4$ is very much alike the one in the aforementioned few-body systems. This assumption is at the heart of this method and we will investigate its validity in more detail. However, there are some further arguments that solidify this assumption. First, the nuclei in the training set cover a broad range of convergence rates from rather slow (^{2}H) to very fast (^{4}He). This is in favor of the method due to the outstanding interpolation capabilities of ANNs. Following the same line of reasoning, one can further increase the quantity and quality of the training data by performing calculations of the training nuclei for multiple realistic interactions, which are additionally SRG evolved to multiple different flow parameters, as they also exhibit different convergence patterns. Here, the universality of this ANN method w.r.t. interaction can be exploited. However, we will only employ the Entem-Machleidt-Nosyk (EMN) family of interactions [85] in order to strictly separate training and evaluation data where the latter will be obtained with a different family.

In particular, the full set of training data consists of Jacobi-NCSM calculations for ^{2}H, ^{3}H, and ^{4}He up to $N_{\max} = 50, 40,$ and 24 respectively, for seven HO frequencies $\hbar\Omega = 12, 14, 16, 20, 24, 28,$ and 32 MeV. These calculations are performed with non-local chiral NN+3N interactions from Hüther et al. [78] at N^{2}LO, N^{3}LO, and N^{4}LO each with three different cutoffs $\Lambda_{\mathrm{C}} = 450, 500,$ and 550 MeV. All nine interactions are employed bare and SRG evolved to different flow parameters $\alpha = 0.02, 0.04,$ and 0.08 fm^4 resulting in a total of 36 different interactions or 756 converging sequences with different convergence patterns.

However, we exclude some of the data based on unphysical behavior or incomplete convergence. We, therefore, discard all results obtained at $N_{\max} = 0$ and exclude all sequences that are not converged within 5% of the limit at the largest available $N_{\max}$.

Since we use supervised learning we need to extract labels, i.e., the actual converged values, for the training data. As the data is basically fully converged to the accuracy we aim for, we extract the label for the ground-state energy from the lowest value at the highest $N_{\max}$, thus, always considering the optimal $\hbar\Omega$. For other observables we consider the exponential extrapolation of the flattest sequence regarding the six highest $N_{\max}$.

Next, the data needs to be converted into input samples. For each combination of nucleus and interaction we construct all subsets of F frequencies and L consecutive $N_{\max}$ values. We,

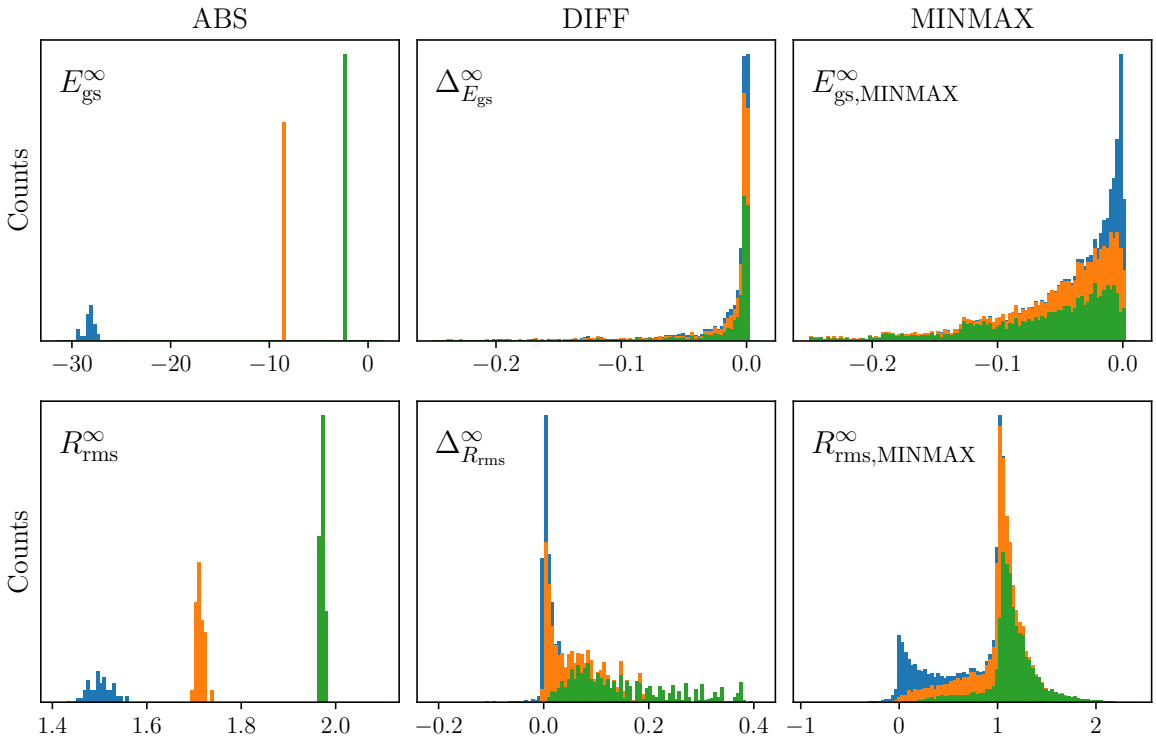

Figure 5.4: Distributions of target values in the training samples for ground-state energies (upper panels) and rms mass radii (lower panels) for all three input modes. The different colors indicate contributions from the different training nuclei ^{2}H (green), ^{3}H (orange), and ^{4}He (blue).

Nucleus	Samples E_{gs}	Samples R_{rms}
^{2}H	166320	166320
^{3}H	126990	126990
^{4}He	60642	62262

Table 5.1: Numbers of training samples per nucleus for ground-state energy and rms radius. Differences arise from the different N_{max} ranges for the individual nuclei.

further, consider all permutations of the HO frequencies in order to become independent of the order in which they are piped into the ANN. Finally, the samples themselves can be formatted according to the input mode.

As a result we obtain 353,952 (355,572) energy (radius) samples and the distributions of their target values are shown in Fig. 5.4. Table 5.1 shows the corresponding numbers of samples per training nucleus. For the ABS mode we see three peaks emerge, which correspond to the three training nuclei. The distributions for the DIFF and MINMAX modes appear much smoother and more widely spread. Moreover, the distributions for the individual nuclei show strong overlap and very similar shape, which indicates nucleus-independent features. For the ground-state energies the MINMAX distribution is more evenly spread over the interval, which is expected to be beneficial over the high peak structure we find for DIFF. Both exhibit a sharp edge at 0 which is due to the variational principle. For the radii, on the other hand, we find rather different distributions for DIFF and MINMAX. In the DIFF case there is again a single peak structure around 0 from the almost fully converged samples.

Other than that most targets are above 0 and only very few below, which indicates that the training set continues much more upward converging sequences for radii than downward converging ones. For MINMAX we have a double-peak structure with peaks at the boundaries of the interval we are scaling the data to, i.e. 0 and 1. The targets above 1 again resemble mostly upward converging samples (mostly $^2\mathrm{H}$ and $^3\mathrm{H}$) and the few values below 0 reflect the imbalance compared to downward converging ones ($^4\mathrm{He}$). Moreover, with the MINMAX mode we can also resolve samples with mixed convergence given by the part of the distribution between 0 and 1.

Let us reconsider the training set for the ABS mode. The structure of the distributions indicate that the ANN either has to learn how to deal with possible input values that span a range of values in the training and an even wider range in subsequent applications to p-shell nuclei or, in the worst case, it will end up learning to reproduce three specific values for the training nuclei which would resemble a certain kind of overfitting w.r.t. the nucleus. We can try to facilitate this task for the ANN by artificially spreading the training set over a wider range of values. This is done by randomly scaling the samples. Our target nuclei are p-shell nuclei which exhibit a range of about 0 to -150 MeV in binding energy. We, therefore, ideally spread the training set over the same interval. As the DIFF samples also exhibit a certain range dependence we scale the samples for both ABS and DIFF such that

$$a \cdot \mathcal{S}_{\mathrm{ABS}}^{\mathcal{N}_{\max}} + b \quad \text{and} \quad c \cdot \mathcal{S}_{\mathrm{DIFF}}^{\mathcal{N}_{\max}}, \tag{5.8}$$

where $a \in (0.25, 4)$, $b \in (-20, 20)$, and $c \in (0.5, 2)$ are random numbers drawn from a uniform distribution over the given intervals. This way we can obtain an arbitrary number of samples that span a wider energy range and slightly differ in their convergence pattern. We can apply the same scaling for radii. However, they exhibit a different range and we, therefore, choose $b \in (-1, 1)$ as negative or very large radii are unphysical. Figure 5.5 shows the new distributions of the target values after scaling. We find that the target values in the ABS mode now spread the desired energy range and similarly a wider range of radii is covered. The DIFF distributions appear rather unchanged, however they now also include samples that cover much larger values.

While this concludes the construction of the sets of training data used in the following applications, we have not yet discussed, whether the initial assumption about the training data resembling actual use cases holds. Figure 5.6 shows a comparison of evaluation samples constructed from NCSM data for $^{16}\mathrm{O}$ with a small subset of the training samples. For better comparability we look at the normalized MINMAX samples that serve as direct input to the ANNs. We find for both, ground-state energies and rms radii, that the $^{16}\mathrm{O}$ samples lie well within the training data. Most importantly the overall trend of the sequences matches and one is not to be concerned by minor derivations due to the excellent interpolation qualities of ANNs. For clarity, this is only presented for one sample each, however, this holds for most if not all samples for other p-shell nuclei as well. Hence, we can conclude that the convergence pattern in p-shell nuclei appears to be universal and the few-body systems considered for training are, therefore, sufficient to provide training data of good quality.

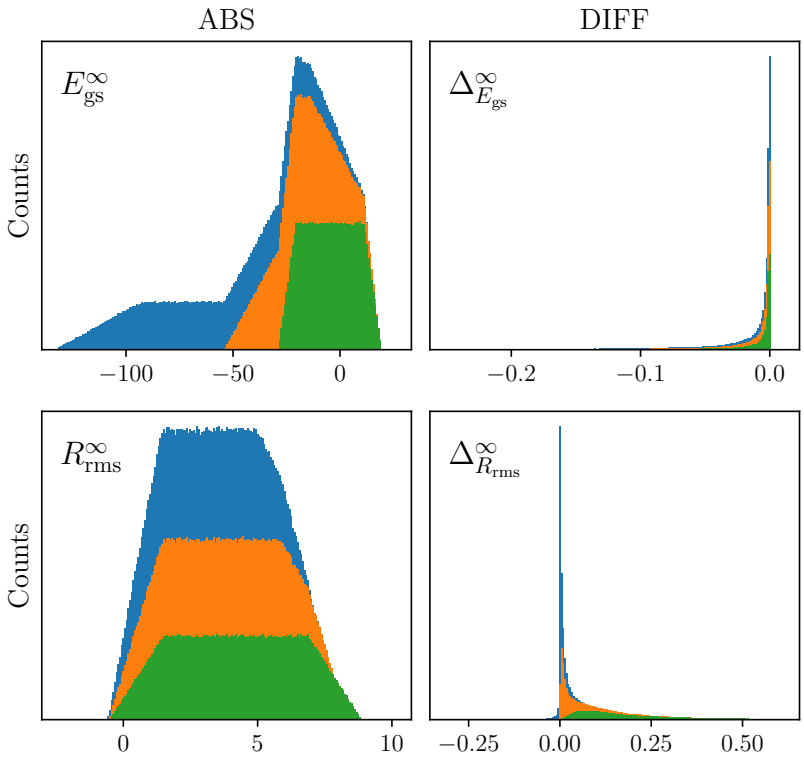

Figure 5.5: Same as Fig. 5.4 but for the scaled samples.

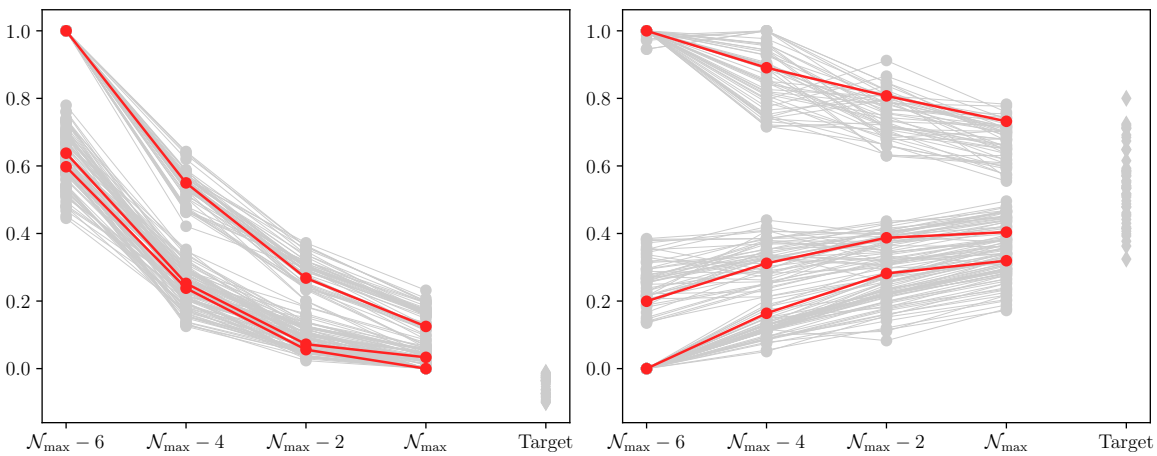

Figure 5.6: Evaluation sample for ^{16}O (red) embedded in the 50 nearest training samples (gray) for ground-state energies (left-hand panel) and rms radii (right-hand panel) in MINMAX mode. The samples are of similar shape, hence, the convergence patterns in heavier p-shell nuclei are very much alike the ones in the training data.

5.3 Training

With the training data prepared, we can turn to the actual training of the ANNs. For that, we divide the available data into three disjoint subsets, i.e., a training set, which is used for the actual training and consists of 1,000,000 samples, a development set made from 10,000 samples, which is evaluated multiple times during the training in order to adjust parameters such as the learning rate, and finally a validation set consisting of 50,000 samples on which the quality of the fully trained ANN will be evaluated. Note that the given numbers of samples are obtained through random scaling for ABS and DIFF, while the samples for MINMAX are randomly drawn from the original set of samples, which results in the same samples occurring multiple times.

The overall procedure for the training is given as follows:

1. Iterate through the whole training set in batches of size N_{batch}.
   Such an iteration is called an epoch.

2. Evaluate the development set and adjust the learning rate.

3. Repeat 1. and 2. for a given number of epochs.

4. Evaluate the validation set and keep or discard the ANN.

Since each of these steps contains hyperparameters that need to be chosen in advance or adapted during the training, we will discuss them in more detail. Note that the adjustment of hyperparameters is an extensive iterative process that mostly builds on experience. We will, therefore, not go through the whole process, which would require the discussion of hundreds of combinations of those hyperparameters, but instead start with the configuration that worked best in applications and show how a variation of the individual hyperparameters affects the training process.

We define the following as the starting point for our investigations and will refer to it as the standard configuration: For the ANN topology we first of all choose an input of three sequences ($F = 3$) with four consecutive N_{max} values each ($L = 4$), which results in an input layer with 12 neurons. We further choose three hidden layers with $4 \cdot F \cdot L$, $4 \cdot F \cdot L$, and $8 \cdot F$ neurons respectively. The size of the output layer remains at one single neuron. As activation function we employ a rectified linear unit (ReLU) [126] and for backpropagation we use the AdamW algorithm [127]. Our loss function of choice is the MSE loss function as given in Eq. (4.2). For the training process we start with an initial learning rate of $\eta = 0.001$, which is adapted by a learning rate scheduler that reduces the learning rate by 50% if the change in loss after a training epoch is less than 10% compared to the previous epoch. The training is performed with a batch size of 512, i.e., the ANN is trained on 512 samples simultaneously.

Let us first take a look at the training performance for the standard configuration. In Fig. 5.7 the average loss for 10 ANNs over a period of 50 epochs is shown for all input modes. Apparently, the DIFF and MINMAX trainings converge faster than the ABS mode, which aligns with our assumption of facilitating the training through normalization of the data. Since we want to avoid the risk of overfitting the data, we stop the training after 20 epochs, indicated by the gray dashed line, just before the loss is converged. Training a single ANN then takes approximately 80 seconds on a desktop computer.

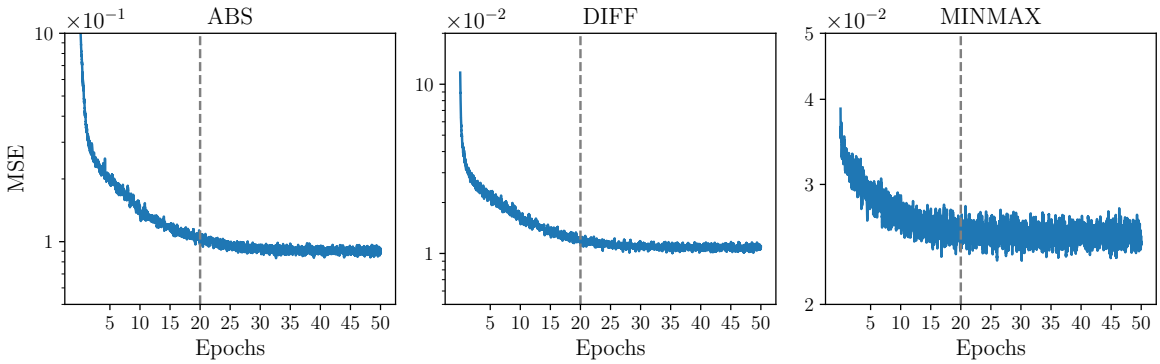

Figure 5.7: Change of loss on the development data set during the training process for the standard configuration for ground-state energies averaged over 10 ANNs. The dashed grey line at 20 epochs is chosen as the stopping point for the training in actual applications. The different panels correspond to different input modes.

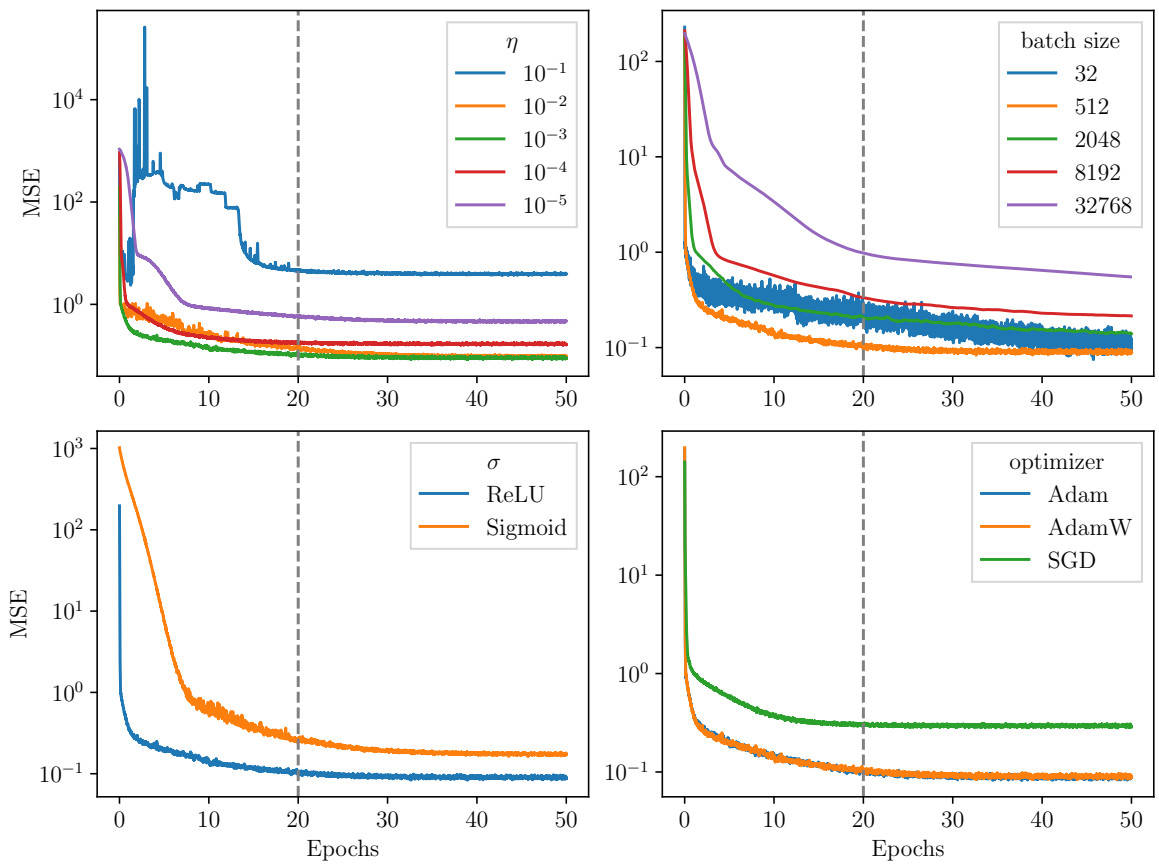

Figure 5.8: Same as Fig. 5.7 but for variations of different hyperparameters: initial learning rate (upper left), batch size (upper right), activation function (lower left), and optimizer (lower right).

We can further investigate the dependence on the initial learning rate η, which is shown in the upper left panel of Fig. 5.8. Again, the average loss for 10 ANNs in ABS mode is shown, this time for initial learning rates at various orders of magnitude. We find that both $\eta = 0.01$ and $\eta = 0.001$ converge toward a similar level of accuracy on the development set, however, $\eta = 0.001$ converges faster and performs much better at epoch 20 in particular. If the learning rate is increased even further the step size in the optimization process becomes too large to be meaningful and the exhibits almost random behavior before approaching a high-lying local minimum eventually, probably due to the learning rate scheduler. Going to significantly smaller learning rates the training is slower and the ANNs also end up in local minima due to the small step size.

The upper right panel of Fig. 5.8 shows variations of the batch size, i.e., the number of samples that are piped into the network simultaneously. We find the best training for a batch size of 512. For a batch size of 32 the loss converges toward a similar limit however it does so much slower and the training time increases to around 280 seconds. Any larger batch sizes seem to end up in local minima or converge far more slowly while the training time remains at around 80 to 90 seconds.

Regarding the activation function (lower left panel of Fig. 5.8), we find superior convergence for ReLU over a Sigmoid. Finally, the lower right panel shows that the Adam optimization algorithms perform far better than stochastic gradient descent (SGD). Note that for the latter the learning rate had to be adjusted to $\eta = 0.00001$ which we have found to be optimal for this algorithm.

All of the hyperparameters above have been optimized for the network topology that we have found to work best in our applications. However, all of these may or may not change if the number of hidden layers or their size is changed. Finding the right topology can be time consuming as it is an iterative and rather experimental process. We will, therefore, refrain from a detailed discussion but want to mention that the performance of the ANNs discussed here is quite robust against changes of the topology and under-(over-)fitting only occurs for drastically smaller (bigger) numbers of nodes. Between these two extremes the topology will have a significant impact on the performance of the training, however, the final predictions remain of similar quality.

As a last step in the training process, we need to address the validation of an ANN once it is trained. The decision whether to keep or discard an ANN is based on two criteria. One one hand, we define a validation threshold, i.e., a target accuracy for the ANN on the validation data. This threshold varies based on observable and input mode. The validation thresholds employed in this work are given in Tab. 5.2. On the other hand, we can look for indicators of overfitting by comparing the ANNs performance on the training set and the validation set. A significantly smaller loss on the training data is a strong sign of overfitting.

Mode	E_{gs}	R_{rms}
ABS	0.05	0.01
DIFF	0.01	0.001
MINMAX	0.05	0.1

Table 5.2: Validation thresholds for the ANN training for different observables and input modes.

5.4 Statistical Evaluation

The remaining questions are how to employ trained ANNs to unseen data, which we will refer to as evaluation data from now on, and how to extract a robust prediction along with a reliable uncertainty estimate. Again, we first need to format the evaluation data into samples that match the design of the input layer of the ANN. This can be done analogously to the preparation of the training data discussed beforehand and results in a multitude of evaluation samples that all yield different, equally plausible predictions when passed through the ANN. We can sort those by $\mathcal{N}_{\max}$ in order to assess predictions for different model-space sizes separately, yet this still leaves us with multiple predictions at a given $\mathcal{N}_{\max}$. Moreover, it is difficult to quantify the quality and associated uncertainty for the prediction of a single network. We will, therefore, train multiple ANNs and turn to a statistical evaluation process. In particular, we employ up to 1000 valid ANNs that have all been initialized randomly, train them individually, and evaluate them with all possible evaluation samples. As shown in Fig. 5.9, the resulting distributions of predictions are reminiscent of a normal distribution, hence, we can extract the final prediction and statistical uncertainty by fitting a Gaussian to the distribution of predictions and extracting the mean and standard deviation. Thus, the estimated uncertainty can be interpreted as a 68% or 1σ degree of believe interval. We can further see how the number of ANNs contributes to the stability of the statistical evaluation. The leftmost panel shows the predictions obtained from a single ANN, hence, the spread of predictions arises from the different evaluation samples that can be constructed from the evaluation data. It also serves as a rough estimate on how large the uncertainty from this source is. Looking at the other panels that show the same distribution for an increasing number of ANNs we find that the prediction and uncertainty remain very stable, which means that the ANNs are in very good agreement with each other and the main uncertainty arises from the different evaluation samples. Yet, the distributions become smoother and grow closer to a normal distribution, which results in a more robust fit. Overall we see some fluctuations of the predictions for small numbers of ANNs which eventually saturate at around 100 ANNs. To ensure minimal dependence on statistical errors we will employ 1000 ANNs for actual predictions.

In some cases the distribution of predictions does not resemble a normal distribution but exhibits multimodality to different extends, as illustrated in Fig. 5.10. Consequently, this leads to a breakdown of the Gaussian fit method. As already discussed in [57], the occurrence of such multi-peak structures can be attributed to an imbalance in the evaluation data, which can be avoided through a thorough selection or rebalancing of the data. Either way, it serves

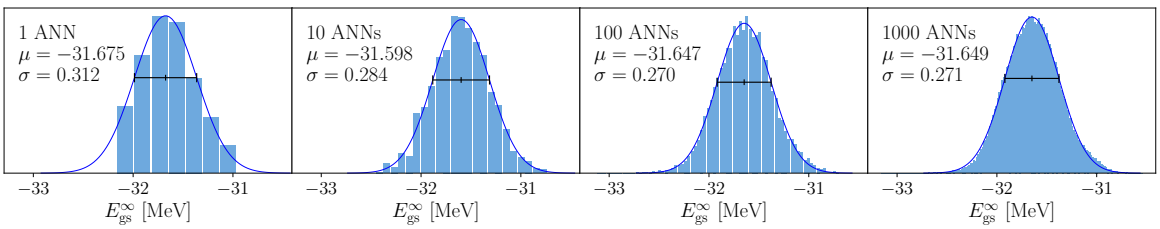

Figure 5.9: Distributions of predictions for the ground-state energy of ^{6}Li at $\mathcal{N}_{\max} = 8$ along with fitted Gaussian, obtained from different numbers of ANNs in ABS mode.

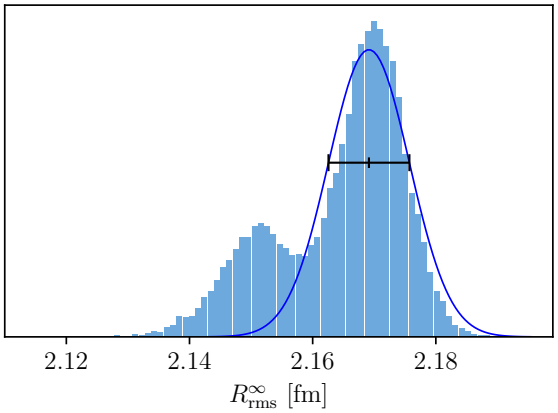

Figure 5.10: Distributions of predictions for imbalanced evaluation data for the rms radius of $^{16}\mathrm{O}$ at $\mathcal{N}_{\mathrm{max}} = 10$ along with fitted Gaussian, obtained from 1000 ANNs in MINMAX mode.

as a strong indicator for unwanted biases in the data. In our case, such structures only occur for very sparse sets of evaluation data.

6

Predictions of Energies and Spectra

In the previous chapters we have discussed how to construct and train ANNs for NCSM model-space extrapolation. With the trained networks on hand we can now investigate their performance and predictive power. Results presented here have led to publication [122]. If not stated otherwise, all predictions shown here are obtained from 1000 ANNs.

In order to avoid the reproduction of training data and ensure that the evaluation data is truly unseen by the ANNs, we strictly separate the evaluation data from the training data by using another family of interactions. In addition, this will automatically proof the universal applicability of the developed ANN topology. The interaction family for evaluation is the semi-local momentum space regularized set of chiral NN + 3N interactions developed by the LENPIC collaboration [128, 129]. Those are available up to N^{4}LO on the two-body level and N^{2}LO on the three-body level and will be employed for the cutoff values $\Lambda_\mathrm{C} = 450, 500$ MeV.

6.1 Benchmark in Few-Body Systems

Before we look at actual applications of the ANNs we first quantify the performance of the ANNs on data for which we know the converged values, in order to compare different ANN setups. This requires a set of benchmark data along with a metric that determines the accuracy across multiple predictions. The benchmark data consists of Jacobi-NCSM calculations for ^{2}H, ^{3}H, and ^{4}He with semi-local interactions at N^{2}LO and N^{3}LO in the NN sector accompanied by a 3N interaction at N^{2}LO [52], each for both cutoff values $\Lambda_\mathrm{C} = 450, 500$ MeV and SRG evolved to $\alpha = 0.04, 0.08$ fm^4, resulting in 8 different interactions. For every interaction, data for $\hbar\Omega = 12, 14, 16, 20, 24, 28,$ and 32 MeV is available and we will look at the performance of the ANNs at $\mathcal{N}_\mathrm{max} = 8, 10$ and 12.

Since the different nuclei exhibit very different binding energies, we estimate the total accuracy a on a relative measure given by

$$a = 1 - \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - y_i^\mathrm{t})^2}{(y_i^\mathrm{t})^2}}, \tag{6.1}$$

where y_i and y_i^t are the predicted value and the target value respectively and the sum runs over all samples.

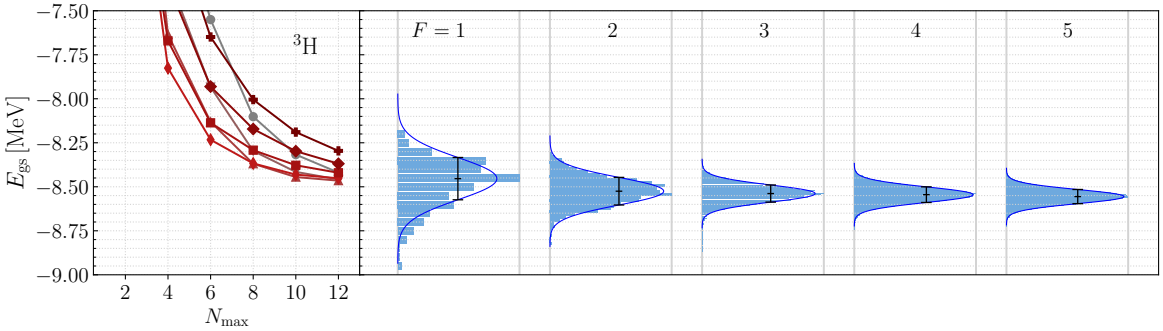

Figure 6.1: Predictions of the ground-state energy of ^{3}H at $\mathcal{N}_{max} = 10$ for different number of input sequences F for 100 ANNs in ABS mode each.

Nucleus	$\mathcal{N}_{max}$	$F = 1$	$F = 2$	$F = 3$	$F = 4$	$F = 5$
^{2}H	8	0.9259	0.9750	0.9752	0.9802	0.9819
	10	0.9667	0.9821	0.9907	0.9902	0.9844
	12	0.9915	0.9954	0.9956	0.9962	0.9982
^{3}H	8	0.9943	0.9994	0.9978	0.9959	0.9946
	10	0.9889	0.9916	0.9928	0.9918	0.9904
	12	0.9969	0.9974	0.9969	0.9969	0.9960
^{4}He	8	0.9944	0.9981	0.9977	0.9981	0.9976
	10	0.9958	0.9978	0.9978	0.9983	0.9978
	12	0.9984	0.9986	0.9985	0.9988	0.9983
Total	8	0.9570	0.9855	0.9856	0.9883	0.9890
	10	0.9796	0.9885	0.9931	0.9926	0.9893
	12	0.9947	0.9969	0.9968	0.9971	0.9973

Table 6.1: Accuracy of 100 ANNs in ABS mode on the benchmark set for different numbers of input sequences F.

Durning the previous discussion of hyperparameters we have chosen $L = 4$, since we want the ANNs to be able to predict a wide range of p-shell nuclei for which we are often limited to $N_{max} = 8$. The choice of the number of input sequences, however, is not constraint by such considerations, as the computational cost only scales linearly with the number of frequencies. To get an idea of the dependence of the predictions on F, we can discuss Fig. 6.1, which shows predictions for the ^{3}H ground-state energy for different F. While this is only an example, we already see that a single input sequence significantly reduces the stability of the statistical evaluation. The shape of the histogram and the extracted prediction improves with increasing F up to $F = 3$ and remains stable after. In order to avoid any bias from a single case, we can additionally consider Tab. 6.1, which shows the accuracy on the benchmark set for the different ANN topologies at $\mathcal{N}_{max} = 8, 10,$ and 12. Similar to the previous example we find overall increasing accuracy up to $F = 3$. For higher numbers of input sequences the accuracy at $\mathcal{N}_{max} = 8$ keeps improving at the cost of the $\mathcal{N}_{max} = 10$ accuracy. We, therefore, choose $F = 3$ to be optimal as it allows high quality predictions with a rather small amount of input data. Moreover, it should be noted that we reach an accuracy on the benchmark set above 99%.

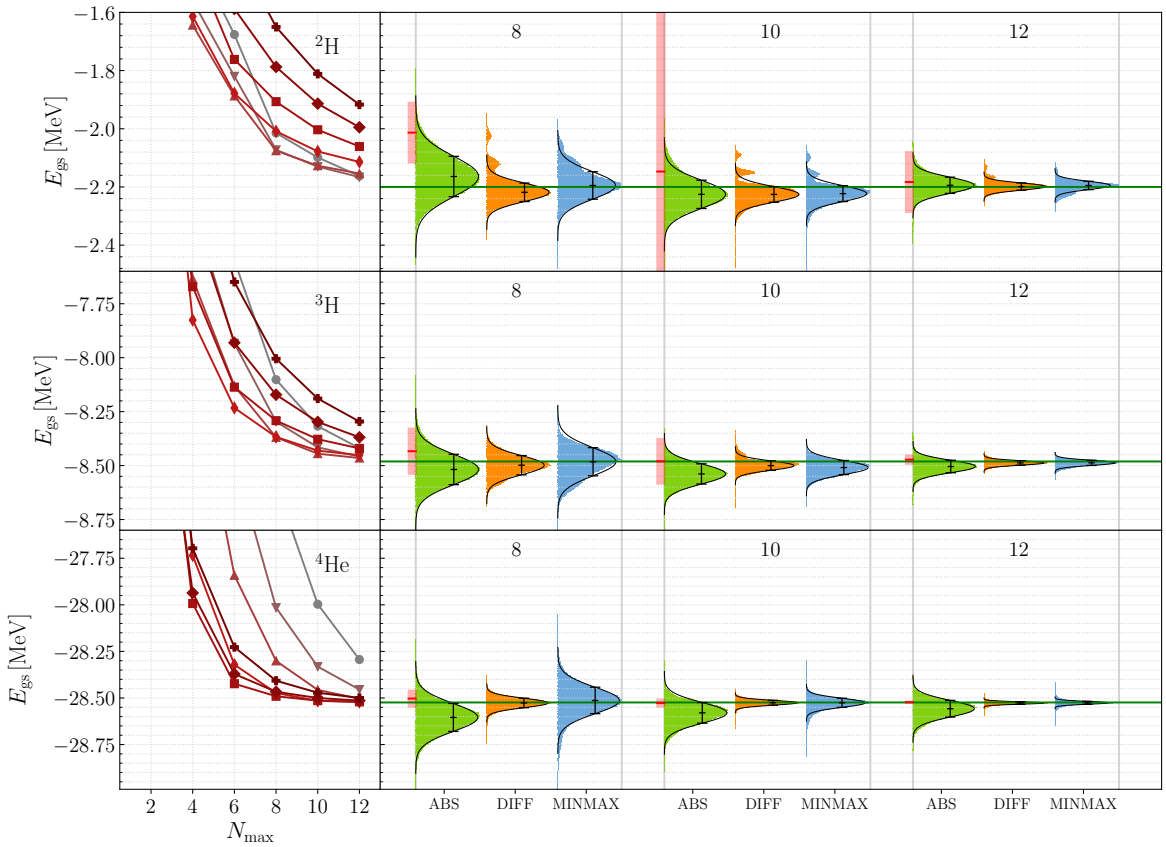

Figure 6.2: Evaluation data (left) and ANN predictions (right) for the ground-state energies of ^{2}H, ^{3}H, and ^{4}He at N^{2}LO with $\Lambda_C = 450$ MeV and $\alpha = 0.08$ fm^4. The evaluation data consists of NCSM calculations for $\hbar\Omega = 12, 14, 16, 20, 24, 28,$ and 32 MeV (gray to red) while the predictions are grouped in three sections corresponding to the respective $\mathcal{N}_{max} = 8, 10$ or 12. The colored histograms with Gaussian fits correspond to the different input modes ABS (green), DIFF (orange), and MINMAX (green). The darker green horizontal lines indicate the converged value for the respective nucleus. Classical extrapolations according to the procedure described in Sec. 3.2 are given in red with red error bands.

With the topology set, we can now investigate how the ANNs perform on the selected few-body systems. Figure 6.2 shows the predictions of the ground-state energies for the three nuclei but with the semi-local interaction at N^{2}LO with $\Lambda_C = 450$ MeV and SRG evolved to $\alpha = 0.08$ fm^4. The left-hand panels show the NCSM calculations used as evaluation data, while the right-hand side depicts histograms of predictions for the ABS (green), DIFF (orange), and MINMAX (green) modes at different $\mathcal{N}_{max}$. The darker green lines additionally indicate the actual converged ground-state energies for the respective nucleus and interaction. Classical extrapolations according to the procedure described in Sec. 3.2 are given in red with red error bands.

First of all, we notice that all predictions are in the vicinity of the respective target values. Hence, the ANNs are able to capture the convergence pattern and produce reasonable predictions in the first place. We, further, find that especially for ^{2}H and ^{3}H the predictions are more accurate and in the case of the DIFF and MINMAX mode also more precise than the classical extrapolations. Comparing the different input modes we see that DIFF comes with the smallest uncertainties while being very accurate. MINMAX has a similar level of

Nucleus	$\mathcal{N}_{\max}$	ABS	DIFF	MINMAX
^{2}H	8	0.9775	0.9887	0.9852
	10	0.9900	0.9909	0.9937
	12	0.9962	0.9926	0.9953
^{3}H	8	0.9972	0.9984	0.9988
	10	0.9927	0.9986	0.9953
	12	0.9970	0.9993	0.9990
^{4}He	8	0.9976	0.9999	0.9961
	10	0.9977	0.9996	0.9977
	12	0.9985	0.9997	0.9991
Total	8	0.9869	0.9934	0.9911
	10	0.9927	0.9947	0.9953
	12	0.9971	0.9957	0.9972

Table 6.2: Accuracy of 1000 ANNs on the ground-state energy benchmark set for the different input modes.

accuracy but larger uncertainties for smaller $\mathcal{N}_{\max}$. The ABS mode lacks both accuracy and precision compared to the other modes which was expected from the previous discussion of the input modes. We can verify these findings based on the accuracy on the benchmark set given in Tab. 6.2, for which MINMAX seems slightly superior. Note that the accuracy measure does not take the uncertainties into account and might, therefore, be misleading. However, the uncertainties are intentionally omitted in the measure as we do not know whether larger uncertainties are more realistic.

6.2 Application to p-Shell Nuclei

So far, we have demonstrated that the ANNs are able to capture the convergence pattern for the systems they are trained on, even for a different family of interactions. We can now extend these investigations to unseen p-shell nuclei, which resemble actual applications and allow us to validate the assumption of universality as discussed previously. Figure 6.3 shows ANN predictions for the p-shell nuclei ^{6}Li, ^{7}Li, ^{8}Li, ^{9}Be, and ^{12}C across a wider mass range than the few-body systems. Again, all three input modes have been evaluated and classical extrapolations are given for comparison. First of all, the ANNs in all modes produce reasonable predictions comparable to intuitive expectations or classical extrapolations. Compared to the latter, the ANN predictions yield uncertainties of similar or even smaller size depending on the input mode. Considering how well the classical extrapolation works for ground-state energies this proofs that the ANN extrapolation does work at least equally well.

When comparing the different input modes in Fig. 6.3, we observe indications of an energy dependence for the ABS or DIFF mode. For ^{6}Li, ^{7}Li, ^{8}Li, and ^{9}Be we find that the MINMAX predictions are very consistent across different $\mathcal{N}_{\max}$ while the results for ABS and DIFF, just like the classical extrapolations, exhibit a downward trend. In the DIFF case, which yields the smallest uncertainty estimates, this downward trend comes with partially inconsistent predictions. Further, this downward trend is a potential indicator for an energy dependence,

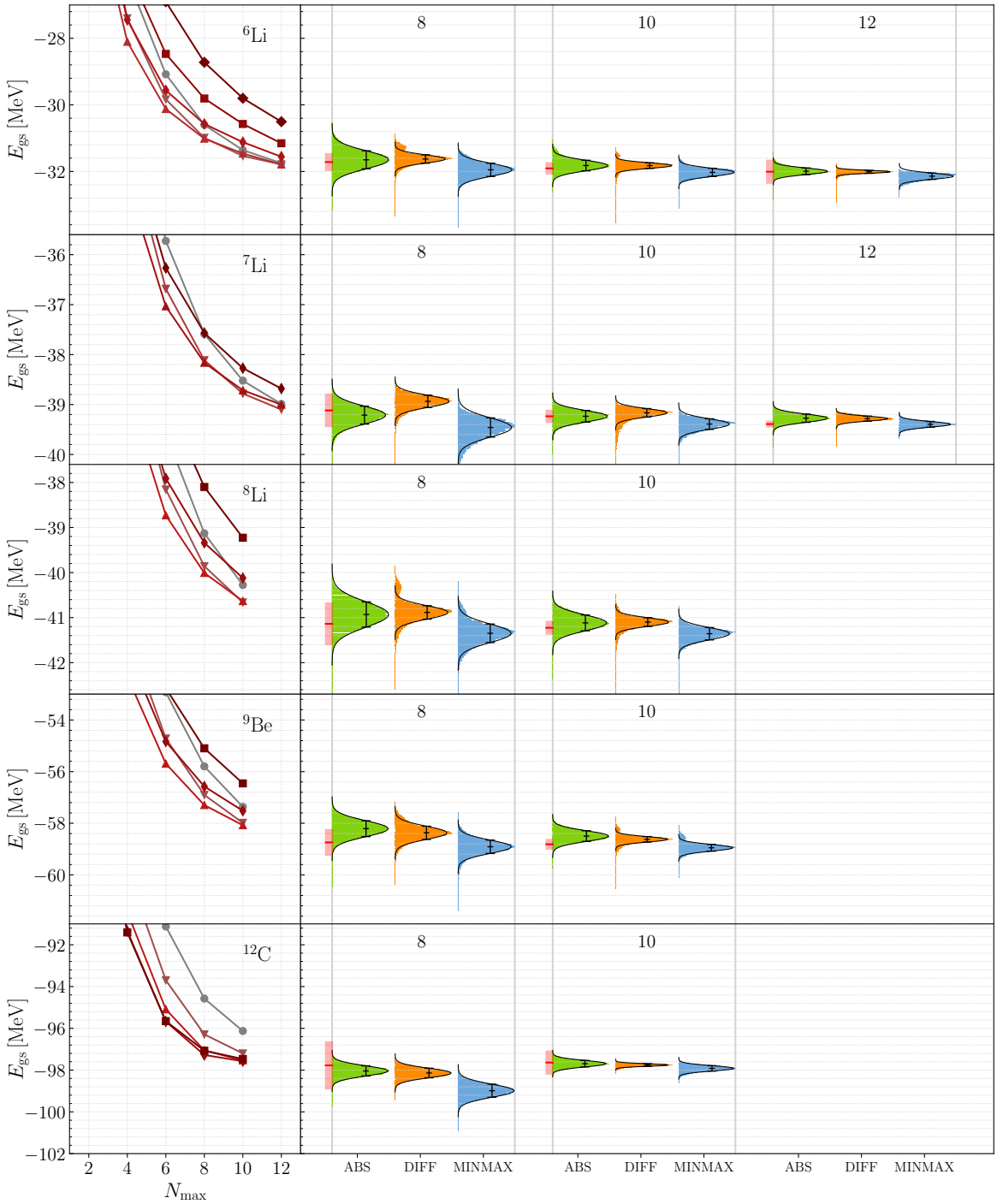

Figure 6.3: Same as Fig. 6.2 but for selected p-shell nuclei ^{6}Li, ^{7}Li, ^{8}Li, ^{9}Be, and ^{12}C.

which is expected to manifest as a shift towards the energy regime covered in the training. Besides normalization schemes this energy dependence should also be significantly lowered by looking at the binding energy per nucleon E/A instead, which is rather consistent across the whole nuclear chart. Since this does not require any conceptual change to the ABS mode, we will refer to this as ABS' meaning ANNs that are trained on E/A. A comparison of these two modes is shown in Fig. 6.4, where we find that the energy dependence is indeed reduced but at the cost of significantly larger uncertainties. The latter can most likely be attributed to the rescaling required to obtain the full ground-state energy. Overall, ABS' performs very much like the MINMAX mode.

Note that we have limited our discussion to $^{6}\mathrm{Li}$, $^{7}\mathrm{Li}$, $^{8}\mathrm{Li}$, and $^{9}\mathrm{Be}$ so far. Surprisingly, the situation is quite opposite for $^{12}\mathrm{C}$, where predictions for all modes exhibit an upward trend resulting in very inconsistent results for the MINMAX mode and rather stable predictions for ABS and DIFF. Upon closer inspection, we have found that this is potentially an artifact of the strong importance truncation staggered across multiple N_{max} steps, which is required to render the $^{12}\mathrm{C}$ model space at $N_{\mathrm{max}} = 10$ tractable. However, to confirm this a more detailed investigation is required.

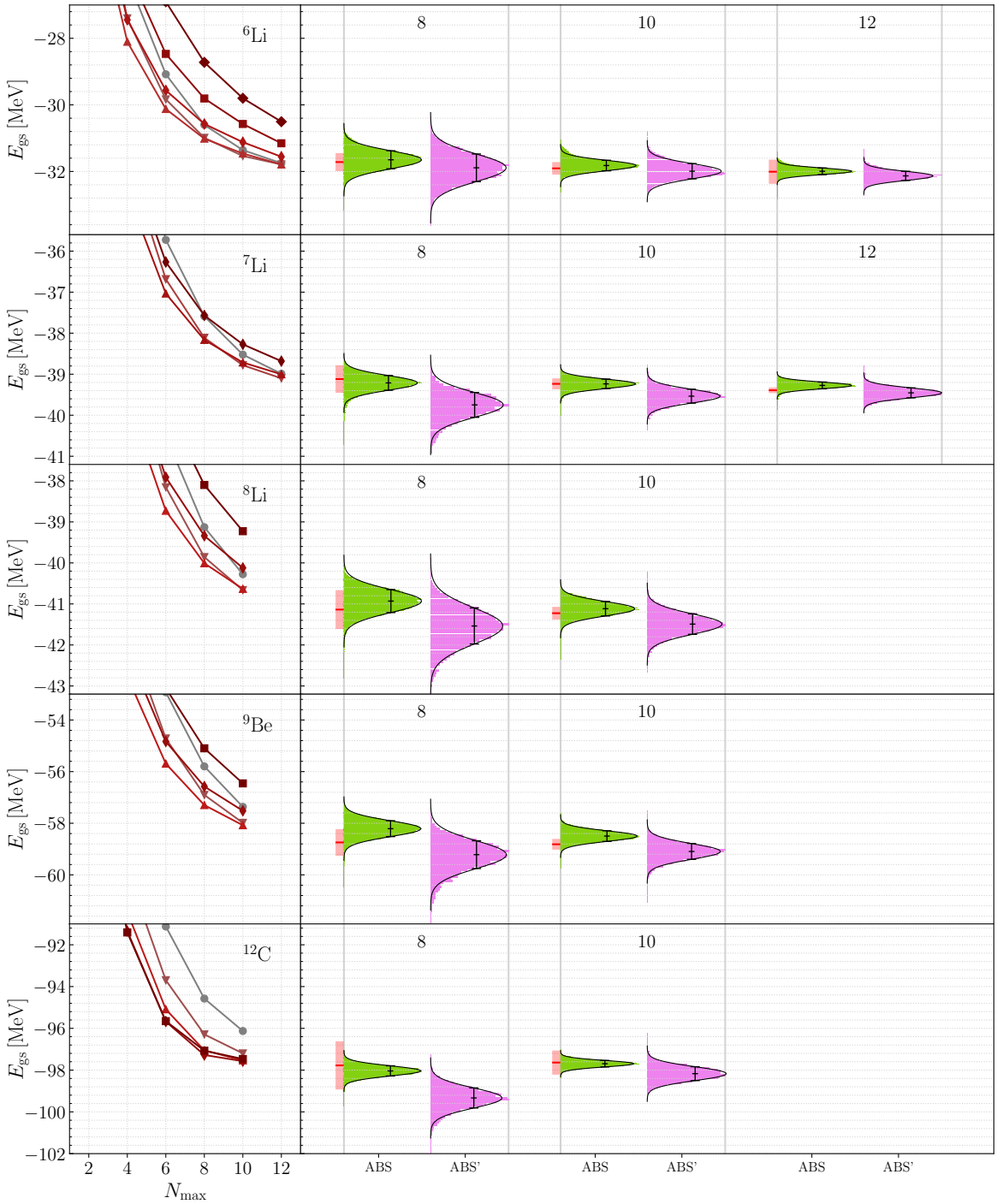

Figure 6.4: Same as Fig. 6.3 but comparing ABS and ABS' input modes (see text for details).

6.3 Extension to Excitation Energies

The previous investigations in this chapter have proven the universality of the ANNs extrapolation capabilities w.r.t. the binding energy of the nucleus and the employed interaction. This leaves us with one final step towards the prediction of full energy spectra, which is the extension to excitation energies. The developments of this extension have led to the publication [123].

Studying the convergence pattern for energies of excited states, we find that they are very much like those for ground-state energies, at least for bound states, which we are interested in. Hence, we can start by simply using the previous ANNs to predict the converged energies of excited states. Due to their interpolation capabilities, we find that the ANNs work equally well for excited states as they did for ground states, as can be seen in the upper panels of Fig. 6.5. Both, the ground-state energy and the energy of the first 3^+ excited state of ^{6}Li have been predicted with the ANNs in MINMAX mode. While this adds to the universality aspect of the ANNs, the observable we are actually interested in is the excitation energy E_{ex}. The straight-forward way to obtain those is a simple subtraction of the extrapolated energies for ground state and excited state. However, providing a meaningful uncertainty is complicated as the uncertainty estimates of the individual energies are strongly correlated. If we recede to the false assumption that both uncertainties are uncorrelated an estimate can be obtained through standard uncertainty propagation. In the given example, this would yield an excitation energy of $E_{\text{ex}} = 2.391(82)$ for $\mathcal{N}_{\text{max}} = 12$. Alternatively, we can extend the idea

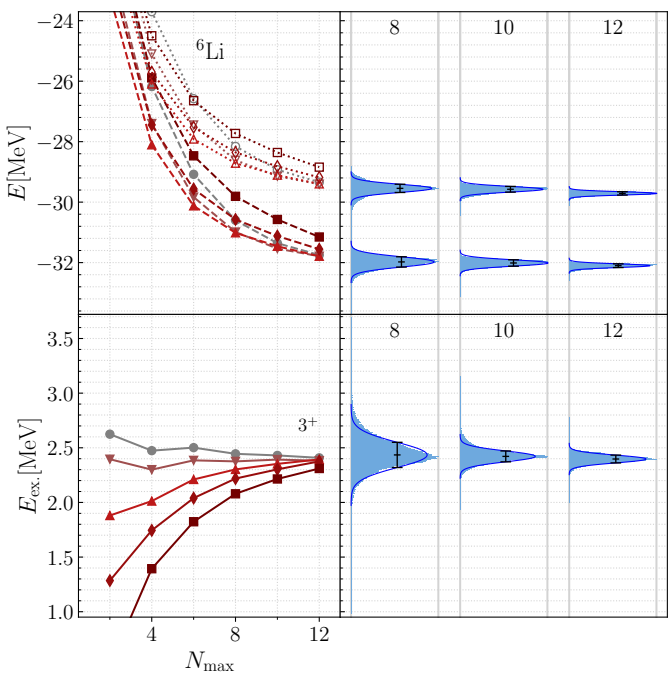

Figure 6.5: Upper panels: NCSM calculations for $\hbar\Omega = 14, 16, 20, 24,$ and 28 MeV and extracted predictions of the ground-state energy (dashed) and first 3^+ excited state energy (dotted) of ^{6}Li from ANNs in MINMAX mode at $\mathcal{N}_{\text{max}} = 8, 10,$ and 12. Lower panels: Corresponding excitation energy with prediction obtained through sample-wise subtraction of the individual predictions (see text for details).

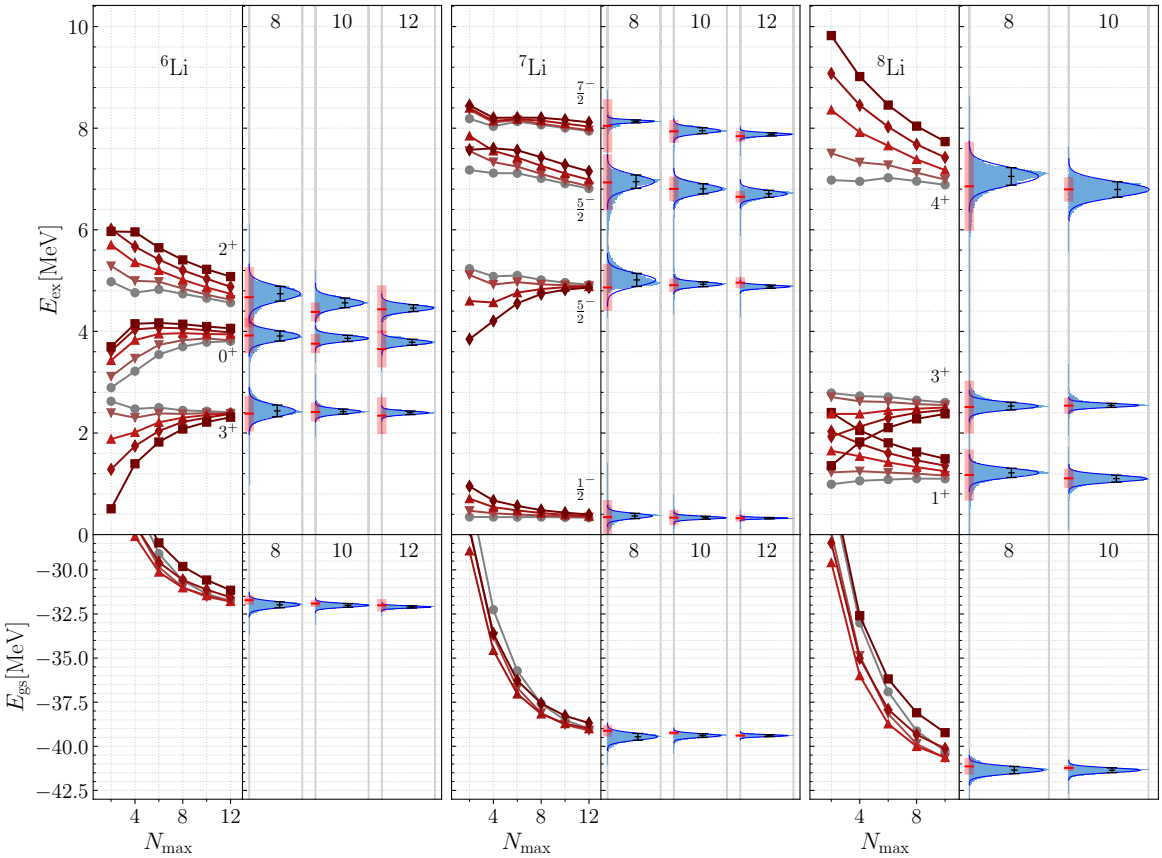

Figure 6.6: NCSM calculations of ground-state and excitation energies for the low-lying natural parity states of ^{6}Li, ^{7}Li, and ^{8}Li along with ANN predictions in MINMAX mode and classical extrapolations.

of the statistical evaluation by constructing a distribution of predictions for the excitation energy, from which we can then extract a more meaningful uncertainty. Such a histogram can be obtained through a sample-wise subtraction of the predictions for the absolute energies of both states. The idea is to construct pairs of evaluation samples for the two different states, which consist of data for the same set of $N_{\max}$ and $\hbar\Omega$. We can then have each ANN predict a converged value for both samples and subtract them obtaining a single estimate for the excitation energy. Doing this for all possible evaluation samples and all ANNs we end up with a distribution as depicted in the lower panels of Fig. 6.5. The predictions are remarkably consistent and for the final prediction at $\mathcal{N}_{\max} = 12$ we find $E_{\text{ex}} = 2.397(36)$, which is about twice as precise than the standard uncertainty propagation while the predictions are in agreement even though the extracted prediction is slightly shifted.

We now have the toolbox to study predictions of energy spectra. Figure 6.6 shows predictions for the ground-state energy and the excitation energies of low-lying natural parity states for ^{6}Li, ^{7}Li, and ^{8}Li along with ANN predictions starting at $\mathcal{N}_{\max} = 8$ together with classical extrapolations for comparison. The latter are obtained through the difference of the extrapolations for the two individual states, where the larger uncertainty is chosen. Across all nuclei and states we find that the ANNs produce very consistent extrapolations with the exception of the $\frac{7}{2}^-$ state in ^{7}Li. Moreover, the predictions of excitation energies are much more accurate that those from classical extrapolations and different convergence patterns are handled well.

Overall, we have shown that ANNs constructed and trained as discussed in the previous chapter feature a great degree of universality regarding different nuclei, states and interactions and can easily compete with classical extrapolation methods. Since they are rather small and only have to be trained once, they provide an efficient tool for postprocessing NCSM calculations.

A collection of the numerical values for all results shown in this chapter can be found in Appendix A.1.

7

Predictions of Radii

In this chapter we will investigate to what extend the properties of the ANNs, which work exceptionally well for energies, can be transferred to another observable, e.g., rms radii. The results presented here have also entered the publication [123].

As we have already discussed, the main challenge here is the less constrained and, therefore, much more complicated convergence pattern that includes upwards and downwards converging as well as non-monotonic sequences. In addition, there are no established extrapolation methods for radii, hence, robust ANN predictions would be very beneficial for future investigations and precision calculations of nuclear radii.

For the following applications we have trained 1000 ANNs analogously to the energy case. Training and evaluation data are obtained with different families of interactions in order to be strictly separated. In particular, the evaluation data for radii is calculated with the semi-local momentum space regularized chiral NN + 3N interaction [128, 129].

7.1 Benchmark in Few-Body Systems

We start our investigations by studying the ANNs performance on few-body nuclei. In order to do so, we construct a set of benchmark data for the same chiral orders, cutoffs, SRG flow parameters and $\hbar\Omega$ as for the ground-state energies in Sec. 6.1. The accuracy of the predictions is, again, estimated by the measure defined in Eq. (6.1).

The performance on the benchmark nuclei for the semi-local interaction at N^{2}LO with $\Lambda_{\mathrm{C}} = 450$ MeV and $\alpha = 0.08$ fm^4 is illustrated in Fig. 7.1. Again, the evaluation data is displayed in the left-hand panels while the right-hand panels show predictions from 1000 ANNs for ABS (green), DIFF (orange), and MINMAX (blue) grouped by $\mathcal{N}_{\mathrm{max}}$. The darker green lines indicate the converged values for the respective nuclei.

One notices right away that the performance of the ANNs is worse and far less consistent across the different input modes compared to the ground-state energies (see Fig. 6.2). This can, in general, be attributed to the more complex convergence behavior. However, looking at the individual input modes we see remarkable differences in the accuracy of the predictions. Starting with ABS we find that $^4\mathrm{He}$ as the simplest case is described reasonably well, while $^2\mathrm{H}$ and $^3\mathrm{H}$ are clearly under- or overestimated. This is even worse for the DIFF mode, where all three nuclei are strongly overestimated. The normalization in the MINMAX mode seems to

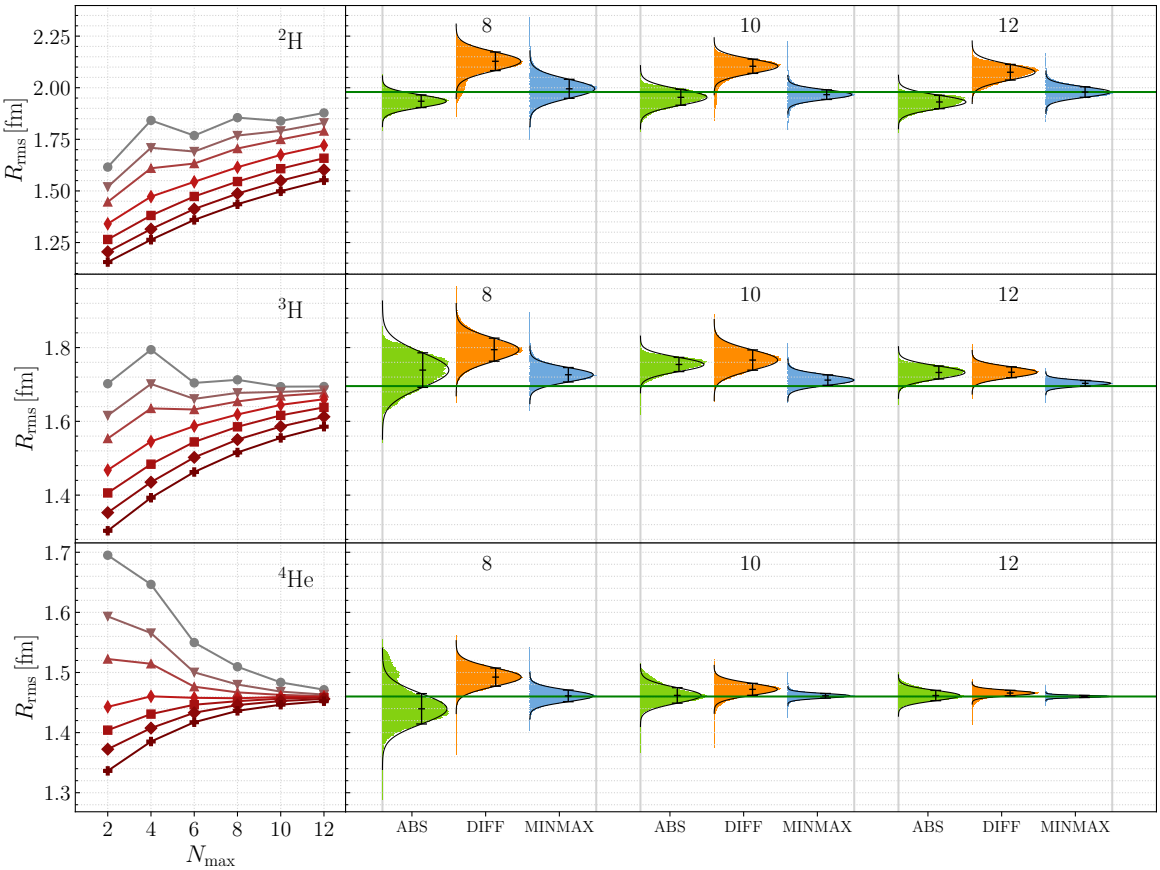

Figure 7.1: Evaluation data (left) and ANN predictions (right) for the mass rms-radii of ^{2}H, ^{3}H, and ^{4}He at N^{2}LO with $\Lambda_{\mathrm{C}} = 450$ MeV and $\alpha = 0.08$ fm^4. The evaluation data consists of NCSM calculations for $\hbar\Omega = 12, 14, 16, 20, 24, 28,$ and 32 MeV (gray to red) while the predictions are grouped in three sections corresponding to the respective $\mathcal{N}_{\max} = 8, 10$ or 12. The colored histograms with Gaussian fits correspond to the different input modes ABS (green), DIFF (orange), and MINMAX (green). The darker green horizontal lines indicate the converged value for the respective nucleus.

Nucleus	$\mathcal{N}_{\max}$	ABS	DIFF	MINMAX
^{2}H	8	0.9743	0.9305	0.9937
	10	0.9811	0.9399	0.9923
	12	0.9691	0.9521	0.9918
^{3}H	8	0.9803	0.9418	0.9799
	10	0.9696	0.9492	0.9897
	12	0.9826	0.9755	0.9966
^{4}He	8	0.9916	0.9734	0.9967
	10	0.9958	0.9872	0.9976
	12	0.9980	0.9922	0.9987
Total	8	0.9807	0.9454	0.9877
	10	0.9792	0.9540	0.9925
	12	0.9795	0.9686	0.9948

Table 7.1: Accuracy of 1000 ANNs on the mass radius benchmark set for the different input modes.

remedy these issues, describing ^{2}H and ^{4}He almost perfectly from small $\mathcal{N}_{\max}$ on, while only exhibiting minor deviations for ^{3}H. Apparently, normalizing the data facilitates the learning process of the convergence patterns, allowing the ANNs to capture more complex patterns. We further notice that the predictions seem to benefit from input data converging from both directions, which is reasonable as it provides an additional constraint on the converged value.

These observations are supported by the accuracy measures on the benchmark set given in Tab. 7.1. We find that all input modes perform best on ^{4}He. Both, ABS and DIFF, show deficits in the description of ^{2}H and ^{3}H, whereas MINMAX is consistent and comes with the highest overall accuracy. From this we conclude that the normalization of the data is not only beneficial but necessary and we will, therefore, continue our investigations with the MINMAX mode only.

7.2 Application to p-Shell Nuclei

While the ANNs succeed in describing the light nuclei, the question remains whether the learned convergence patterns transfer to heavier p-shell nuclei. Figure 7.2 shows predictions of mass rms-radii for ^{6}Li, ^{7}Li, ^{8}Li, and ^{9}Be, which resemble prime examples of interesting p-shell nuclei. In all cases we find robust predictions that are fairly consistent across multiple $\mathcal{N}_{\max}$. On closer inspection, one notices a slight upward trend in the predictions of ^{6}Li and ^{7}Li, which correspond to similar trends in the input data. However, this mainly results in deficiencies in the smallest model spaces and predictions beyond those are consistent again. Alternatively, we would expect this effect to decrease when more calculations at smaller $\hbar\Omega$ are taken into account, as they should result in downward converging series, thus, adding additional information akin to the previous ^{4}He convergence pattern.

Overall, we can conclude that the ANNs can indeed be extended to radii and provide robust predictions with reasonable uncertainty estimates. This allows for precise calculations of radii across the p-shell. Moreover, the universality of the ANNs does not only yield good predictions for mass radii but also other radii, e.g., point-proton radii. We can further extend the difference-based framework for excitation energies and look at differences of radii, which is subject to ongoing developments, as they directly relate to experimental efforts.

In principle, the ANN extrapolation tool can also be applied to other observables such as electromagnetic moments or transitions, but the few-body systems we have trained on exhibit little to none electromagnetic moments. Hence, other means of obtaining training data or the exploitation of correlations between observables are required for future applications along those lines.

A collection of the numerical values for all results shown in this chapter can be found in Appendix A.2.

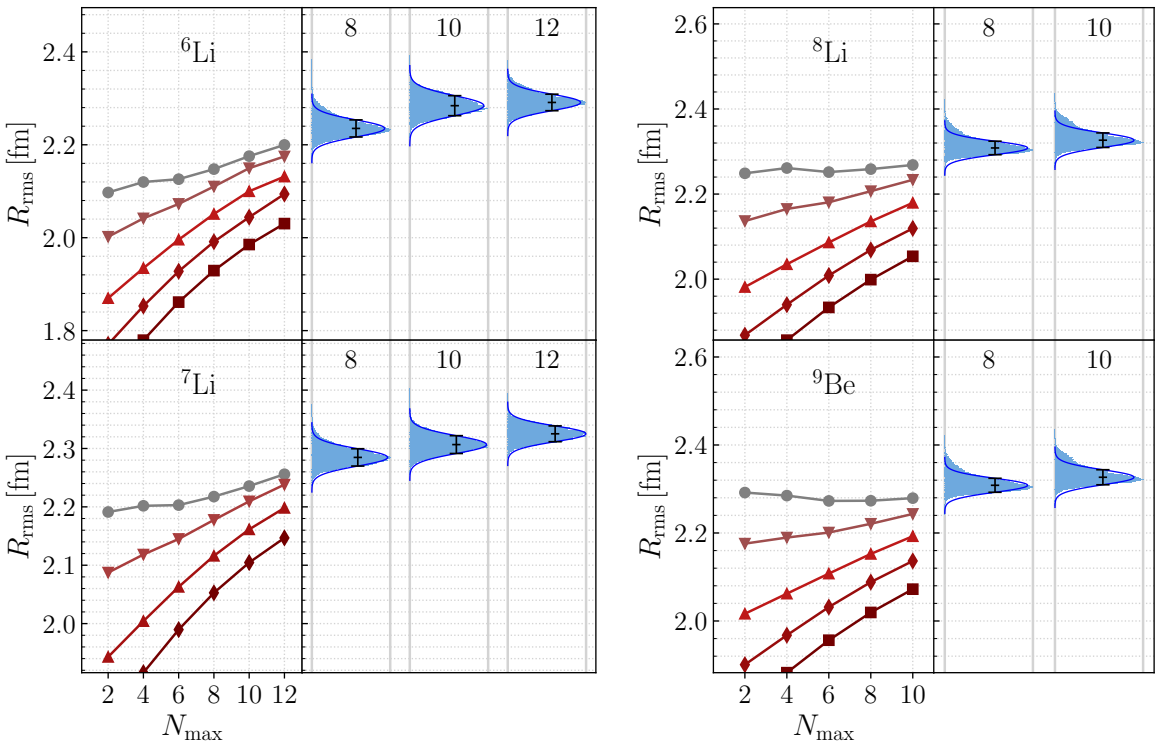

Figure 7.2: Evaluation data along with predictions from ANNs in MINMAX mode at accessible $\mathcal{N}_{\rm max}$ for the mass rms-radii of selected p-shell nuclei ^{6}Li, ^{7}Li, ^{8}Li, and ^{9}Be at N^{2}LO with $\Lambda_{\rm C} = 450$ MeV and $\alpha = 0.08$ fm^4 for $\hbar\Omega = 14, 16, 20, 24,$ and 28 MeV (gray to red). Evaluation data has been provided from [48].

Part III

Ab Initio Theory for p-Shell Hypernuclei

8

The Hypernuclear Hamiltonian

Since the advent of ab initio methods in nuclear structure theory together with HPC facilities, the nuclear chart has been explored up to the medium-mass regime and the quantitative understanding of nuclear structure and insights into nuclear interactions have enabled the field to enter a precision era. However, there is one aspect that has lately gained increasingly more attention in a strive to understand the fundamental aspects of the strong force. It is the field of hypernuclei and other systems with strangeness that are governed by the strong force [18].

Experimentally, hypernuclei have been known and successfully measured since the 1950s [6]. The theoretical description of these systems requires the shift from purely nucleonic approaches to a more general baryonic picture, introducing strangeness $\mathcal{S}$ as an additional degree of freedom. Consequently, hyperons (see Tab. 8.1), i.e., baryons with $\mathcal{S} \neq 0$, enter the particle content of the theoretical models. Hence, the Hamiltonian defining the hypernuclear many-body problem needs to be extended in order to describe these new particle species. In its general form

$$\hat{H}_{\text{int}} = \Delta\hat{M} + \hat{T}_{\text{int}} + \hat{V}_{\text{N}} + \hat{V}_{\text{Y}} \tag{8.1}$$

the intrinsic kinetic energy $\hat{T}_{\text{int}}$ and the nucleonic potential $\hat{V}_{\text{N}}$ remain unchanged and it mainly deviates from the nuclear Hamiltonian Eq. (1.2) by the addition of a potential term $\hat{V}_{\text{Y}}$ that covers the interaction among hyperons and nucleons. The additional mass term $\Delta\hat{M} = \hat{M} - M_0$ accounts for the different rest masses of the particles due to the comparably large mass of the hyperons. Here, $M_0 \equiv Zm_{\text{p}} + Nm_{\text{n}} - \mathcal{S}m_{\Lambda}$ is the reference point for the energy that corresponds to the rest mass of the constituents of the system under consideration. Note that the previous expression for $\hat{T}_{\text{int}}$ in equation Eq. (1.5) remains valid, however, it cannot be simplified further as the equal mass assumption for nucleons does not hold for the heavier hyperons.

The hyperonic interaction

$$\hat{V}_{\text{Y}} = \hat{V}_{\text{YN}} + \hat{V}_{\text{YY}} + \hat{V}_{\text{YNN}} + \hat{V}_{\text{YYN}} + \hat{V}_{\text{YYY}} + \dots\ , \tag{8.2}$$

can again be decomposed by particle rank and species up to A-baryon interactions. Analogously to the nucleonic case we truncate this at the three-body level and omit terms with higher particle rank. We, further, limit our investigations to singly-strange hypernuclei, which leaves us with hyperon-nucleon (YN) and hyperon-nucleon-nucleon (YNN) contributions.

Particle	Mass m [MeV]	Charge Q [e]	Isospin t	Isospin Projection m_t	Strangeness $\mathcal{S}$
n	939.565	0	1/2	−1/2	0
p	938.272	+1	1/2	+1/2	0
Λ	1115.683	0	0	0	−1
Σ^-	1197.449	−1	1	−1	−1
Σ^0	1192.642	0	1	0	−1
Σ^+	1189.37	+1	1	+1	−1

Table 8.1: Properties of nucleons and singly-strange hyperons that are constituents of the systems discussed in this work.

8.1 Baryonic Interactions

Analogously to the nucleonic case, the YN and YNN interactions required for the calculation of hypernuclei are rooted in QCD. While there have been some approaches to directly calculate hypernuclear observables via lattice QCD [130–132], mostly phenomenological interactions based on meson-exchange models such as the Nijmegen [133–135] and the Jülich [136, 137] potentials have been employed in structure calculations. In recent years, hyperonic interactions from chiral EFT have entered the field and allow for ab initio calculations of hypernuclei on the same footing as nucleonic interactions [19–22].

Constructing a baryon-baryon interaction, however, is all but simple. The concept of isospin that arises from the SU(2) symmetry in nucleonic chiral EFT has to be adapted to the SU(3) flavor symmetry required for the baryonic case. Hence, the nucleon and pion degrees of freedom are extended to the full octet baryons and the pseudoscalar mesons which are depicted in Fig. 8.1. We do not go into the details of the derivation of the interactions but some general remarks are in order. The SU(3) flavor symmetry clearly is broken by the mass of the strange quark, but it provides a good starting point for the chiral expansion and reduces the free parameters in the theory. Nevertheless, more LECs than in the nucleonic sector occur and need to be constrained on scattering data. Due to the short lifetime of the Λ hyperon of approx. 260 ps, scattering experiments are extremely challenging. Thus, there is very little scattering data available. We will address this issue in much more detail in Ch. 10. Further, the inclusion of baryon-baryon coupling terms, in particular terms that account for the Λ-Σ mixing, has proven to be crucial for the description of the spectra of singly-strange s-shell hypernuclei [138–141].

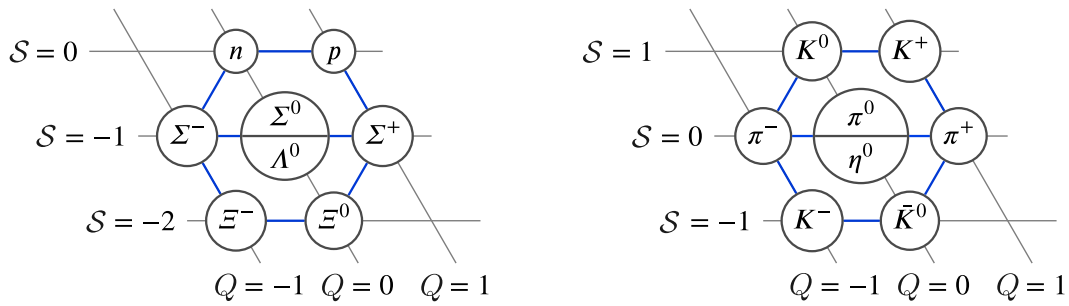

Figure 8.1: Octet baryons (left) and pseudoscalar mesons (right) that resemble the degrees of freedom in baryonic chiral EFT sorted by strangeness $\mathcal{S}$ and charge Q.

To date, YN interactions from chiral EFT have been derived up to N^2LO [22]. Initial YNN forces, which enter at N^2LO, have been derived partially [142], but a full description is not available for now. In this work we limit ourselves to LO YN interactions. At this order no YNN terms contribute. However, induced YNN terms enter again when nucleonic and hyperonic interactions are consistently SRG evolved [143]. The inclusion of these induced YNN terms has been found to be crucial for an accurate description of hypernuclei and hints to a potential solution of the hyperon puzzle in neutron stars [16, 17]. All calculations of systems with strangeness in this work are performed with induced YNN forces.

9

Hypernuclei in the No-Core Shell Model

In an attempt to solve the hypernuclear many-body problem, extensions of conventional ab initio methods are of particular interest. Throughout the past two decades, successes in the s-shell regime have been achieved with Faddeev-Yakubovsky calculations [36, 37], and Gaussian expansion methods [38]. More recently, p-shell hypernuclei have become accessible through extensions of the NCSM [42, 44], Quantum Monte Carlo approaches [39, 40], and nuclear lattice EFT calculations [45].

Our method of choice remains the NCSM and we will discuss the differences to the nucleonic case, following the work of Wirth et al. [41–43, 143–145]. The general setup of the method is the same as discussed in Ch. 2. The starting point is an expansion of the many-body state in Slater determinants that consist of HO single-particle states. The essential step when extending to systems with strangeness is conceptually simple, but has severe practical implications. It is sufficient to allow hyperons as constituents of the single-particle basis by introducing strangeness $\mathcal{S} \in \{-1, 0\}$ as an additional quantum number. Hence, compared to Eq. (2.4) a single-particle state is now characterized by

$$|a_i\rangle = \{|nlm\rangle \otimes |sm_s\rangle \otimes |\mathcal{S}tm_t\rangle\}_i. \tag{9.1}$$

Note that isospin t and its projection m_t can now take on more values as before according to Tab. 8.1. Due to these additional degrees of freedom, the model-space dimensions grow even more rapidly than before as shown in Fig. 9.1. We find that the model spaces for hypernuclei are about an order of magnitude larger compared to nucleonic systems with the same number of particles and nearly two orders of magnitude compared to their nucleonic parents.

When it comes to the application of intrinsic operators such as $\hat{T}_{\text{int}}$, which we have already discussed in the context of the hypernuclear Hamiltonian, the equal mass assumption becomes invalid. Hence, we need to explicitly keep track of the single-particle masses, and the expressions for $\hat{H}_{\text{cm}}$ or $\hat{R}^2_{\text{ms}}$ cannot be simplified beyond Eq. (2.7) and Eq. (2.9) respectively.

Other than that, the truncation scheme remains unchanged and we can exploit the same symmetries as in the nucleonic case in order to reduce the model-space sizes. Similarily, the idea of importance truncation can be transferred to the hypernuclear sector. This, however, requires an adjustment of the perturbative importance measure, again, to account for the different rest masses of the nucleons and hyperons. In particular, we need to redefine Eq. (2.14)

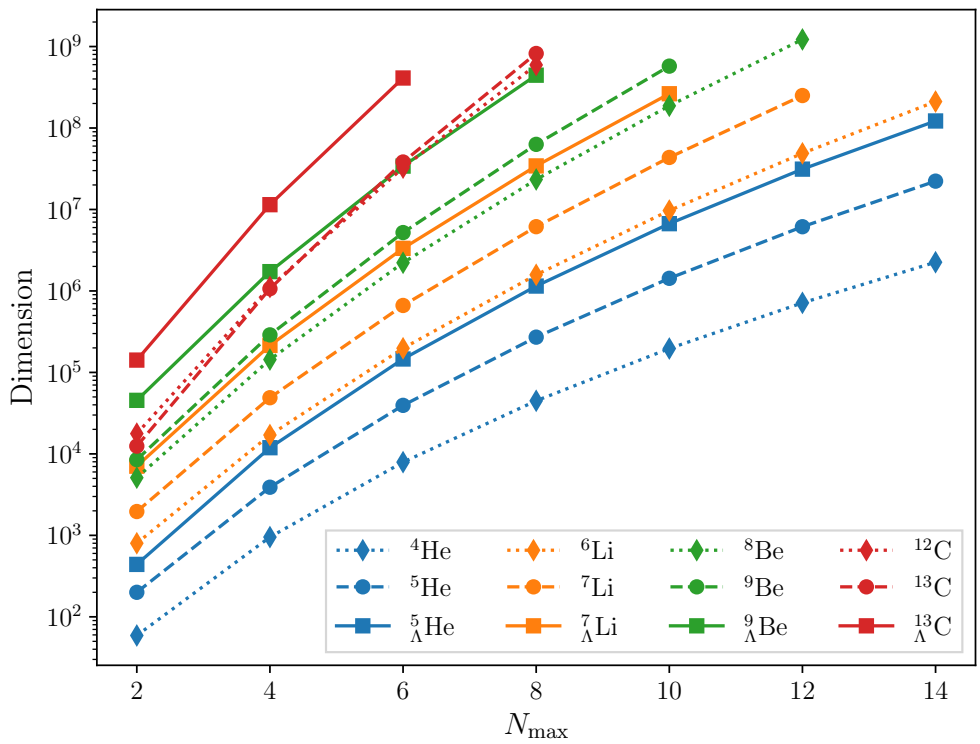

Figure 9.1: Comparison of the growth of NCSM model-space size with $N_{\max}$ between hypernuclei, their nucleonic parents and nuclei with same particle number A.

which now reads

$$\kappa_\nu^m = -\frac{\langle \Phi_\nu | \hat{H} | \Psi_m^{\text{ref}} \rangle}{\Delta E_\nu + \Delta M_\nu} = -\sum_{\mu \in \mathcal{M}_{\text{ref}}} \frac{\langle \Phi_\nu | \hat{H} | \Phi_\mu \rangle \langle \Phi_\mu | \Psi_m^{\text{ref}} \rangle}{\Delta E_\nu + \Delta M_\nu} \quad \text{for} \quad \nu \notin \mathcal{M}_{\text{ref}}, \tag{9.2}$$

where ΔM_ν resembles the difference between the mass associated with $|\Phi_\nu\rangle$ and the reference rest mass M_0. Analogously to the nucleonic case, we define a truncation $\kappa_{\min}$ to limit the size of the IT-NCSM model space and the extrapolation to $\kappa_{\min} \to 0$ is obtained from polynomial fits at different orders.

With these extensions in place, we can study an example calculation of the ground-state and excited state energies of $^7_\Lambda$Li with a LO YN interaction presented in Fig. 9.2. In the left-hand panel we find that the overall convergence behavior is very similar to the nucleonic case. We further find that the spectroscopic data, i.e., the lowest-lying natural parity states are well described considering that these calculations are performed with a LO YN interaction.

For a more detailed discussion of the method we refer to [42].

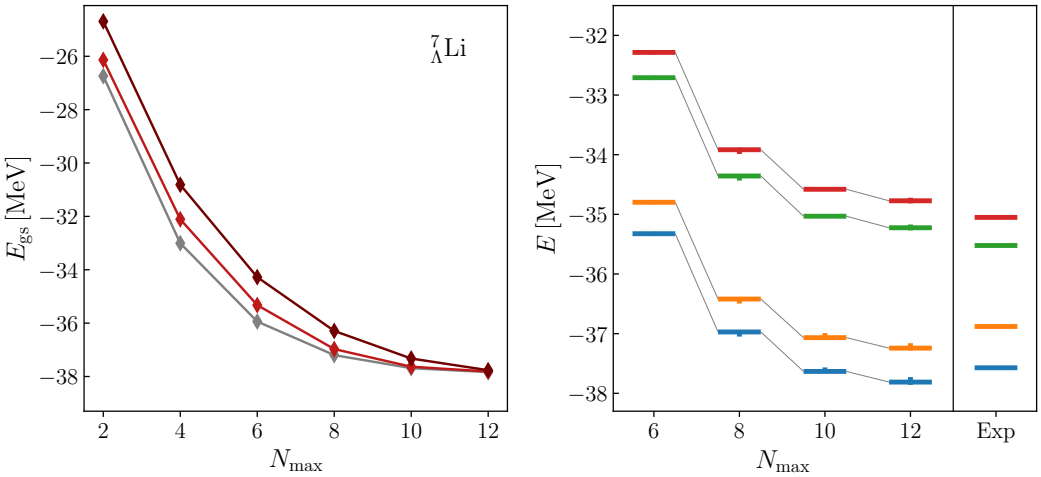

Figure 9.2: IT-NCSM calculations of $^{7}_{\Lambda}$Li with a chiral non-local NN+3N interaction at N^{3}LO with $\Lambda_C = 500$ MeV and $\alpha = 0.08$ fm^4 accompanied by a LO YN interaction with $\Lambda_C = 700$ MeV consistently evolved to $\alpha = 0.08$ fm^4. Shown are ground-state energies for different $\hbar\Omega = 14, 16,$ and 20 MeV (left-hand panel) and bound-state energies of the lowest-lying natural parity states at $\hbar\Omega = 16$ MeV along with experimental values taken from [7, 146] (right-hand panel). Vertical lines indicate IT uncertainties.

9.1 Revisiting Artificial Neural Networks

The main bottleneck of the hypernuclear NCSM remains the rapid growth of model-space dimensions and with that the limited reach w.r.t. $N_{\max}$ and particle number A. As we have shown, this growth is significantly faster than in purely nucleonic systems making the resulting limitation even more severe. In Ch. 5 we have demonstrated that ANNs can provide a robust tool for model-space extrapolations along with meaningful uncertainty estimates. The ANNs have been constructed to learn and adapt to various different convergence patterns resulting in a universality that enables the application across several p-shell nuclei, interactions and states.

It is this universality that suggests the direct transfer of the ANNs to hypernuclear NCSM calculations. From Fig. 9.2 we know that the convergence patterns in hypernuclei are very much alike those we find in nucleonic calculations. Hence, we can assume that, due to their interpolation capabilities, the ANNs, though only trained on purely nucleonic systems, can directly be applied to the convergence patterns found in calculations of hypernucei.

Figure 9.3 shows the predictions from the exact same ANNs employed in Ch. 6 for the ground-state energies of selected hypernuclei. The evaluation data is depicted in the left-hand panels and the corresponding distributions of predictions in the right-hand panels. For both systems we find very robust predictions for the ground-state energies that are in very good agreement across all accessible $\mathcal{N}_{\max}$. This is remarkable considering that these predictions are obtained from three $\hbar\Omega$ only and the ANNs have never been informed with data for hypernuclei.

Another observable of particular interest for hypernuclei is the hyperon separation energy B_Λ, which is defined as the difference of the binding energy of the hypernucleus and the binding energy of its nucleonic parent. As it is very sensitive to the YN interaction, it

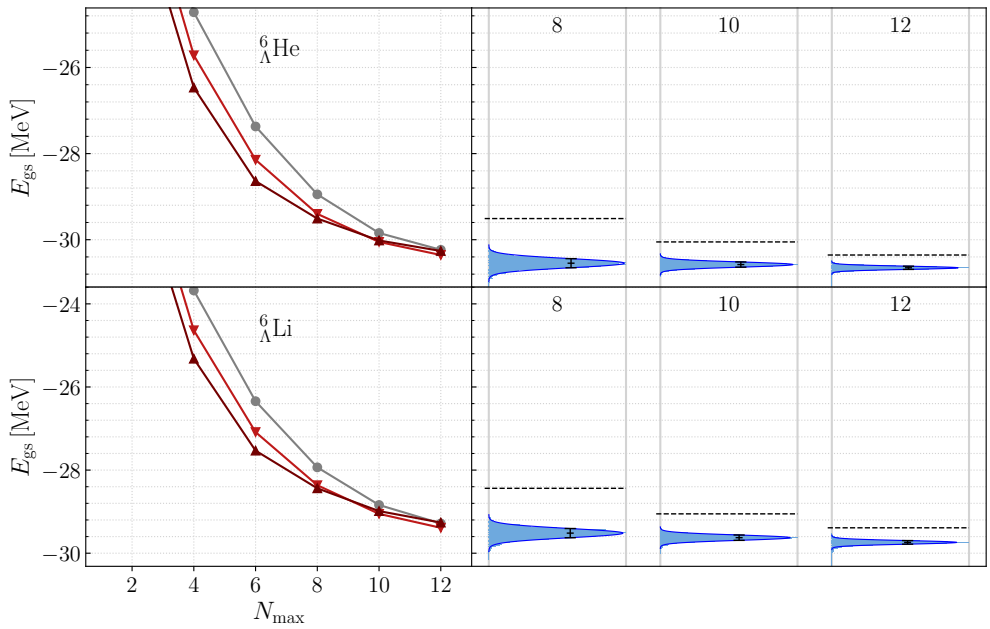

Figure 9.3: Evaluation data (left-hand panels) along with predictions of ground-state energies at different $\mathcal{N}_{\max}$ (right-hand panels) for selected hypernuclei from the same 1000 ANNs in MINMAX mode as employed in Ch. 6. Calculations have been performed for $\hbar\Omega = 14, 16,$ and 20 MeV and horizontal dashed lines indicate the variational boundaries.

provides an ideal tool for studying it. Hence, being able to provide reliable predictions and uncertainty estimates for B_Λ is a crucial aspect for future investigations. The problem is that the constraints on the convergence behavior of bound-state energies do not transfer to B_Λ and training ANNs specifically for the prediction of it is not desirable since we already have ANNs that provide good extrapolations of energies. We can, therefore, make use of it being a difference-based observable and treat it analogously to the excitation energies in Sec. 6.3, i.e., we construct pairs of evaluation samples for a hypernucleus and its nucleonic parent for the same values of $\hbar\Omega$ and $N_{\max}$, which we then use as input for the ANNs and the difference of the ANN outputs yields a single prediction of B_Λ. When we do so for all possible pairs of samples we obtain distributions of predictions from which we can deduce a final prediction along with an uncertainty by statistical means.

An example for the prediction of B_Λ for $^{7}_{\Lambda}$Li is provided in Fig. 9.4. The upper panels depict the evaluation data and predictions for the ground-state energy of $^{7}_{\Lambda}$Li and its nucleonic parent ^{6}Li. By sample-wise subtraction of these distributions we find the predictions for B_Λ as shown in the lower panels. For the unconstrained convergence behavior of B_Λ we find monotonously increasing sequences up to $N_{\max} = 8$ before the larger frequencies exhibit a kink and decrease again. When data from sufficiently large model spaces, here $\mathcal{N}_{\max} \geq 10$, is available the predictions are in good agreement and match the expectations. Smaller model spaces on the other hand result in incompatible predictions. However, when looking at the convergence behavior of B_Λ in the lower left panel and considering the monotonously increasing behavior up to $N_{\max} = 8$ the predictions again match the expectations.

Overall, the ANNs provide a good method for the extrapolation of hyperon separation energies and will, therefore, be valuable for the precision study of hypernuclei in the following chapter.

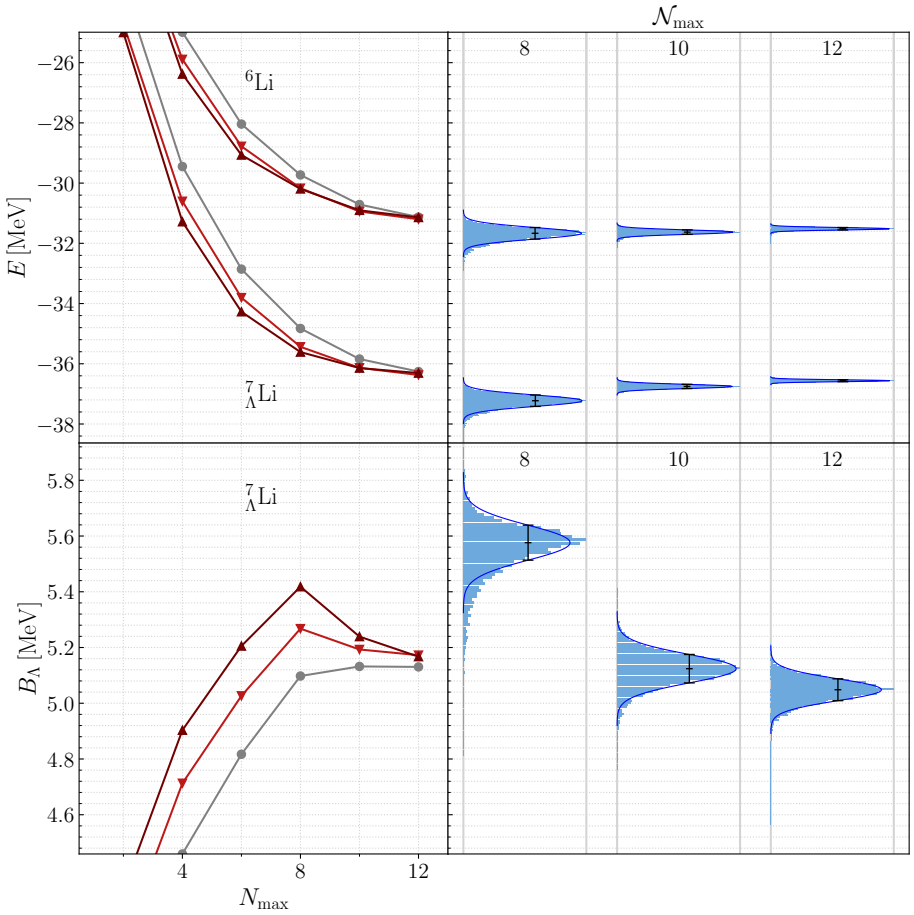

Figure 9.4: Evaluation data along with predictions of ground-state energies (upper panels) and hyperon separation energies (lower panels) of $^7_\Lambda$Li and its parent nucleus ^{6}Li using the same 1000 ANNs in MINMAX mode as employed in Ch. 6. Calculations have been performed for $\hbar\Omega = 14, 16,$ and 20 MeV. See text for details.

10

Hyperon-Nucleon Interaction Constrained on p-Shell Hypernuclei

The construction of a realistic YN interaction is crucial for an accurate description of hypernuclei. As discussed in Ch. 8, the extension of chiral EFT from nucleons to hyperons allows for a systematic construction of such an interaction in analogy to the purely nucleonic sector. By design, these interactions come with a set of LECs that needs to be constrained on experimental data, typically two-baryon scattering data. However, the very limited availability of such scattering data, i.e., a total of 35 YN scattering data in S-wave channels [147–150], presents a major bottleneck for fitting procedures of these LECs [1]. In particular, the YN interactions derived in [19–22] come with 5 LECs at LO and 23 LECs at NLO and N^2LO. While the number of LECs in the latter can be reduced to 10 by imposing additional symmetries, which are known to be broken, the lack of P-wave data remains and one has to resort to NN data in order to fit these LECs. Given the poor quality of the data, even constraining the 5 LECs at LO is challenging and consequently results in a rather ill-constraint interaction, which exhibits deficits in the description of hypernuclei, overbinding the hyperon [42]. Hence, additional constraints on those LECs would provide a significant improvement of YN interaction models.

In recent generations of nucleonic interactions, many-body observables, such as ground-state energies or spectroscopic data of p-shell nuclei, have been employed as additional constraints supplementing the scattering data and yielding a considerably improved description of ground-state energies and radii up to the medium-mass regime [78, 94]. Furthermore, the aforementioned YN interactions already include the hyperon separation energy of the hypertriton in the LEC fitting strategy. Along the same lines we can make use of experimental data for selected p-shell hypernuclei, which is far more precise than the scattering data, aiming to construct a YN interaction that is optimized for hypernuclear structure calculations. In order to keep the number of LECs low, we concentrate our investigations on the LO YN interaction.

The optimization procedure as well as some of the results presented in this chapter have been published in [155].

[1] As of today, a few more YN scattering data have become available [151–154], however, the interactions employed in this work were only fitted to the aforementioned 35 data.

LEC	Λ_C^{YN} [MeV]			
	550	600	650	700
$C^{\Lambda\Lambda}_{^1S_0}$	−0.0466	−0.0403	−0.0322	−0.0304
$C^{\Lambda\Lambda}_{^3S_1}$	−0.0222	−0.0163	−0.0097	−0.0022
$C^{\Sigma\Sigma}_{^1S_0}$	−0.0766	−0.0763	−0.0757	−0.0744
$C^{\Sigma\Sigma}_{^3S_1}$	0.2336	0.2391	0.2392	0.2501
$C^{\Lambda\Sigma}_{^3S_1}$	−0.0016	−0.0019	0.0000	0.0035

Table 10.1: Values of the LECs occurring in the LO chiral YN interaction for different cutoffs Λ_C^{YN} as presented in [19]. The values are obtained through fits to 35 YN scattering data and are given in 10^4 GeV^{-2}.

10.1 The Leading-Order YN Interaction

Starting point for our investigations is the LO YN interaction from chiral EFT as derived by Polinder et al. in [19], which we will denote as $\mathrm{YN_P}$, that is SRG evolved to flow parameter $\alpha = 0.08$ fm^4. As already mentioned, it comes with 5 LECs associated with particle species and partial waves as given in Tab. 10.1. By construction of the interaction, Λ_C^{YN} resembles an additional free parameter which subsequent calculations will be sensitive to. Hence, the LECs are given for different cutoffs $\Lambda_C^{YN} = 550, 600, 650,$ and 700 MeV. The corresponding values of the LECs are then obtained through fits of calculated scattering cross sections to the 35 YN scattering data available. Note that these fits are of similar quality for all cutoffs.

In a first step towards a YN interaction optimized for calculations of many-body systems, we want to understand how calculations for p-shell hypernuclei depend on Λ_C^{YN} and the corresponding LECs. We, therefore, start by studying calculations of ${}^5_\Lambda$He, ${}^7_\Lambda$Li, and ${}^9_\Lambda$Be for different cutoff values $\Lambda_C^{YN} = 600, 700$ MeV as shown in Fig. 10.1. All three hypernuclei are overbound and this overbinding is far more pronounced for $\Lambda_C^{YN} = 600$ MeV (dashed lines) than for $\Lambda_C^{YN} = 700$ MeV (solid lines). Hence, there is a strong cutoff dependence and one

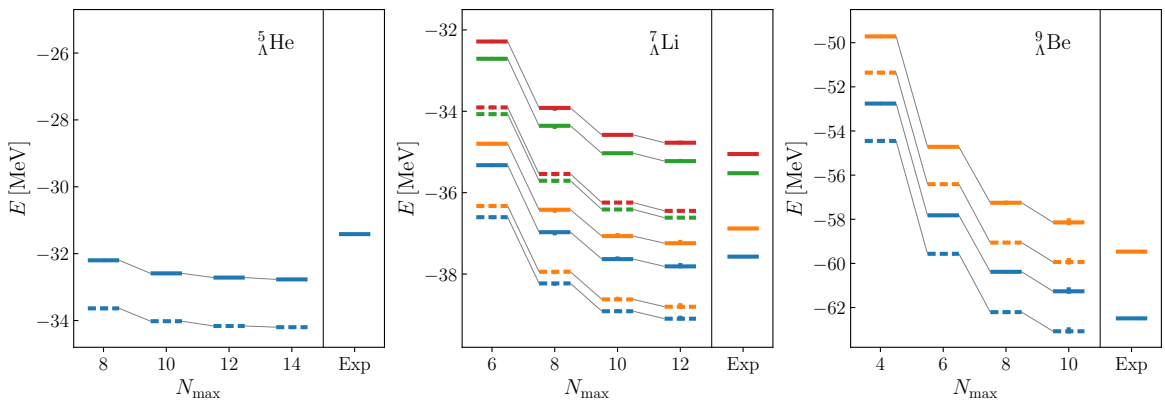

Figure 10.1: IT-NCSM calculations of ${}^5_\Lambda$He, ${}^7_\Lambda$Li, and ${}^9_\Lambda$Be with $\hbar\Omega = 16$ MeV for $\mathrm{YN_P}$ with $\Lambda_C^{YN} = 600$ MeV (dashed lines) and $\Lambda_C^{YN} = 700$ MeV (solid lines). Experimental values taken from [7, 146].

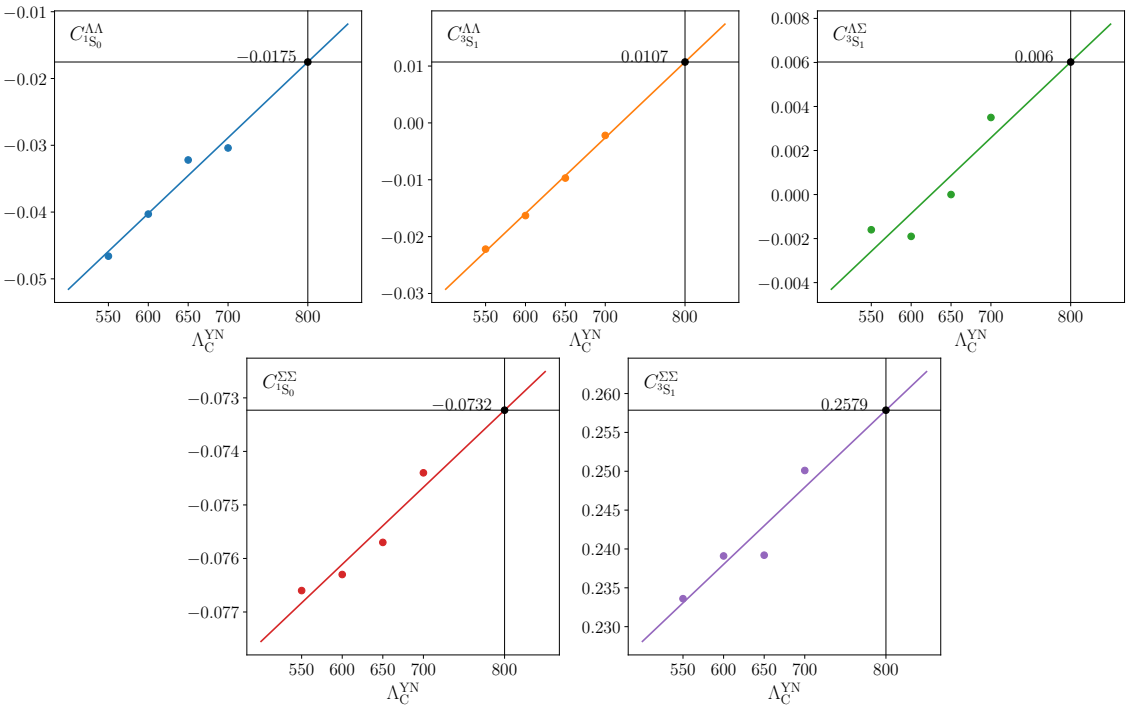

Figure 10.2: Values for the 5 LECs in the LO YN interaction for different cutoffs $\Lambda_C^{YN} = 550, 600, 650,$ and 700 MeV as given in Tab. 10.1 along with linear extrapolations to $\Lambda_C^{YN} = 800$ MeV.

would expect the many-body calculations to further improve when the cutoff is increased. Following these findings in a simple and rather heuristic manner we construct an interaction at $\Lambda_C^{YN} = 800$ MeV and adjust the LECs by a linear extrapolation as shown in Fig. 10.2. Depicted are the values for all 5 LECs at different cutoffs $\Lambda_C^{YN} = 550, 600, 650,$ and 700 MeV. The LECs roughly increase linearly with cutoff, which is indicated by the colored lines that have been fitted to the available data. From this we can extract the extrapolated values at $\Lambda_C^{YN} = 800$ MeV:

$$\begin{aligned} C^{\Lambda\Lambda}_{^1S_0} &= -0.0175, \\ C^{\Lambda\Lambda}_{^3S_1} &= 0.0107, \\ C^{\Lambda\Sigma}_{^3S_1} &= 0.0060, \\ C^{\Sigma\Sigma}_{^1S_0} &= -0.0732, \\ C^{\Sigma\Sigma}_{^3S_1} &= 0.2579. \end{aligned} \tag{10.1}$$

With this new set of LECs we again perform calculations for ${}^7_\Lambda$Li presented in Fig. 10.3 and compare the results for $\Lambda_C^{YN} = 700$ MeV (solid lines) and $\Lambda_C^{YN} = 800$ MeV (dotted lines). Contrary to our expectations, we find that this extrapolated interaction yields a stronger overbinding. Moreover, the hyperon induced splittings of the angular momentum partner states, i.e., ground state and first excited state as well as second and third excited state, become too large, which leads to an overall decreased accuracy.

Clearly, such simple adjustments of the YN interactions are insufficient for an accurate description of p-shell hypernuclei and a more sophisticated optimization procedure is required. We will, therefore, construct an improved interaction based on the adjustment of individual LECs in analogy to the procedure employed in [78].

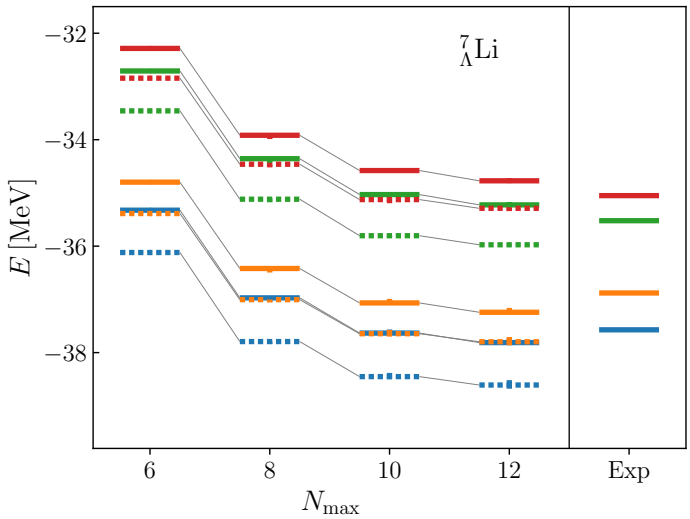

Figure 10.3: IT-NCSM calculations of $^7_\Lambda$Li with $\hbar\Omega = 16$ MeV for $\Lambda_C^{YN} = 700$ MeV (solid lines) and the extrapolated interaction at $\Lambda_C^{YN} = 800$ MeV (dotted lines). Experimental values taken from [7, 146].

10.2 LEC Sensitivity Analysis

Before we attempt an optimization of the LO YN interaction at the level of individual LECs, we first need to investigate how changes of the LECs manifest in structure calculations of p-shell hypernuclei. We conduct this sensitivity analysis on $^7_\Lambda$Li as it is the most precisely measured hypernucleus that additionally features bound excited states. In particular, we take a closer look at the hyperon separation energy B_Λ and the excitation energy of the first excited $\frac{3}{2}^+$ state.

We further need to specify how to vary the LECs on a reasonable scale. In order to identify a variation that we consider to be of natural scale, we again take a look at the change of the LECs for different cutoffs as done in Fig. 10.2. This will prevent unnaturally large or insignificantly small adjustments to the LECs, which is crucial as we aim to maintain a good description of the scattering data to which the LECs have been fitted to initially. Since the previous extrapolation to $\Lambda_C^{YN} = 800$ MeV already resembles a variation that is representative for the average change of the LECs over the given cutoffs we can understand the values in Eq. (10.1) as natural variations. Hence, we study the sensitivity of the aforementioned observables on the LECs by changing them, one by one, to the values in Eq. (10.1) while keeping the others unchanged w.r.t. the YN interaction at $\Lambda_C^{YN} = 700$ MeV, which resembles the starting point for our optimization.

The results of the sensitivity analysis are shown in Fig. 10.4. Note that the results for varied LECs (red) need to be assessed in comparison to the result of the original interaction (green). The experimental results (black) give an orientation for how much adjustment of a given LEC is needed for a subsequent optimization procedure. We find that both, B_Λ and ΔE_1^*, are by far most sensitive to $C^{\Lambda\Lambda}_{^3S_1}$ followed by $C^{\Lambda\Lambda}_{^1S_0}$. This matches the naïve expectation that the $\Lambda\Lambda$ channels are dominant, since the Σ-content of the Λ-Σ mixing to the low-lying states are usually very small. We further find that an increase of $C^{\Lambda\Lambda}_{^3S_1}$ also corrects the result in the direction of the experimental values. The other LECs associated with Σ hyperons have

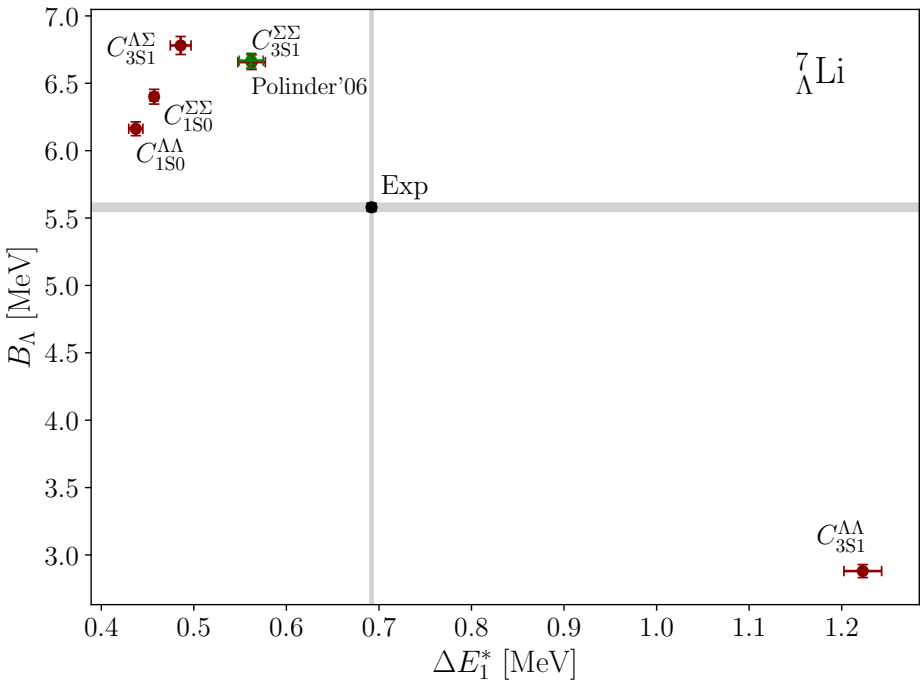

Figure 10.4: IT-NCSM calculations for the hyperon separation energy and fist excitation energy of $^{7}_{\Lambda}$Li at $N_{\max}$ =10 with $\hbar\Omega = 16$ MeV for the original YN interaction with $\Lambda_C^{\mathrm{YN}} = 700$ MeV (green) and with individually adjusted LECs (red). Experimental values (black) from [7, 146] are given for comparison.

much smaller or even negligible effects.

Based on this analysis we choose $C^{\Lambda\Lambda}_{^1\mathrm{S}_0}$ and $C^{\Lambda\Lambda}_{^3\mathrm{S}_1}$ as degrees of freedom for our optimization procedure. Because of the small Σ admixtures these are the dominant interaction channels. Adjusting the Σ-channels could easily lead to an unphysical amplification of Σ-related effects. Moreover, the limitation to only two LECs holds the advantage that stronger constraints can be achieved with less data, which allows us to focus on the most precisely measured hypernuclei.

10.3 Optimization Procedure

For the optimization we first select a set of experimental data for well-known p-shell hypernuclei, which we want to constrain the interaction on. As we are focusing on the YN interaction, we choose observables that are predominantly controlled by the YN interaction, i.e., hyperon separation energies and energy differences between spin-orbit partner states in excitation spectra. In particular we consider B_Λ for $^{3}_{\Lambda}$H, $^{5}_{\Lambda}$He, $^{7}_{\Lambda}$Li, and $^{9}_{\Lambda}$Be along with ΔE for the $\frac{1}{2}^+$ and $\frac{3}{2}^+$ states and the $\frac{5}{2}^+$ and $\frac{7}{2}^+$ states in $^{7}_{\Lambda}$Li as well as the $\frac{3}{2}^+$ and $\frac{5}{2}^+$ states in $^{9}_{\Lambda}$Be. Note that, while these are amongst the best-studied hypernuclei, the experimental values for the above observables scatter significantly and different experiments are not necessarily consistent with one another. Hence, the selection of experimental values will affect the optimization procedure. The selection of values used in this work is listed in Tab. 10.2. A more complete overview of the current experimental situation can be found in [156].

In addition to the data, a metric is required. We choose to perform the optimization w.r.t.

Obs	Exp [MeV]	Obs	Exp [MeV]
$B_\Lambda\left({}^3_\Lambda\mathrm{H}\right)$	0.41(23)	$\Delta E\left({}^7_\Lambda\mathrm{Li}\right)_{\frac{1}{2}^+\to\frac{3}{2}^+}$	0.692(2)
$B_\Lambda\left({}^5_\Lambda\mathrm{He}\right)$	3.12(2)	$\Delta E\left({}^7_\Lambda\mathrm{Li}\right)_{\frac{5}{2}^+\to\frac{7}{2}^+}$	0.471(2)
$B_\Lambda\left({}^7_\Lambda\mathrm{Li}\right)$	5.58(3)	$\Delta E\left({}^9_\Lambda\mathrm{Be}\right)_{\frac{3}{2}^+\to\frac{5}{2}^+}$	0.043(2)
$B_\Lambda\left({}^9_\Lambda\mathrm{Be}\right)$	6.59(15)		

Table 10.2: Experimental values for hyperon separation energies B_Λ and energy differences of spin-orbit partner states ΔE used as constraints for the YN interaction [7, 146].

a χ^2 metric

$$\chi^2 = \sum \frac{(o - o_{\mathrm{exp}})^2}{\sigma^2_{\mathrm{exp}} + \sigma^2_{\mathrm{theo}}} \tag{10.2}$$

in which we include experimental and theoretical uncertainties.

In order to find the optimal values for the two selected LECs, our strategy is to construct a grid in the $C^{\Lambda\Lambda}_{{}^1\mathrm{S}_0}$ - $C^{\Lambda\Lambda}_{{}^3\mathrm{S}_1}$ plane, perform IT-NCSM calculations for the fit observables at the grid points, and interpolate the results. Since the computational cost for calculating a SRG evolved YN interaction along with multiple subsequent IT-NCSM calculations is high, we limit ourselves to a single HO frequency of $\hbar\Omega = 16$ MeV. Note that this prohibits the application of the ANN extrapolation tool discussed in Ch. 5. We will, therefore, use the results at the largest model space accessible, which is $N_{\mathrm{max}} = 14, 12, 8$ for ${}^5_\Lambda\mathrm{He}$, ${}^7_\Lambda\mathrm{Li}$, and ${}^9_\Lambda\mathrm{Be}$, respectively. For the theoretical uncertainties, however, we perform calculations for three HO frequencies $\hbar\Omega = 14, 16,$ and 20 MeV for the original YN interaction. From this we can estimate many-body uncertainties via the ANN tool, which we assume to be the same at all grid points. This allows us to provide realistic many-body uncertainties, while keeping the computational costs at a minimum. Results for ${}^3_\Lambda\mathrm{H}$ are obtained with Jacobi-NCSM calculations at $N_{\mathrm{max}} = 40$ and extrapolated by fitting an exponential function. From the fit we can extract the final prediction along with a fit uncertainty that serves as theoretical error.

Once we have calculated all observables at the grid points, we interpolate them individually, using the LinearNDInterpolator function provided by the scipy Python package [2], before evaluating the metric.

The combined results of this analysis are presented in Fig. 10.5, which shows the χ^2 metric as a contour plot constructed from the interpolation between the grid points marked as black crosses. The original LEC values are indicated by a blue circle in the lower left corner. We find that this metric shows one pronounced minimum within the considered range. It is indicated by the purple star and located at

$$C^{\Lambda\Lambda}_{{}^1\mathrm{S}_0} = -0.0146, \quad C^{\Lambda\Lambda}_{{}^3\mathrm{S}_1} = 0.0004 \tag{10.3}$$

with $\chi^2 = 96$. While this change of 48% in $C^{\Lambda\Lambda}_{{}^1\mathrm{S}_0}$ and 118% in $C^{\Lambda\Lambda}_{{}^3\mathrm{S}_1}$ compared to the original values might seem large, it is well within the previously considered natural range of the

[2]The interpolation has additionally been performed using the interp2d and CloughTocher2DInterpolator functions from the same package in order to check for robustness. All three methods have resulted in identical values at the level of precision we are aiming for.

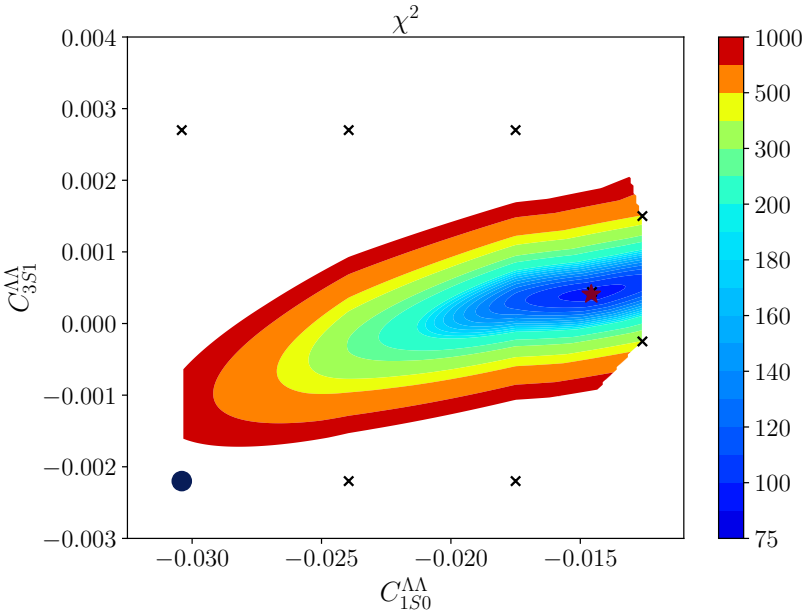

Figure 10.5: Contour plot of the χ^2 metric in the LEC plane spanned by $C^{\Lambda\Lambda}_{^1\mathrm{S}_0}$ and $C^{\Lambda\Lambda}_{^3\mathrm{S}_1}$. The black crosses mark the grid points for the actual calculations, the blue circle indicates the original values of the LECs and the point for the minimal χ^2 is given by the purple star. The colored area is obtained through interpolation between the black crosses.

parameters. Yet, we emphasize that $C^{\Lambda\Lambda}_{^3\mathrm{S}_1}$ changes its sign. The optimization further retains a reasonably good description of the scattering data with only minor deviations from results obtained with the original LECs. This is illustrated in Fig. 10.6 which shows results for the Λp elastic scattering cross sections for the original and the adjusted YN interaction obtained from solving the Lippmann-Schwinger equation given in [19], which have been performed and kindly provided by J. Haidenbauer [157]. We find that the overall shape of the cross sections does not change, but features a minor shift to smaller values. Generally speaking, the optimization results in a slightly weakened ΛN interaction, which is compatible with the systematic overbinding of the hyperon we found with the original YN interaction.

Indeed, looking at hypernuclear structure calculations for the observables included in the fit, which are shown in Fig. 10.7 we find that, except for $^3_\Lambda$H, the hyperon separation energies are in much better agreement with experiment, while the energy differences in the spin-orbit partner states change very little. The hypertriton, however, becomes unbound. While this is unwanted it does not indicate a major deficiency of the interaction. Note that we are only considering a LO YN interaction, for which the appropriate interaction uncertainties are much larger than the difference between the two results we find with the different interactions. Hence, taking all sources of uncertainty into account, the hypertriton can easily be shifted back to a bound system through a NLO or N^{2}LO correction. We further find a tension between the hyperon separation energies for $^5_\Lambda$He and $^7_\Lambda$Li, which are slightly over- and underbound. Apparently, this cannot be resolved by an adjustment of the dominant LECs in the LO interaction. However, these deviations are, again, well within the expected interaction uncertainties. Overall, we can conclude that the optimized YN interaction provides a significantly improved description of the considered hypernuclei.

As already mentioned, the optimization result depends on the chosen experimental val-

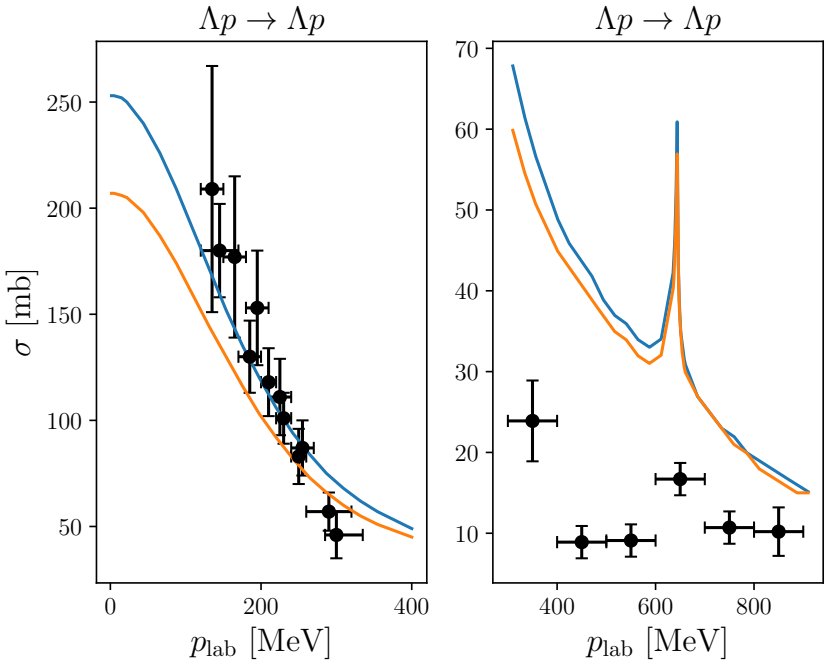

Figure 10.6: Λp elastic scattering cross sections for different momenta $p_{\rm lab}$ obtained from solving a Lippmann-Schwinger equation as given in [19] for the original YN interaction (—) and the optimized interaction (—). Black markers denote experimental data taken from [147, 148, 158, 159].

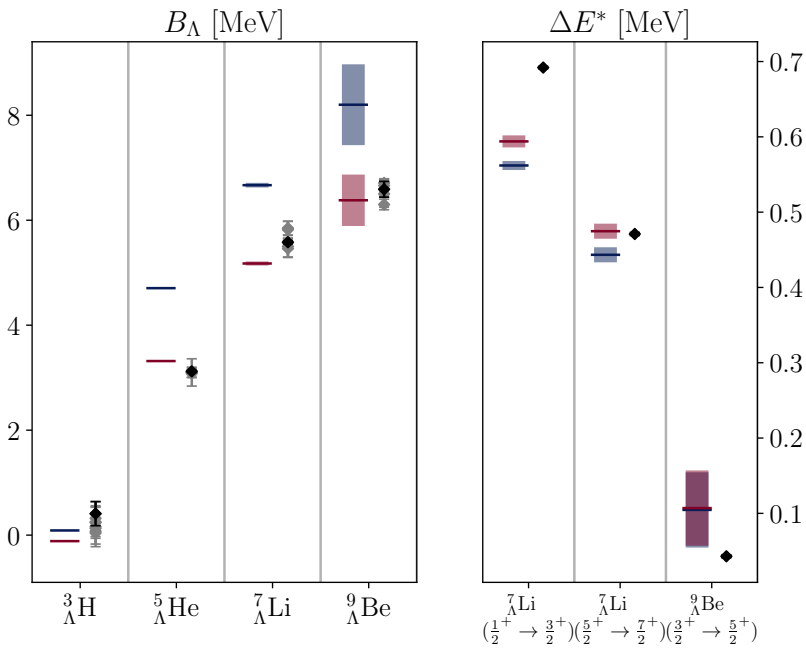

Figure 10.7: IT-NCSM calculations for the observables used as constraints for the LEC optimization along with many-body uncertainties obtained with the ANN tool based on calculations for $\hbar\Omega = 14, 16,$ and 20 MeV. Results in blue are obtained with the original YN interaction while results in purple indicate results for the optimized YN interaction. Black and gray markers indicate experimental results from [7, 146, 156], where the black markers resemble the values chosen for the optimization according to Tab. 10.2.

ues. It turns out that the chosen value for the hyperon separation energy of ${}^{3}_{\Lambda}\mathrm{H}$ is disputed amongst experimentalist. We, therefore, investigate the dependence on this particular datum by repeating the optimization with a value of $B_\Lambda\left({}^{3}_{\Lambda}\mathrm{H}\right) = 0.148(40)$ MeV which is currently recommended by [156]. This results in a minimum at

$$C^{\Lambda\Lambda}_{{}^1\mathrm{S}_0} = -0.0153, \quad C^{\Lambda\Lambda}_{{}^3\mathrm{S}_1} = 0.0004 \tag{10.4}$$

with $\chi^2 = 132.5$. The deviation of $C^{\Lambda\Lambda}_{{}^1\mathrm{S}_0}$ from the previously obtained result is very small compared to the change of the LEC w.r.t. the original YN interaction. Hence, any effects on the many-body calculations are expected to remain within our uncertainties.

We will, therefore, stick with the LECs as given in Eq. (10.4) and refer to the optimized YN interaction as $\mathrm{YN}_{\mathrm{opt}}$ from now on.

10.4 Results for p-Shell Hypernuclei

In order to investigate the quality of the optimized interaction we conduct a study across various p-shell hypernuclei beyond the ones that the interaction has been constrained with. For that we first study a selection of well-known hypernuclei up into the mid-p-shell, followed by closer investigations of the ${}_{\Lambda}\mathrm{He}$ and ${}_{\Lambda}\mathrm{Li}$ isotopic chains up to neutron-rich isotopes. The results for the first two have been published in Ref. [155] and are supplemented by new ${}_{\Lambda}\mathrm{Li}$ calculations. We provide an overview over calculations of excitation spectra for the hypernuclei and their nucleonic parents along with hyperon separation energies for the original YN interaction by Polinder et al. [19] with $\Lambda_{\mathrm{C}}^{\mathrm{YN}} = 700$ MeV ($\mathrm{YN_P}$) and $\mathrm{YN}_{\mathrm{opt}}$. Both are accompanied by the non-local chiral NN+3N interactions from Hüther et al. [78] at N^{3}LO with $\Lambda_{\mathrm{C}} = 500$ MeV ($\mathrm{NN_{EMN}} + \mathrm{3N_H}$). We further employ the ANN extrapolation to obtain predictions for the converged values with many-body uncertainties. Figures illustrating the quality of the ANN extrapolation are provided in Appendix B.1. In addition we compare our calculations to previous calculations by Wirth et al. [42, 43, 145], which are obtained with a non-local NN interaction at N^{3}LO by Entem and Machleidt [160] together with a local 3N interaction at N^{2}LO by Navrátil [161] both with $\Lambda_{\mathrm{C}} = 500$ MeV ($\mathrm{NN_{EM}} + \mathrm{3N_N}$) and the $\mathrm{YN_P}$ interaction in the hyperonic sector. The error bands for these results resemble importance truncation uncertainties. All interactions are consistently SRG evolved to $\alpha = 0.08$ fm^4 and induced forces are included up to the three-body level.

The results are presented in Figs. 10.8 to 10.10, which are structured as follows: The center and upper panels show the ground-state energies and lowest-lying excited states of the hypernuclei and the nucleonic parents, where different (hyper-)nuclei are separated by vertical gray lines. In the lower panel the resulting hyperon-separation energies are given. For each (hyper-)nucleus the different columns correspond to different interactions, i.e., $\mathrm{NN_{EM}} + \mathrm{3N_N} + \mathrm{YN_P}$ (left-hand column), $\mathrm{NN_{EMN}} + \mathrm{3N_H} + \mathrm{YN_P}$ (center column), and $\mathrm{NN_{EMN}} + \mathrm{3N_H} + \mathrm{YN}_{\mathrm{opt}}$ (right-hand column) and are supplemented with experimental values indicated by diamonds.

Again, different experimental results exhibit a significant spread beyond their associated uncertainties, particularly for hyperon separation energies. This disagreement makes a direct comparison to experiment rather challenging. Hence, we depict all reasonable experimental

results in gray in addition to a selected value given in black [3].

We start our discussion by assessing the results for ${}^{5}_{\Lambda}\mathrm{He}$, ${}^{7}_{\Lambda}\mathrm{Li}$, ${}^{9}_{\Lambda}\mathrm{Be}$, and ${}^{13}_{\Lambda}\mathrm{C}$ shown in Fig. 10.8. Here, we expect an improvement of the optimized interaction compared to $\mathrm{YN_P}$ as all of those hypernuclei except for ${}^{13}_{\Lambda}\mathrm{C}$ are part of the data set that $\mathrm{YN_{opt}}$ has been optimized on. First of all, we find that the spectra for both, the hypernuclei and their nucleonic parents are in overall good agreement with experiment and, moreover, discrepancies in the hypernuclear spectra correspond to analogous discrepancies in the nuclear spectra and can, therefore, be attributed to the nucleonic interaction. This holds for all interactions and the dependence on the interaction itself is rather small in general. This changes drastically for the hyperon separation energies, which validates our initial assumption that this observable is far more sensitive to changes in the YN sector. Studying the center and right-hand columns that only differ by the YN interaction, we find an overall reduction of the hyperon separation energy leading to significantly better agreement with experiment except for ${}^{13}_{\Lambda}\mathrm{C}$, which also exhibits the widest spread of experimental values. We further find, that the dependence on the nucleonic interaction, which has also been found in other applications [95, 166], increases drastically with increasing particle number, which can be seen from the comparison of the left-hand and center columns that share the same YN interaction. As a consequence, ${}^{13}_{\Lambda}\mathrm{C}$ does not provide additional insight on the quality of $\mathrm{YN_{opt}}$ compared to $\mathrm{YN_P}$ as the choice of the nucleonic interaction might easily favor one over the other.

We, therefore, turn our investigations to the ${}_{\Lambda}\mathrm{He}$ isotopic chain depicted in Fig. 10.9. An assessment of the hypernuclear spectra w.r.t. experiment is not possible here since there is no experimental data available beyond ${}^{4}_{\Lambda}\mathrm{He}$, for which the agreement is very good. We further find a stronger interaction dependence of the spectra compared to the previous set of hypernuclei, though they are more sensitive to the nucleonic interaction than the hyperonic one, as the center and right-hand columns agree within their uncertainties. Additionally, the changes in the spectra of hypernuclei for the different nucleonic interactions do not always correspond to the shifts of the excited states in the parent nuclei. For the latter, the results up to ${}^{6}\mathrm{He}$ are in good agreement with experiment, while they deviate for heavier isotopes, which hints at deficiencies of the employed nucleonic interactions. Note that the odd isotopes above ${}^{4}\mathrm{He}$ are particle unstable and for the more neutron-rich isotopes a description in NCSM calculations becomes increasingly difficult. This difficulty might translate to the hyperonic sector. The same reasoning applies to ${}^{9}_{\Lambda}\mathrm{He}$ and ${}^{10}_{\Lambda}\mathrm{He}$, which are themselves particle unstable and neutron rich. For the hyperon separation energies, we find an overall decrease for $\mathrm{YN_{opt}}$. This results in a remarkably accurate description of the light isotopes, while the heavier isotopes tend to be underbound. However, we also find the dependence on the nucleonic interaction as well as the spread of experimental results to increase again with particle number.

Let us finally take a closer look at the ${}_{\Lambda}\mathrm{Li}$ isotopic chain shown in Fig. 10.10. Similar to the ${}_{\Lambda}\mathrm{He}$ chain there is very little experimental data for the spectra available. However, shifts in the spectra can again be attributed to the nucleonic interaction as they correspond to the shifts in the nucleonic spectra. Other than that there is very little variation with interaction especially for the lowest excited states. When studying the hyperon separation energies the picture is consistent with the previous figures, yielding a significantly improved description

[3] We emphasize that this illustration might be misleading as the lack of grey markers does not necessarily indicate very consistent and accurate experimental results but can also stem from the lack of more than the one experimental value shown.

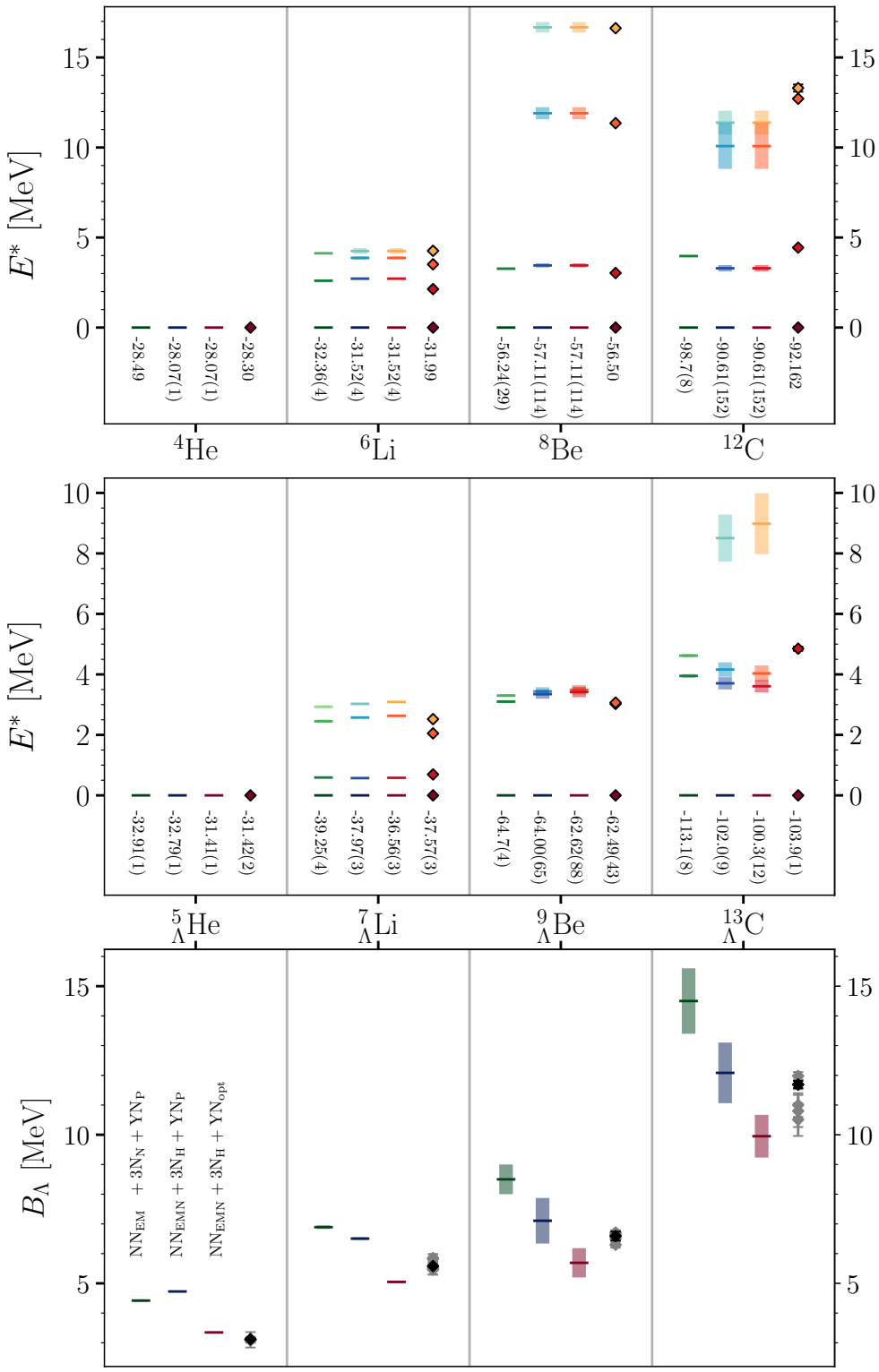

Figure 10.8: ANN extrapolated excitation spectra and hyperon separation energies of the low-lying natural-parity states for selected well-known hypernuclei (center and lower panel) and their nucleonic parents (upper panel) along with many-body uncertainties based on IT-NCSM calculations for $\hbar\Omega = 14, 16,$ and 20 MeV. The different columns for each nucleus indicate different combinations of NN+3N and YN interactions (see text for details) while the right-hand column shows experimental results, where available. Experimental results given in black are selected from [7, 146, 162–165] while the gray markers indicate alternative experimental results taken from [156].

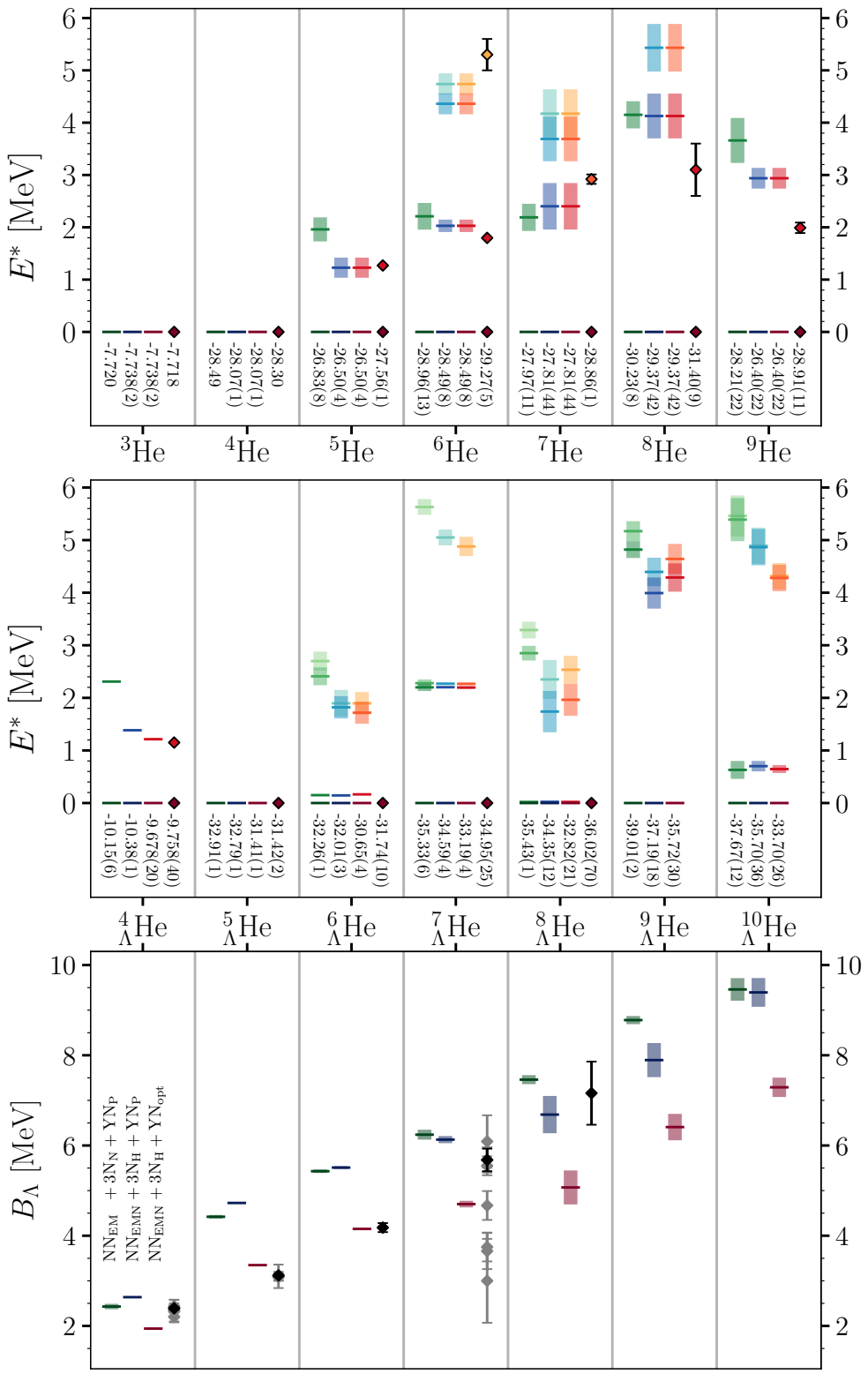

Figure 10.9: Same as Fig. 10.8 but for the $_\Lambda$He isotopic chain. Experimental results given in black are selected from [7, 11, 146, 162–164].

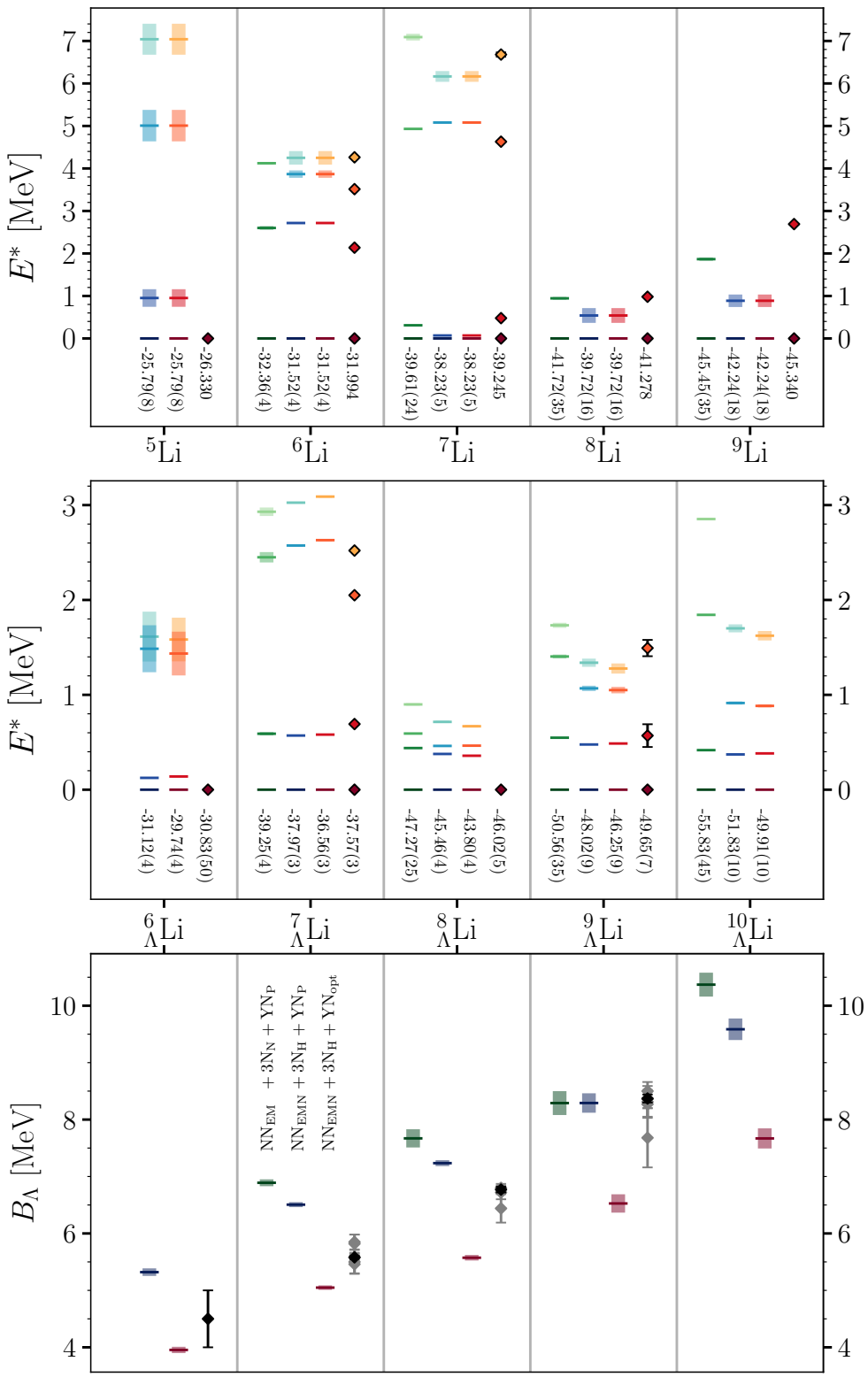

Figure 10.10: Same as Fig. 10.8 but for the $_\Lambda$Li isotopic chain. Experimental results given in black are selected from [7, 146, 162–164, 167].

of the lighter isotopes for the optimized YN interaction, while more neutron-rich isotopes are being underbound. But, contrarily to the previous figures, the dependence on the nucleonic interaction does not increase with particle number.

Overall, we can conclude that the optimization of the LO YN interaction provides more accurate results in particular for light hypernuclei. This is remarkable considering that this was achieved with the adjustment of only two LECs in a LO interaction. Hence, we have demonstrated that nuclear structure data for p-shell hypernuclei provides very valuable additional constraints on hypernuclear interactions that are otherwise poorly constrained from the very limited scattering data available. However, we have also seen that a simultaneous description of light p-shell hypernuclei and more neutron-rich isotopes appears to be difficult. Recalling the tension between the ${}^{5}_{\Lambda}\mathrm{He}$ and ${}^{7}_{\Lambda}\mathrm{Li}$ results for the hyperon separation energies we find that a LO YN interaction is insufficient to capture all relevant physics and higher chiral orders are required to alleviate this.

Considering the difficult experimental situation, calculations with $\mathrm{YN}_{\mathrm{opt}}$ along with reasonable many-body uncertainties obtained via the ANN extrapolation tool are a great asset to hypernuclear structure theory and can provide important insights into baryonic interactions beyond the nucleonic sector. Moreover, they enable predictions of observables which are essential for the design of future experiments. Nevertheless, the significant dependence on the nucleonic interaction as well as the lack of interaction uncertainties need to be considered and should be main subjects of future investigations and developments. Here, the optimized interaction or future optimization procedures could tie in with other efforts towards a meaningful uncertainty estimation based on chiral order-by-order calculations [166] or Bayesian parameter estimation [95].

A complete collection of the numerical values for the results shown in this chapter can be found in Appendix B.2.

11

Natural Orbitals for Hypernuclei

In the previous chapters we have demonstrated that the IT-NCSM is a powerful method for ab initio structure calculations for, both, nuclei and hypernuclei. However, even when combined with sophisticated extrapolation methods like the ANN extrapolation tool, its reach is severely limited by the combinatorial growth of the many-body basis that spans the model space. This is partially due to the rather slow convergence of the HO single-particle basis w.r.t. $N_{\max}$. As discussed in Ch. 2 this choice of basis is reasonable based on the analytical properties of the HO basis, but it is certainly not the ideal single-particle basis in terms of convergence. In recent years, the search for a better-suited single-particle basis has led to so-called natural orbitals (NAT) which provide a computationally efficient yet powerful basis that has successfully been employed in various many-body methods [168, 169]. It is constructed from a one-body density matrix obtained with Hartree-Fock many-body perturbation theory (HF-MBPT) [170–172] and its most prominent features are the rapid convergence and the independence of the underlying HO frequency.

In the following we construct natural orbitals for hypernuclei in order to transfer the features of this improved basis to the hypernuclear regime.

11.1 Conceptual Basics

Before we can derive the hypernuclear natural orbitals we first need to introduce some mathematical basics and notation.

Second Quantization. Second quantization is a useful formalism when working in an anti-symmetric Fock space. At its core are creation and annihilation operators that are constructed to change the particle number of the system and are defined by their action on a Slater determinant. According to their names the creation operator $\hat{a}_a^\dagger$ creates a particle in state $|a\rangle$, while its annihilation counterpart $\hat{a}_a$ removes such a particle from a given state. Put in mathematical terms their action is described by

$$\hat{a}_a^\dagger |a_1, \ldots, a_A\rangle_a = \begin{cases} |a, a_1, \ldots, a_A\rangle_a & \text{if } a \notin \{a_1, \ldots, a_A\} \\ 0 & \text{else} \end{cases}, \tag{11.1}$$

$$\hat{a}_a|a_1,\dots,a_A\rangle_a = \begin{cases}(-1)^{k-1}|a_1,\dots,a_{k-1},a_{k+1},\dots,a_A\rangle_a & \text{if } a=a_k\\ 0 & \text{if } a\notin\{a_1,\dots,a_A\}\end{cases}. \tag{11.2}$$

In addition they fulfill the anti-commutator relations

$$[\hat{a}_a^\dagger,\hat{a}_b^\dagger]_+ = [\hat{a}_a,\hat{a}_b]_+ = 0, \tag{11.3}$$

$$[\hat{a}_a^\dagger,\hat{a}_b]_+ = \delta_{ab}. \tag{11.4}$$

Furthermore, any A-body operator $\hat{O}^{[A]}$ can be expressed in second quantization via

$$\hat{O}^{[A]} = \frac{1}{(A!)^2}\sum_{\substack{a_1\dots a_A\\ a'_1\dots a'_A}} O^{[A]}_{a_1\dots a_A a'_1\dots a'_A}\,\hat{a}^\dagger_{a_1}\dots\hat{a}^\dagger_{a_A}\hat{a}_{a'_A}\dots\hat{a}_{a'_1}, \tag{11.5}$$

where $O^{[A]}_{a_1\dots a_A a'_1\dots a'_A} \equiv {}_a\langle a_1\dots a_A|\hat{O}|a'_1\dots a'_A\rangle_a$ denotes an A-body matrix element of $\hat{O}$.

Particle-Hole Formalism. A Slater determinant can be constructed by applying a sequence of creation operators to the vacuum state $|0\rangle$

$$|\Phi\rangle = |a_1\dots a_A\rangle_a = \hat{a}^\dagger_{a_1}\dots\hat{a}^\dagger_{a_A}|0\rangle, \tag{11.6}$$

which is convenient when constructing, e.g., a NCSM model space. However, several other many-body methods are formulated w.r.t. a reference Slater determinant $|\Phi_{\text{ref}}\rangle$ instead. In this case a model space can be constructed by applying one or multiple pairs of creation and annihilation operators to the reference state

$$\begin{aligned}|\Phi^p_q\rangle &= \hat{a}^p_q|\Phi_{\text{ref}}\rangle \equiv \hat{a}^\dagger_q\hat{a}_p|\Phi_{\text{ref}}\rangle,\\ |\Phi^{pq}_{rs}\rangle &= \hat{a}^{pq}_{rs}|\Phi_{\text{ref}}\rangle \equiv \hat{a}^\dagger_r\hat{a}^\dagger_s\hat{a}_q\hat{a}_p|\Phi_{\text{ref}}\rangle,\\ &\vdots\end{aligned} \tag{11.7}$$

which again yields Slater determinants that are characterized by lower indices (upper indices) referring to the creation (annihilation) operators. In the particle-hole picture the single-particle states that are occupied in $|\Phi_{\text{ref}}\rangle$ are called hole-states and the unoccupied ones are called particle-states respectively. For a better distinction in the subsequent derivations we will use the following index convention

$$\begin{aligned}\text{Holes}: &\quad (i,j,k,\dots)\\ \text{Particles}: &\quad (a,b,c,\dots)\\ \text{General}: &\quad (p,q,r,\dots)\end{aligned}$$

whenever we switch to particle-hole notation.

11.2 The Hartree-Fock Method

The state which the one-body density for the natural orbitals is constructed from is based on the Hartree-Fock (HF) solution for the (hyper)nucleus of interest. The HF method is a variational mean-field method build on the equivalence of the solution of the Schrödinger equation to the stationary point of the energy functional

$$\delta E[|\Psi\rangle] = 0, \tag{11.8}$$

where the latter is given by

$$E[|\Psi\rangle] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{11.9}$$

We approximate this solution by requiring our trial state $|\Psi\rangle$ to be equal to a single Slater determinant $|\Phi_{\rm HF}\rangle$. According to the Ritz variational principle this approximation yields an upper bound to the exact ground-state energy of the system.

In a next step, the HF state can be expressed in a HF single-particle basis $\{b_i\}$ and using second quantization we can write

$$|\Phi_{\rm HF}\rangle = |b_1, b_2, \ldots, b_A\rangle = \prod_{p=1}^{A} \hat{b}_p^\dagger |0\rangle \tag{11.10}$$

with $\{\hat{b}_p^\dagger\}$ being the creation operators corresponding to the HF basis. These can be related to a known reference basis $\{|a_p\rangle\}$, typically the HO basis, with corresponding creation operators $\{\hat{a}_p^\dagger\}$ via

$$\hat{b}_p^\dagger = \sum_q C_{pq} \hat{a}_q^\dagger, \tag{11.11}$$

where C_{pq} denotes the matrix elements of the unitary transformation that connects both bases. We can further express the energy functional in second quantization w.r.t. the same reference basis, which yields

$$\begin{aligned} E[|\Phi_{\rm HF}\rangle] = H^{[0]} &+ \sum_{pq} H^{[1]}_{p,q} \langle\Phi_{\rm HF}|\hat{a}_p^\dagger \hat{a}_q|\Phi_{\rm HF}\rangle + \frac{1}{4}\sum_{pqrs} H^{[2]}_{pq,rs} \langle\Phi_{\rm HF}|\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r|\Phi_{\rm HF}\rangle \\ &+ \frac{1}{36}\sum_{pqrstu} H^{[3]}_{pqr,stu} \langle\Phi_{\rm HF}|\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r^\dagger \hat{a}_u \hat{a}_s \hat{a}_t|\Phi_{\rm HF}\rangle, \end{aligned} \tag{11.12}$$

where we have assumed a Hamiltonian with up to three-body forces. The remaining expectation values are one-body, two-body, and three-body density matrices $\rho^{[i]}$. Since $|\Phi_{\rm HF}\rangle$ is a single Slater determinant the higher-body densities can be decomposed into one-body density matrices

$$\rho^{[2]}_{pq,rs} = \rho^{[1]}_{p,r}\rho^{[1]}_{q,s} - \rho^{[1]}_{p,s}\rho^{[1]}_{qr,} \tag{11.13}$$

$$\rho^{[3]}_{pqr,stu} = \rho^{[1]}_{p,u}\rho^{[1]}_{q,t}\rho^{[1]}_{r,s} \pm \text{permutations} \tag{11.14}$$

and we can, therefore, express the energy functional in terms of one-body density matrices

$$E[\rho^{[1]}] = H^{[0]} + \sum_{pq} H^{[1]}_{p,q}\rho^{[1]}_{p,q} + \frac{1}{2}\sum_{pqrs} H^{[2]}_{pq,rs}\rho^{[1]}_{pr}\rho^{[1]}_{qs} + \frac{1}{6}\sum_{pqrstu} H^{[3]}_{pqr,stu}\rho^{[1]}_{p,s}\rho^{[1]}_{q,t}\rho^{[1]}_{r,u}, \tag{11.15}$$

where we have exploited the antisymmetry of the Hamilton matrix elements. Performing the variation of this energy functional and neglecting quadratic and higher orders in $\delta\rho^{[1]}$ we find

$$\delta E[\rho^{[1]}] = \sum_{pq}\left(H^{[1]}_{p,q} + \sum_{rs} H^{[2]}_{pr,qs}\rho^{[1]}_{rs} + \frac{1}{2}\sum_{rstu} H^{[3]}_{psr,qtu}\rho^{[1]}_{s,t}\rho^{[1]}_{r,u}\right)\delta\rho^{[1]}_{p,q} = 0 \tag{11.16}$$

for the stationary condition. By introducing an auxiliary mean-field Hamiltonian

$$h_{p,q}[\rho^{[1]}] = H^{[1]}_{p,q} + \sum_{rs} H^{[2]}_{pr,qs}\rho^{[1]}_{r,s} + \frac{1}{2}\sum_{rstu} H^{[3]}_{psr,qtu}\rho^{[1]}_{s,t}\rho^{[1]}_{r,u} \tag{11.17}$$

the previous expression can be simplified to

$$\sum_{pq} h_{p,q}[\rho^{[1]}]\delta\rho^{[1]}_{p,q} = 0. \tag{11.18}$$

Since this must hold for an arbitrary variation $\delta\rho^{[1]}$ that is consistent with the constraints from idempotence and hermiticity, we can immediately conclude that

$$h_{p,q}[\rho^{[1]}] = 0, \tag{11.19}$$

which is equivalent to the statement that the commutator of the corresponding operators $\hat{h}[\rho^{[1]}]$, the so-called Fock operator, and $\hat{\rho}$ vanishes. Hence, we can find a common eigenbasis for which we can construct the eigenvalue problem

$$\hat{h}[\rho^{[1]}]|b_p\rangle = \epsilon_p|b_p\rangle. \tag{11.20}$$

Here, the ϵ_p emerge as eigenvalues of the single-particle basis. They are interpreted as the HF single-particles energies to the corresponding HF basis states $|b_p\rangle$. The eigenvalue problem can be solved by transforming it to the reference basis

$$\sum_{r} h_{q,r}[\rho^{[1]}]C_{pq} = \epsilon_p C_{pq} \tag{11.21}$$

yielding the so-called HF equation, a non-linear equation that can be solved iteratively. Note that this equation separates into blocks with distinct particles species. Thus, the extension to hypernuclei does not require additional adjustments. The HF ground state $|\Phi_{\mathrm{HF}}\rangle$ is then constructed from the eigenstates that correspond to the lowest energy eigenvalues w.r.t. the particle content of the system under consideration. For the corresponding HF ground-state energy we find

$$E_{\mathrm{HF}} \equiv E[|\Phi_{\mathrm{HF}}\rangle] = \langle\Phi_{\mathrm{HF}}|\hat{H}|\Phi_{\mathrm{HF}}\rangle = \sum_{p}\epsilon_p - \frac{1}{2}\sum_{pq} H^{[2]}_{pq,pq}, \tag{11.22}$$

which is clearly distinct from a simple sum over the single-particle energies.

Equal-Filling Approximation. The HF method we have discussed so far is only applicable to closed shell nuclei since the HF ground-state is given by a single Slater determinant. But, in the case of singly-strange hypernuclei we are inevitably dealing with open-shell systems. Hence, we need to employ an extension of the HF method called equal-filling approximation [173]. It assumes fractional occupation numbers in the open shells such that all single-particle states within an open shell are fractionally filled. Here, we follow the derivation in [42].

Conceptually, we make a general ansatz for the density operator $\hat{\rho}$ such that the energy functional in terms of the calculation basis becomes

$$E[\hat{\rho}] = H^{[0]} + \sum_{pq} H^{[1]}_{p,q} \mathrm{tr}(\hat{\rho}\hat{a}^\dagger_p \hat{a}_q) + \frac{1}{4}\sum_{pqrs} H^{[2]}_{pq,rs} \mathrm{tr}(\hat{\rho}\hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_s \hat{a}_r) + \frac{1}{36}\sum_{pqrstu} H^{[3]}_{pqr,stu} \mathrm{tr}(\hat{\rho}\hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}^\dagger_r \hat{a}_u \hat{a}_t \hat{a}_s). \tag{11.23}$$

In order to relate this to an energy functional of the same form as in Eq. (11.15) we need to constrain the density operator and require that all higher-body density matrices factorize into one-body density matrices, which requires some occupation numbers to become fractional.

When constructing the general ansatz for the density operator we distinguish between a core $|\mathrm{core}\rangle$ of fully-occupied orbitals with occupation number $n = 1$ for all included states, unoccupied orbitals with $n = 0$, and partially-occupied valence orbitals v with total orbital angular momentum j_v that contain $2j_v + 1$ states with non-negative fractional occupation numbers $n^{(v)}_k$. Hence, $\hat{\rho}$ exhibits an equivalent partitioning and is given by

$$\hat{\rho} = p_{\mathrm{core}}|\mathrm{core}\rangle\langle\mathrm{core}| + \sum_{n=1}^{2j_v+1} \sum_{v_1<\cdots<v_n} p_{\{v_1\cdots v_n\}} \hat{b}^\dagger_{v_1} \cdots \hat{b}^\dagger_{v_n} |\mathrm{core}\rangle\langle\mathrm{core}| \hat{b}_{v_n} \cdots \hat{b}_{v_1} \equiv \sum_\mu p_\mu |\Phi_\mu\rangle\langle\Phi_\mu|, \tag{11.24}$$

where $p_{\{v_1\cdots v_n\}} \geq 0$ are symmetric under the exchange of any two indices. In analogy to Eq. (11.13) and Eq. (11.14) the density matrices can be re-expressed in terms of occupation number operators

$$\hat{n}_p = \hat{b}^\dagger_p \hat{b}_p \quad \text{with} \quad \hat{n}_p|\Phi_\mu\rangle = n^{(\mu)}_p|\Phi_\mu\rangle \tag{11.25}$$

such that

$$\gamma^{[1]}_{p,q} = \sum_\mu p_\mu \langle\Phi_\mu|\hat{b}^\dagger_p \hat{b}_q|\Phi_\mu\rangle = \mathrm{tr}(\hat{\rho}\hat{n}_p)\delta_{pq}, \tag{11.26}$$

$$\gamma^{[2]}_{pq,rs} = \sum_\mu p_\mu \langle\Phi_\mu|\hat{b}^\dagger_p \hat{b}^\dagger_q \hat{b}_s \hat{b}_r|\Phi_\mu\rangle = \mathrm{tr}(\hat{\rho}\hat{n}_p\hat{n}_q)(\delta_{pr}\delta_{qs} - \delta_{ps}\delta_{qr}), \tag{11.27}$$

$$\gamma^{[3]}_{pqr,stu} = \sum_\mu p_\mu \langle\Phi_\mu|\hat{b}^\dagger_p \hat{b}^\dagger_q \hat{b}^\dagger_r \hat{b}_u \hat{b}_t \hat{b}_s|\Phi_\mu\rangle = \mathrm{tr}(\hat{\rho}\hat{n}_p\hat{n}_q\hat{n}_r)(\delta_{ps}\delta_{qt}\delta_{ru} \pm \text{permutations}), \tag{11.28}$$

where we have named them $\gamma^{[i]}$ to emphasize the formulation w.r.t. the HF single particle basis. Density matrices containing occupied or unoccupied states factorize due to the eigenvalue relation of the corresponding occupation number operator, while those with all indices in

valence orbitals remain nontrivial. The requirement that all higher-order density matrices need to factorize yields a set of constraints on the diagonal matrix elements

$$\mathrm{tr}(\hat{\rho}) = p_{\mathrm{core}} + \sum_{n=1}^{2j_v+1} \sum_{v_1<\cdots<v_n} p_{\{v_1\cdots v_n\}} \stackrel{!}{=} 1, \tag{11.29}$$

$$\gamma^{[1]}_{q,q} = p_q + \sum_{n=2}^{2j_v+1} \sum_{v_2<\cdots<v_n} p_{\{v_2\cdots v_n\}} \stackrel{!}{=} \frac{N_v}{2j_v+1}, \tag{11.30}$$

$$\gamma^{[2]}_{qr,qr} = p_{qr} + \sum_{n=3}^{2j_v+1} \sum_{v_3<\cdots<v_n} p_{\{v_3\cdots v_n\}} \stackrel{!}{=} \left(\frac{N_v}{2j_v+1}\right)^2 (1-\delta_{qr}), \tag{11.31}$$

$$\gamma^{[2]}_{qrs,qrs} = p_{qrs} + \sum_{n=4}^{2j_v+1} \sum_{v_4<\cdots<v_n} p_{\{v_4\cdots v_n\}} \stackrel{!}{=} \left(\frac{N_v}{2j_v+1}\right)^3 (1-\delta_{qr}-\delta_{qs}-\delta_{rs}+2\delta_{qr}\delta_{rs}), \tag{11.32}$$

$$\vdots$$

with $0 < N_v < 2j_v + 1$ being the number of particles in the valence orbital, that allow to uniquely determine the p_{core} and $p_{\{v_1\cdots v_n\}}$ coefficients in Eq. (11.23). Solving this set of equations we obtain

$$p_{\mathrm{core}} = (1+x)^{-(2j_v+1)}, \tag{11.33}$$

$$p_{\{v_1\cdots v_n\}} = \frac{x^n}{(1+x)^{2j_v+1}}[v_1\cdots v_n], \quad [v_1\cdots v_n] = \begin{cases} 1 & \text{if } v_k\text{'s are pairwise distinct} \\ 0 & \text{else} \end{cases} \tag{11.34}$$

with $x = N_v/(2j_v+1-N_v)$. Note that this solution correctly collapses to the closed-shell case for $N_v \to 0$ and $N_v \to 2j_v + 1$.

This concludes the generalized ansatz for the energy functional that allows us to construct a HF ground state for an open-shell system and, in particular, singly-strange hypernuclei.

11.3 Many-Body Perturbation Theory

The MBPT is a conceptually simple yet powerful method to calculate perturbative corrections to the known solution of an unperturbed problem [171, 174, 175]. Starting point for the MBPT is once more the many-body Schrödinger equation as given in Eq. (1.1). The Hamiltonian

$$\hat{H} = \hat{H}_0 + \lambda \hat{W} \tag{11.35}$$

is partitioned into an unperturbed part $\hat{H}_0$, for which the solution of the many-body problem is given by

$$\hat{H}_0 |\Psi_n^{(0)}\rangle = E_n^{(0)} |\Psi_n^{(0)}\rangle \quad \text{with} \quad \langle \Psi_n^{(0)} | \Psi_m^{(0)} \rangle = \delta_{nm}, \tag{11.36}$$

and a perturbation $\hat{W}$. The auxiliary parameter λ controls the perturbation and allows to organize the perturbative corrections by orders of λ. For the energy eigenvalues and eigenstates we, therefore, employ an expansion in a power series in terms of λ such that

$$E_n(\lambda) = \sum_{p=0}^{\infty} \lambda^p E_n^{(p)}, \tag{11.37}$$

$$|\Psi_n\rangle = \sum_{p=0}^{\infty} \lambda^p |\Psi_n^{(p)}\rangle. \tag{11.38}$$

Substituting this into the Schrödinger equation and separating by orders of λ leads to

$$\hat{H}_0 |\Psi_n^{(0)}\rangle = E_n^{(0)} |\Psi_n^{(0)}\rangle, \tag{11.39}$$

$$\hat{H}_0 |\Psi_n^{(p)}\rangle + \hat{W} |\Psi_n^{(p-1)}\rangle = \sum_{j=0}^{p} E_n^{(j)} |\Psi_n^{(p-j)}\rangle, \tag{11.40}$$

where the 0-th order correctly reproduces the eigenvalue relation for the unperturbed Hamiltonian. From the second equation we can derive a general expression for the p-th order energy correction by projecting it onto a reference state, i.e., a non-degenerate eigenstate of the unperturbed Hamiltonian $|\Psi_n^{(0)}\rangle$ and require intermediate normalization

$$\langle \Psi_n^{(0)} | \Psi_m^{(p)} \rangle = \begin{cases} \delta_{nm} & \text{if } p = 0 \\ 0 & \text{else} \end{cases}, \tag{11.41}$$

such that

$$E_n^{(p)} = \langle \Psi_n^{(0)} | \hat{W} | \Psi_n^{(p-1)} \rangle. \tag{11.42}$$

In order to evaluate this, an expression for the state corrections is required. They can be expended in terms of the basis of the unperturbed Hamiltonian

$$|\Psi_n^{(p)}\rangle = \sum_m |\Psi_m^{(0)}\rangle \langle \Psi_m^{(0)} | \Psi_n^{(p)} \rangle \equiv \sum_m |\Psi_m^{(0)}\rangle C_{nm}^{(p)} \tag{11.43}$$

such that the evaluation of both, the energy and state corrections, is reduced to the determination of the expansion coefficients $C_{nm}^{(p)}$. Projecting Eq. (11.40) once more onto $|\Psi_{m\neq n}^{(0)}\rangle$ yields

$$C_{nm}^{(p)} = \frac{1}{E_n^{(0)} - E_m^{(0)}} \left(\sum_{l\neq m} \langle \Psi_m^{(0)} | \hat{W} | \Psi_l^{(0)} \rangle C_{nl}^{(p-1)} - \sum_{j=1}^{p} E_n^{(j)} C_{nm}^{(p-j)} \right). \tag{11.44}$$

With this we have a recursive set of equations that allows us to compute perturbative corrections up to an arbitrary order.

Hartree-Fock Many-Body Perturbation Theory. Up to this point the derivation has been completely general. In order to construct the perturbation series required for the calculation of the natural orbitals the Hamiltonian and its partitioning need to be specified. Since we want to construct the MBPT on top of a HF state, we consider a two-body Hamiltonian in HF basis [1]. For the partitioning we employ a Møller-Plesset ansatz

$$\hat{H}_0 = E_{\mathrm{HF}} + \sum_{pq} h_{p,q} \hat{b}_p^\dagger \hat{b}_q, \tag{11.45}$$

$$\hat{W} = \frac{1}{4} \sum_{pqrs} H_{pq,rs}^{[2]} \hat{b}_p^\dagger \hat{b}_q^\dagger \hat{b}_s \hat{b}_r, \tag{11.46}$$

where E_{HF} and h are the previously derived HF energy and mean-field Hamiltonian. We further denote the eigenbasis of the unperturbed Hamiltonian as $\{|\Phi_i\rangle\}$ with corresponding eigenvalues E_i and choose $|\Phi_0\rangle \equiv |\Phi_{\mathrm{ref}}\rangle$ as reference state with $E_0 \equiv E_{\mathrm{ref}}$. With this we can evaluate the first-order state corrections and find

$$\begin{aligned} |\Psi_{\mathrm{HF}}^{(1)}\rangle &= \sum_{m\neq n} |\Phi_m\rangle \frac{\langle \Phi_m | \hat{W} | \Phi_{\mathrm{ref}} \rangle}{E_{\mathrm{ref}} - E_m} \\ &= \sum_{\substack{i<j \\ a<b}} |\Phi_{ij}^{ab}\rangle \frac{H_{ij,ab}^{[2]}}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \end{aligned} \tag{11.47}$$

where we have switched to particle-hole formalism in order to simplify the expressions. Analogously we can derive the second-order correction

$$\begin{aligned} |\Psi_{\mathrm{HF}}^{(2)}\rangle &= \sum_{\substack{m\neq n \\ l\neq m}} |\Phi_m\rangle \frac{\langle \Phi_m | \hat{W} | \Phi_l \rangle \langle \Phi_l | \hat{W} | \Phi_{\mathrm{ref}} \rangle}{(E_{\mathrm{ref}} - E_m)(E_{\mathrm{ref}} - E_l)} \\ &= \frac{1}{4} \sum_{\substack{i<j \\ a<b}} \sum_{\substack{k<l \\ c<d}} \sum_{\substack{pq \\ rs}} |\Phi_{kl}^{cd}\rangle \frac{H_{pq,rs}^{[2]} H_{ij,ab}^{[2]}}{(\epsilon_k + \epsilon_l - \epsilon_c - \epsilon_d)(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)} \langle \Phi_{kl}^{cd} | \hat{b}_p^\dagger \hat{b}_q^\dagger \hat{b}_s \hat{b}_r | \Phi_{ij}^{ab} \rangle. \end{aligned} \tag{11.48}$$

[1] Note that these consideration can be generalized to include three-body or higher forces, but a two-body Hamiltonian is sufficient for our applications.

11.4 Hypernuclear Natural Orbitals

Starting point for our derivation of the natural orbitals is the one-body density matrix in terms of HF creation and annihilation operators

$$\rho_{p,q} = \langle \Psi | \hat{b}_p^\dagger \hat{b}_q | \Psi \rangle \tag{11.49}$$

with hypernuclear many-body state $|\Psi\rangle$, which is approximated as the second-order corrected HF ground state using HF-MBPT

$$|\Psi\rangle \approx |\Phi_{\mathrm{HF}}\rangle + \lambda |\Psi_{\mathrm{HF}}^{(1)}\rangle + \lambda^2 |\Psi_{\mathrm{HF}}^{(2)}\rangle. \tag{11.50}$$

In this approximation the perturbatively corrected density becomes

$$\begin{aligned}
\rho_{p,q} &\approx \langle \Phi_{\mathrm{HF}} | \hat{b}_p^\dagger \hat{b}_q | \Phi_{\mathrm{HF}} \rangle + \langle \Psi_{\mathrm{HF}}^{(1)} | \hat{b}_p^\dagger \hat{b}_q | \Phi_{\mathrm{HF}} \rangle + \langle \Phi_{\mathrm{HF}} | \hat{b}_p^\dagger \hat{b}_q | \Psi_{\mathrm{HF}}^{(1)} \rangle \\
&\quad + \langle \Psi_{\mathrm{HF}}^{(1)} | \hat{b}_p^\dagger \hat{b}_q | \Psi_{\mathrm{HF}}^{(1)} \rangle + \langle \Psi_{\mathrm{HF}}^{(2)} | \hat{b}_p^\dagger \hat{b}_q | \Phi_{\mathrm{HF}} \rangle + \langle \Phi_{\mathrm{HF}} | \hat{b}_p^\dagger \hat{b}_q | \Psi_{\mathrm{HF}}^{(2)} \rangle + \mathcal{O}(\lambda^3) \\
&\equiv \rho_{p,q}^{\mathrm{HF}} + \rho_{p,q}^{(10)} + \rho_{p,q}^{(01)} + \rho_{p,q}^{(11)} + \rho_{p,q}^{(20)} + \rho_{p,q}^{(02)},
\end{aligned} \tag{11.51}$$

where we have neglected any third- or higher-order terms in λ. Apparently, the corrected density decomposes into the unperturbed HF density and corrections corresponding to different orders, which we can address individually.

Since we are working in a HF-MBPT framework we can employ Brillouin's theorem that prevents the direct mixing of singly-excited HF Slater determinants with the HF ground state [2]. Hence, the corrections at $\lambda = 1$, i.e., $\rho^{(10)}$ and $\rho^{(01)}$ vanish. We can further exploit the symmetry of ρ which holds

$$\rho_{p,q}^{(02)} = \rho_{q,p}^{(20)}. \tag{11.52}$$

Thus, we are left with the determination of $\rho^{(11)}$ and $\rho^{(02)}$. Starting with the former, we can plug in the expression for the first-order state correction from Eq. (11.47) and find

$$\rho_{p,q}^{(11)} = \sum_{\substack{i<j \\ a<b}} \sum_{\substack{i'<j' \\ a'<b'}} \langle \Phi_{i'j'}^{a'b'} | \hat{b}_p^\dagger \hat{b}_q | \Phi_{ij}^{ab} \rangle \frac{H_{i'j',a'b'}^{[2]} H_{ij,ab}^{[2]}}{(\epsilon_{i'} + \epsilon_{j'} - \epsilon_{a'} - \epsilon_{b'})(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)} \tag{11.53}$$

and analogously under the use of Eq. (11.48) the second correction yields

$$\begin{aligned}
\rho_{p,q}^{(02)} &= \frac{1}{4} \sum_{\substack{i<j \\ a<b}} \sum_{\substack{k<l \\ c<d}} \sum_{\substack{rs \\ tu}} \langle \Phi_{\mathrm{HF}} | \hat{b}_p^\dagger \hat{b}_q | \Phi_{kl}^{cd} \rangle \frac{H_{rs,tu}^{[2]} H_{ij,ab}^{[2]}}{(\epsilon_k + \epsilon_l - \epsilon_c - \epsilon_d)(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)} \langle \Phi_{kl}^{cd} | \hat{b}_r^\dagger \hat{b}_s^\dagger \hat{b}_u \hat{b}_t | \Phi_{ij}^{ab} \rangle \\
&= \frac{1}{4} \sum_{\substack{i<j \\ a<b}} \sum_{\substack{k \\ c}} \sum_{\substack{rs \\ tu}} \langle \Phi_{\mathrm{HF}} | \hat{b}_r^\dagger \hat{b}_s | \Phi_k^c \rangle \frac{H_{rs,tu}^{[2]} H_{ij,ab}^{[2]}}{(\epsilon_k - \epsilon_c)(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)} \langle \Phi_k^c | \hat{b}_r^\dagger \hat{b}_s^\dagger \hat{b}_u \hat{b}_t | \Phi_{ij}^{ab} \rangle,
\end{aligned} \tag{11.54}$$

[2] Note that indirect mixing can and does still occur in higher orders of the perturbation.

where we have used that a single pair of creation and annihilation operators can only connect the ground state to singly-excited Slater determinants. Conveniently, these terms can be further simplified by separating certain particle-hole terms as done in [168] and take the form

$$\rho^{(02)} = D^{(A)} + D^{(B)}, \tag{11.55}$$

$$\rho^{(11)} = D^{(C)} + D^{(D)}, \tag{11.56}$$

where

$$D^{(A)}_{ia} = \frac{1}{2}\sum_{jbc} \frac{H^{[2]}_{ij,ab} H^{[2]}_{bc,aj}}{(\epsilon_i - \epsilon_a)(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)}, \tag{11.57}$$

$$D^{(B)}_{ia} = -\frac{1}{2}\sum_{jkb} \frac{H^{[2]}_{ia,jk} H^{[2]}_{jk,ab}}{(\epsilon_i - \epsilon_a)(\epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b)}, \tag{11.58}$$

$$D^{(C)}_{ij} = -\frac{1}{2}\sum_{abk} \frac{H^{[2]}_{ik,ab} H^{[2]}_{ab,jk}}{(\epsilon_i + \epsilon_k - \epsilon_a - \epsilon_b)(\epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b)}, \tag{11.59}$$

$$D^{(D)}_{ab} = \frac{1}{2}\sum_{cij} \frac{H^{[2]}_{ac,ij} H^{[2]}_{ij,bc}}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_c)(\epsilon_i + \epsilon_j - \epsilon_b - \epsilon_c)}. \tag{11.60}$$

By diagonalizing the resulting corrected density matrix we obtain a new single-particle basis $\{|c_p\rangle\}$, which are the natural orbitals we are looking for. They can directly be calculated from the HF basis $\{|b_p\rangle\}$

$$|c_p\rangle = \sum_q \tilde{C}_{pq} |b_q\rangle, \tag{11.61}$$

or from the HO basis $\{|b_a\rangle\}$

$$|c_p\rangle = \sum_{qr} \tilde{C}_{pq} C_{qr} |a_r\rangle \equiv \sum_r \tilde{C}'_{pr} |a_r\rangle \tag{11.62}$$

respectively. Due to the symmetries of the Hamiltonian the density matrix is block-diagonal in $l, j, m_j, \mathcal{S}, t$, and m_t. In particular, there is no particle mixing and the derivation does, therefore, not require any adjustments when considering hyperons as additional constituents. Hence, it can directly by applied to hypernuclei.

Note that in actual calculations we include 3N forces in the initial Hamiltonian for the HF calculation in order to obtain a realistic HF ground-state. For simplicity and computational efficiency, however, we limit the subsequent calculation of perturbative corrections to a two-body Hamiltonian, which we obtain by normal ordering the Hamiltonian and truncating at the two-body rank, known as the normal-ordered two-body (NO2B) approximation, which we will discuss in detail in the context of in-medium methods in the final part of this work. Technically, this does not only truncate the Hamiltonian but additionally induces one-body terms in Eq. (11.46), which arise from the normal-ordering and complicate the derivation of the state corrections (see Ch. 6 in [42]). We, therefore, choose to neglect these one-body terms, as we assume the error to be very small. This does not effect the completeness of the resulting NAT basis, which remains a valid choice of single-particle basis.

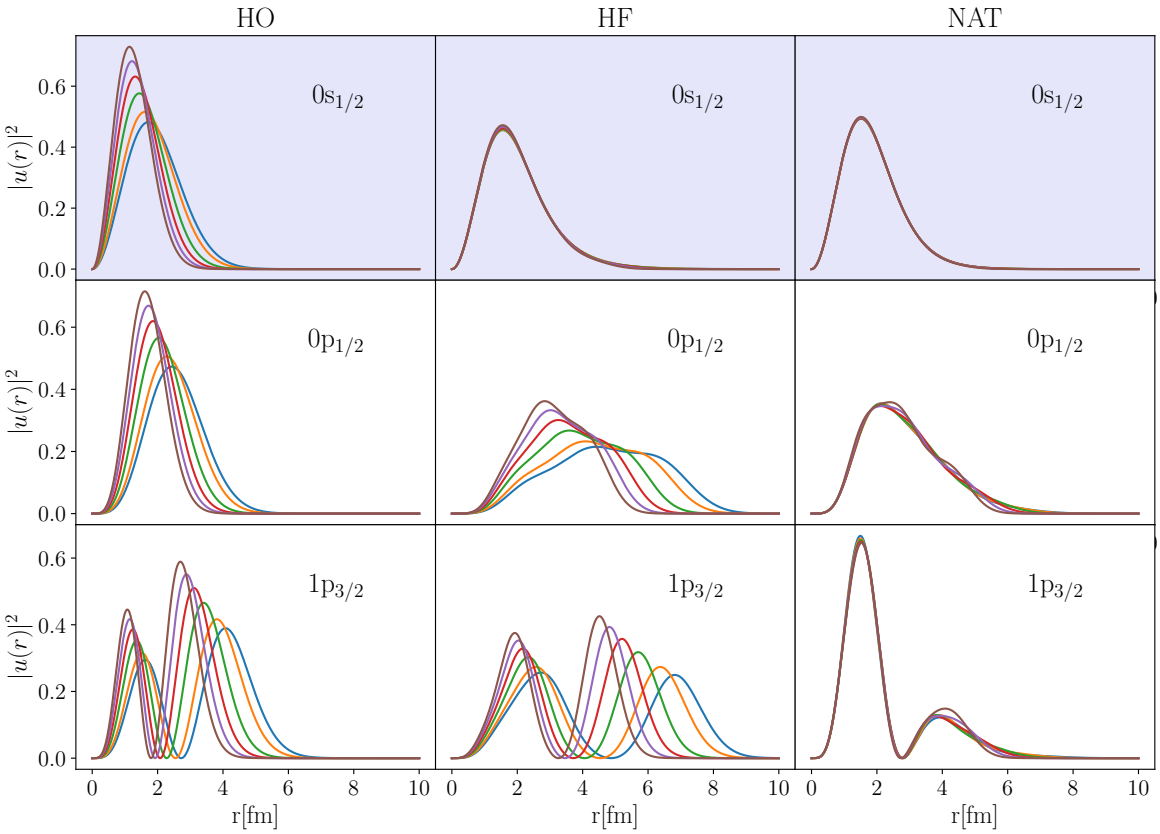

Figure 11.1: Squared radial wavefunctions for the Λ particle in selected single-particle orbitals (top to bottom) in the HO, HF, and NAT bases (left to right) for ${}^{5}_{\Lambda}\mathrm{He}$. Different colors indicate different HO frequencies with $\hbar\Omega = 14$ (—), 16 (—), 20 (—), 24 (—), 28 (—), and 32 MeV (—). Blue background marks occupied orbitals.

Single-Particle Wavefunctions. In order to investigate the different single-particle bases, and the natural orbitals in particular, we study the squared radial wavefunctions $|u(r)|^2$ for a given particle species and orbital. The following calculations have been performed with the $\mathrm{NN_{EMN}} + 3\mathrm{N_H} + \mathrm{YN_{opt}}$ combination of interactions consistently SRG evolved to $\alpha = 0.08$ fm^4. We further employ single-particle truncations $e_{\max} = 12$ and $l_{\max} = 8$ and three-body truncations $E_{3\max} = 14$ in the 3N interaction and $E_{3\max} = 12$ for the induced YNN forces. Figure 11.1 shows wavefunctions for the Λ particle for the HO, HF, and NAT bases in selected orbitals. Moreover, wavefunctions for six different HO frequencies are given. Since the HF and NAT bases are obtained from a HO basis the choice of $\hbar\Omega$ transfers to those bases as well. For the occupied $0s_{1/2}$ orbital we find that the strong frequency dependence in the HO basis, which arises from the direct correspondence of $\hbar\Omega$ to the width of the radial wavefunction, vanishes in the HF and NAT bases. They further do not differ much from the HO wavefunctions in shape or spatial extend. For the higher-lying unoccupied orbitals the picture is very different. The shape of the wavefunction changes drastically in the HF and NAT bases and the HF wavefunctions retain a very strong $\hbar\Omega$ dependence. The natural orbitals, however, are again independent of the HO frequency, which is one of their key features. Any remaining $\hbar\Omega$ dependencies, which mostly occur at longer ranges can be attributed to the initial truncation of the HO basis since wider HO wavefunctions are cut from the basis. Hence, they are not available for the expansion of the NAT states in the HO basis, which results in the

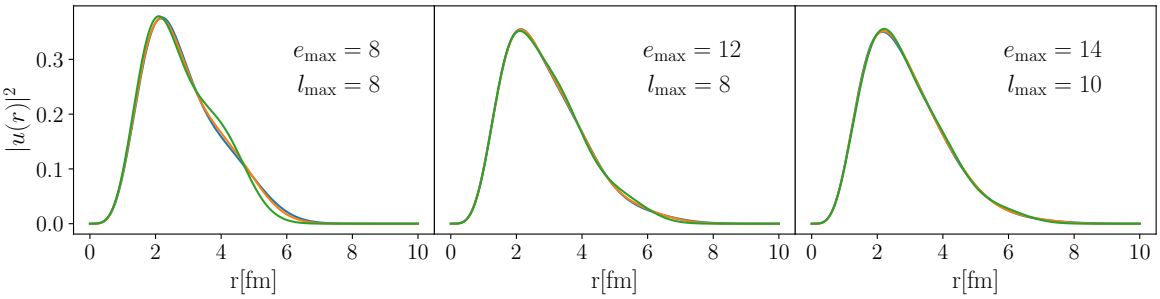

Figure 11.2: Squared NAT radial wavefunctions for the Λ particle in the $0\mathrm{p}_{1/2}$ orbital in ${}^{5}_{\Lambda}\mathrm{He}$ for different model-space truncations $e_{\max}$ and $l_{\max}$. Same color scheme for HO frequencies as in Fig. 11.1.

residual frequency dependence at longer distances. Based on calculations in different model spaces shown in Fig. 11.2 we can confirm that this is a truncation issue as the $\hbar\Omega$ dependencies are smoothed out successively with increasing model-space size. We can, therefore, conclude that the concept of natural orbitals can easily be transferred to hypernuclei and the benefit of frequency independence carries over to hyperons. While we have limited our discussion to Λ orbitals so far, we can also look at all six particle species that are active in our calculations. In Fig. 11.3 the $1\mathrm{p}_{3/2}$ wavefunctions for ${}^{5}_{\Lambda}\mathrm{He}$ are depicted for the six particle species. Regardless of the particular shape of the wavefunction we find that the results for neutron and proton look very much alike, which is to be expected as they mainly differ by the Coulomb interaction. The same holds for the three Σ wavefunctions, which are again nearly identical. We should also note that there are no Σ hyperons present in the HF ground state. Hence, they only enter perturbatively through the Λ-Σ mixing, which typically causes effects at a percentage level and is, thus, very weak. We, therefore, expect the NAT wavefunctions for the Σ particles to be at least partially dominated by the orthogonality constraints on the basis.

Now that we have studied the systematic behavior of the NAT single-particle basis, we can investigate different hypernuclei and use the wavefunctions to learn about the structure of

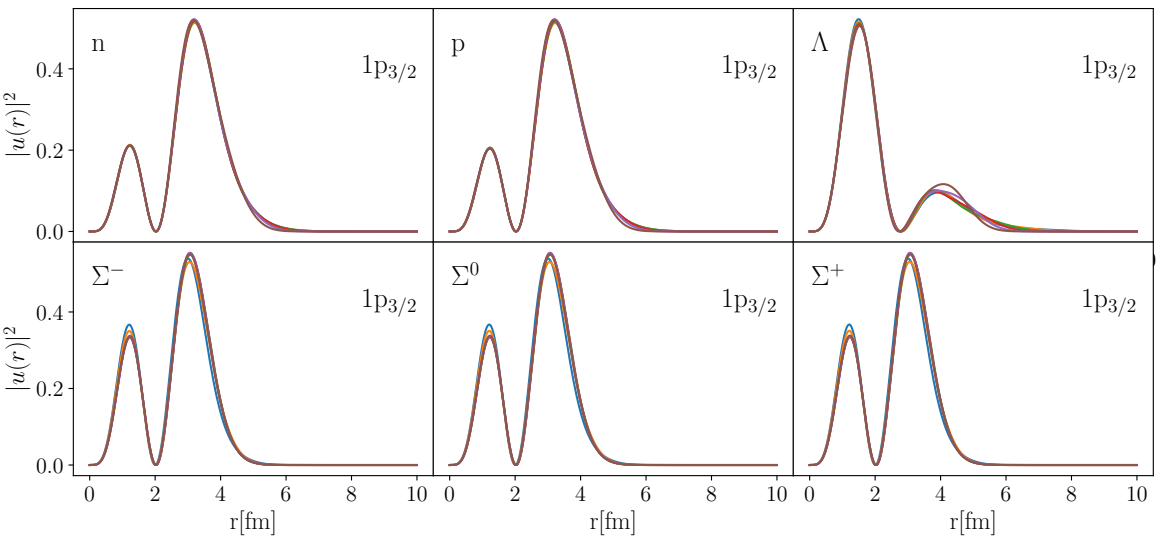

Figure 11.3: Squared NAT radial wavefunctions for all particle species in the $1\mathrm{p}_{3/2}$ orbital in ${}^{5}_{\Lambda}\mathrm{He}$. Same color scheme for HO frequencies as in Fig. 11.1.

those. Since the effect of the Coulomb interaction on the radial wavefunctions is negligible, we limit the discussion to the three neutral particles n, Λ, and Σ^0. One of the most interesting cases is ${}^{5}_{\Lambda}\mathrm{He}$ for which a complete overview of the five lowest-lying orbitals up to the $1\mathrm{s}_{1/2}$ orbital is shown in Fig. 11.4. The occupied orbitals are highlighted with a blue background. For the Λ wavefunctions in the center column we find that they are stretched to larger radii, especially in contrast to the nucleonic wavefunctions in the left-hand column. This stretching is particularly pronounced in the $1\mathrm{s}_{1/2}$ orbital, where the suppression of the short-range part even results in a strong HO frequencies of the wavefunctions due to the lack of long-range states in the HO basis. Apparently, the Λ particle tends to sit at larger distances, which is contrary to the naïve expectation that it would sit in the center of the nucleus as it is not restrained by the Pauli principle [3]. Given that the hyperon separation energy of ${}^{5}_{\Lambda}\mathrm{He}$ with $B_\Lambda = 3.12(2)$ MeV [7] is much smaller than the neutron separation energy of $S_\mathrm{n} = 20.578$ MeV [176] it seems likely that ${}^{5}_{\Lambda}\mathrm{He}$ can be understood as a hyperon coupled to an α particle, maybe even hinting at a Λ halo. The Σ^0 wavefunctions, however, show no anomalies and appear to be slightly less extended compared to the neutron orbitals.

When studying heavier nuclei such as ${}^{17}_{\Lambda}\mathrm{O}$, for which the wavefunctions are depicted in Fig. 11.5 we find that the Λ orbitals are much less extended compared to ${}^{5}_{\Lambda}\mathrm{He}$ and even favor shorter distances than the neutron orbitals. This is more compatible with the Pauli picture and is additionally supported by the fact that the mass radius of ${}^{17}_{\Lambda}\mathrm{O}$ is smaller than the mass radius of ${}^{16}\mathrm{O}$. It further results in a reduction of the $\hbar\Omega$ dependence, which only remains visible for the long-range part of the $1\mathrm{s}_{1/2}$ orbital. The shape of the Σ^0 orbitals remains about the same as before. These trends continue for the radial wavefunctions of ${}^{41}_{\Lambda}\mathrm{Ca}$, which are shown in Fig. 11.6. Again, the Λ orbitals favor shorter radii compared to the neutron orbitals and even higher-lying orbitals like the $1\mathrm{s}_{1/2}$ are dominated by the short-range component. We now find some minor $\hbar\Omega$ dependence across all orbitals and particles. This can most likely also be attributed to the limited model space, which seems to be insufficient to fully describe this significantly larger system.

Overall the natural orbitals provide a computationally efficient yet powerful method to optimize the single-particle basis not only for nucleonic systems but also hypernuclei. While the most striking feature is the independence of the HO frequency, they also serve as a great diagnostic tool to study the radial extend of different particle species, giving inside into basic structural aspects of the respective hypernuclei.

[3] This only holds for the singly-strange hypernuclei under consideration. Since hyperons are fermions they clearly must obey the Pauli principle as soon as there are more than one present.

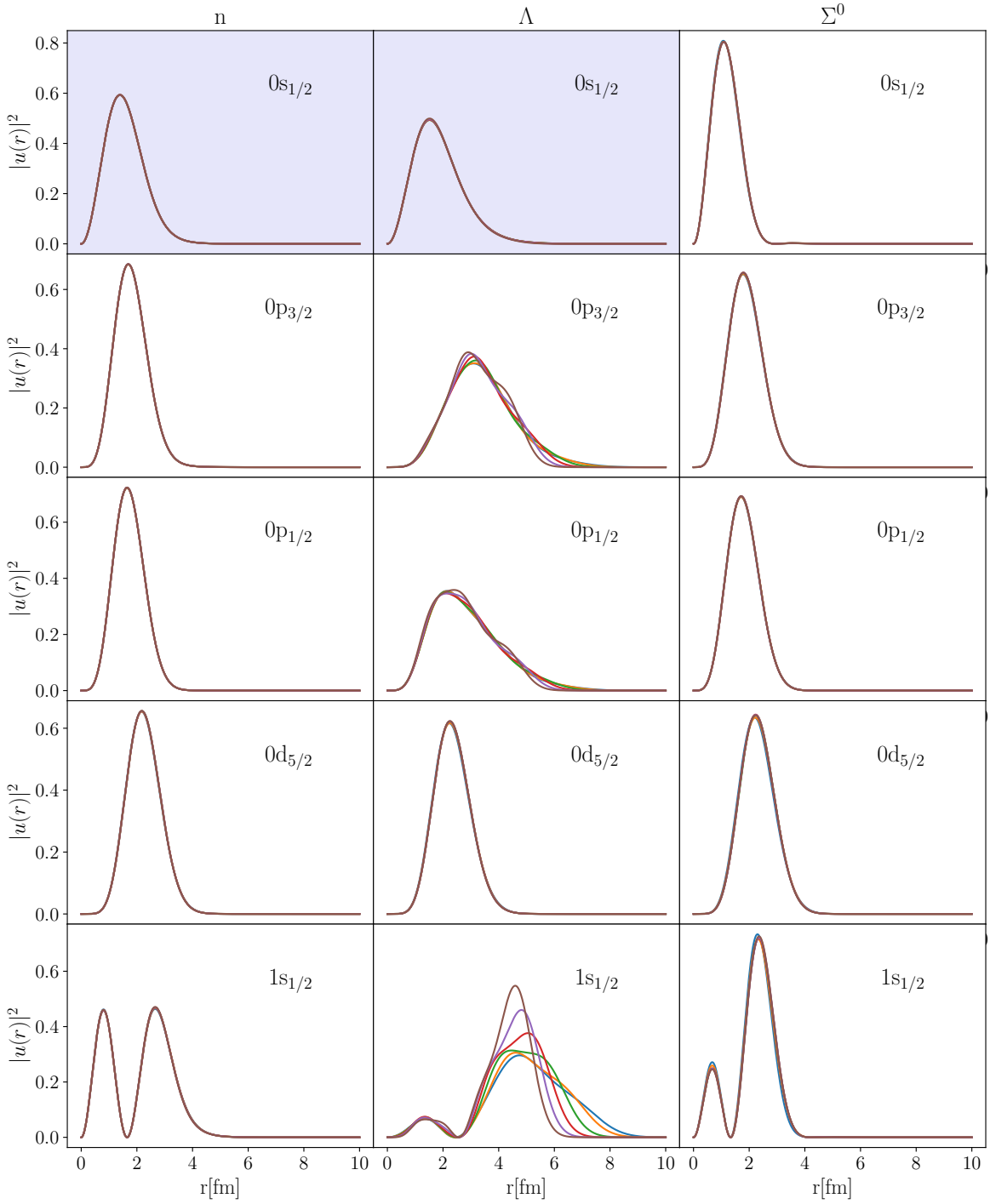

Figure 11.4: Squared NAT radial wavefunctions for neutrons, Λ, and Σ^0 particles in the lowest-lying orbitals for ${}^{5}_{\Lambda}$He. Blue background indicates occupied orbitals. Same color scheme for HO frequencies as in Fig. 11.1.

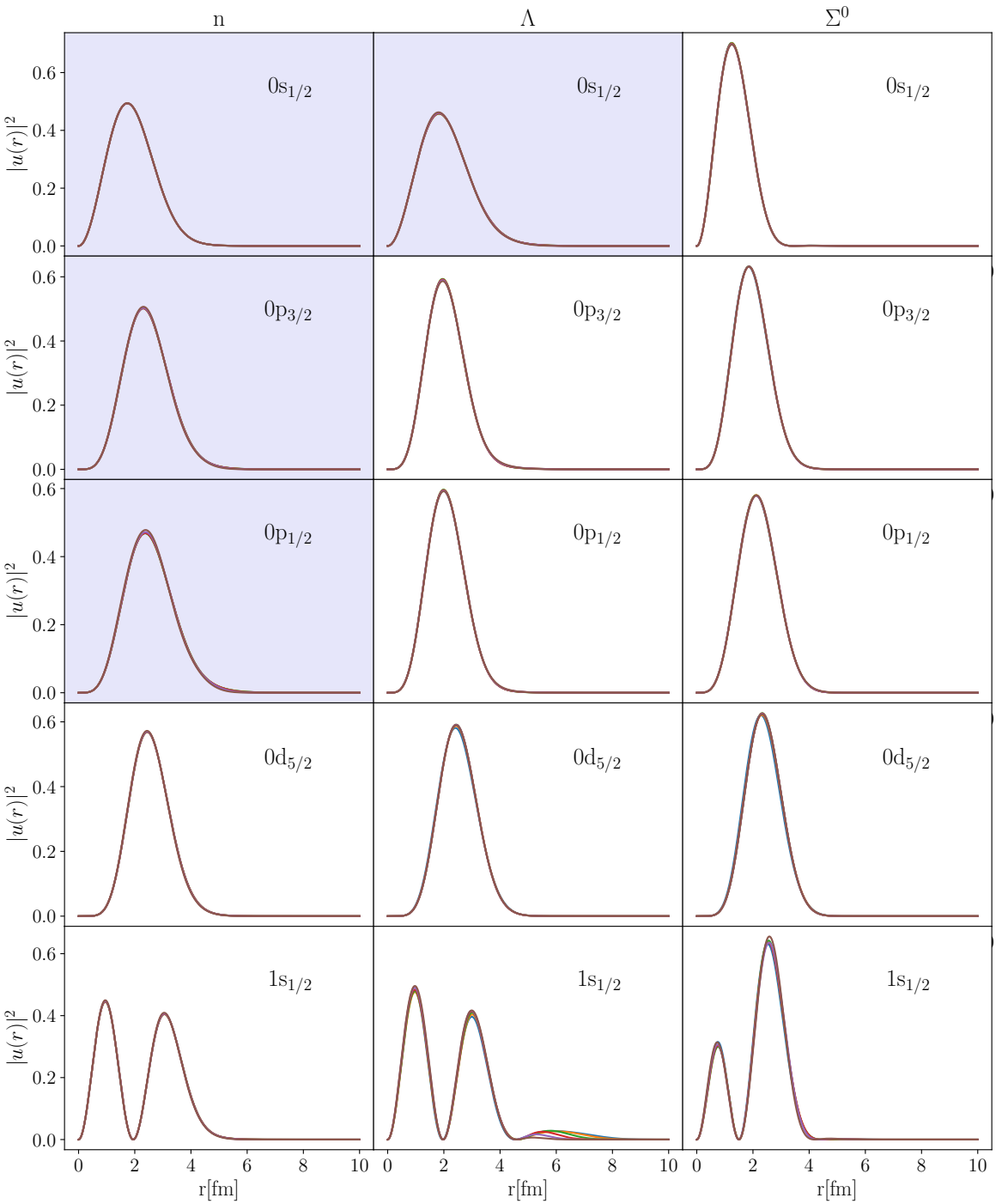

Figure 11.5: Squared NAT radial wavefunctions for neutrons, Λ, and Σ^0 particles in the lowest-lying orbitals for ${}^{17}_{\Lambda}\mathrm{O}$. Blue background indicates occupied orbitals. Same color scheme for HO frequencies as in Fig. 11.1.

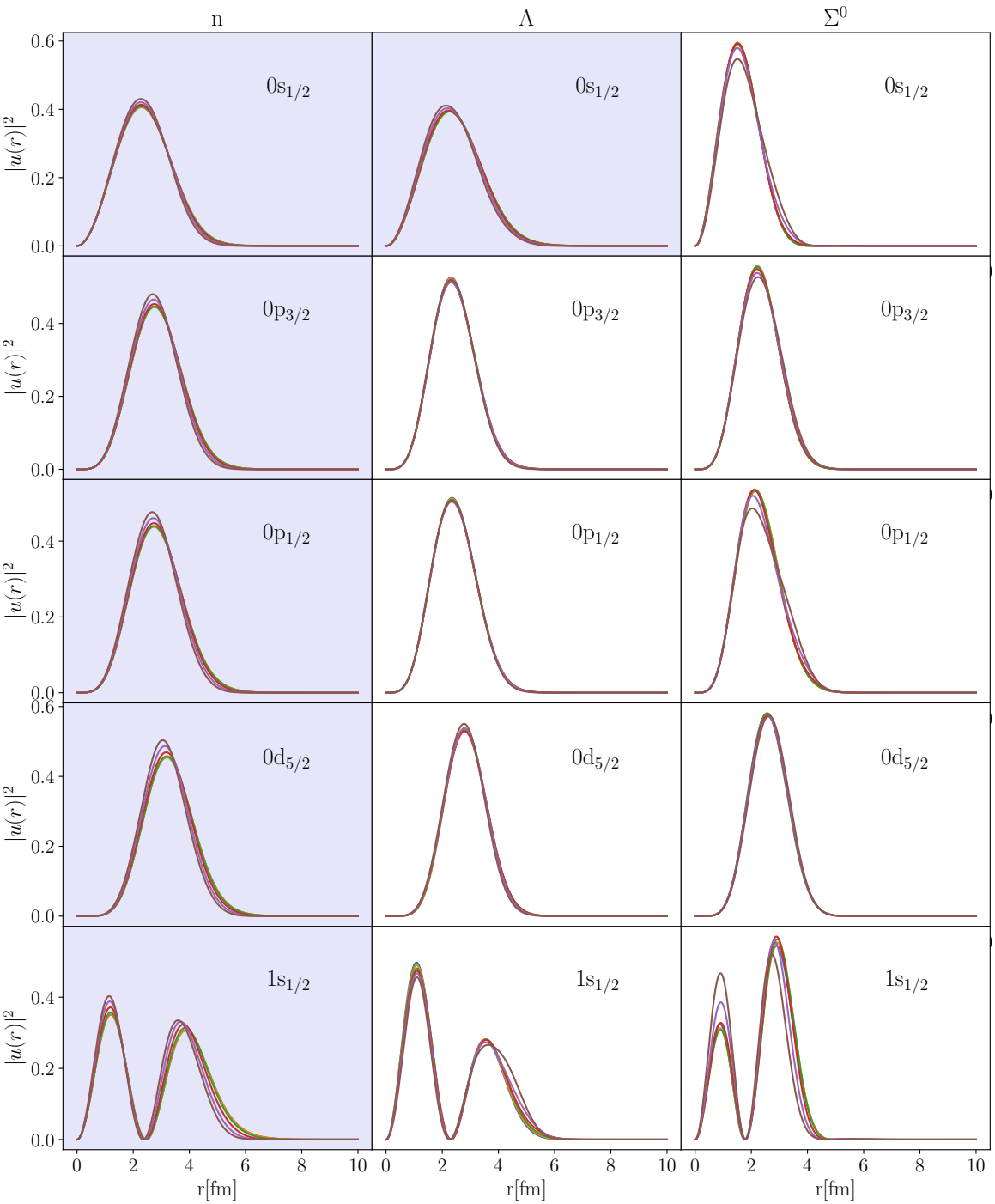

Figure 11.6: Squared NAT radial wavefunctions for neutrons, Λ, and Σ^0 particles in the lowest-lying orbitals for ${}^{41}_{\Lambda}\mathrm{Ca}$. Blue background indicates occupied orbitals. Same color scheme for HO frequencies as in Fig. 11.1.

11.5 Natural Orbitals in the Hypernuclear NCSM

While the NAT basis and the associated single-particle wavefunctions yield interesting physics insights themselves, our initial goal was to employ them as an optimized single-particle basis in many-body calculations. We keep the previous choice of interaction and truncation, i.e., the $\mathrm{NN_{EMN}} + 3\mathrm{N_H} + \mathrm{YN_{opt}}$ combination with $e_{\max} = 12$ and $l_{\max} = 8$. In the NAT basis we limit ourselves to calculations with NO2B approximation, which reduces the computational effort and allows us to access larger model spaces since we do not have to deal with explicit three-body forces. It further permits us to reuse the matrix elements in other applications discussed in part IV.

Before we can employ the NAT basis in NCSM calculations we have to reconsider the aspects of the NCSM that exploit properties of the HO basis, since some of them are not valid for the NAT basis anymore. In particular, the removal of center-of-mass contributions was based on the separation of intrinsic and center-of-mass parts in the HO many-body wavefunctions. This separation is not given in the NAT basis, however, we can adapt the concept of shifting the spurious center-of-mass excitations to higher parts of the spectrum in a similar manner as for the HO basis. Thus, we again modify our Hamiltonian the same way as given in Eqs. (2.6) and (2.7) keeping in mind that this is not exact and the obtained states will contain residual center-of-mass contributions. In order to keep the latter minimal we have to choose small values of λ_{cm}. We, therefore, use $\lambda_{\mathrm{cm}} = 0.3$ in the subsequent calculations.

We start our investigations by comparing NCSM calculations for ${}^{5}_{\Lambda}\mathrm{He}$ in, both, HO and NAT basis as shown in Fig. 11.7. We immediately find that the independence of $\hbar\Omega$ in the NAT wavefunctions carries over to the many-body calculations, which is in strong contrast to the HO calculations.

Furthermore, we find that the convergence rate in the NAT basis is significantly increased even compared to the optimal HO frequency. This goes beyond the expectations from calculations for nucleonic systems, where the convergence rate in the NAT basis is commonly found

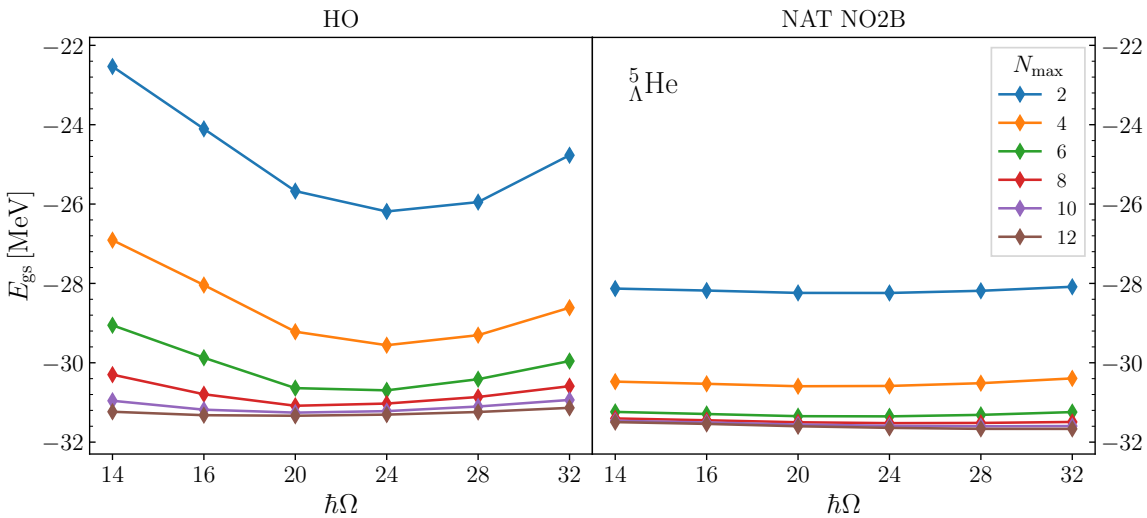

Figure 11.7: NCSM calculations for ${}^{5}_{\Lambda}\mathrm{He}$ in HO basis (left-hand panel) and in NAT basis with NO2B approximation (right-hand panel). The calculations in HO basis contain 3N and YNN forces explicitly.

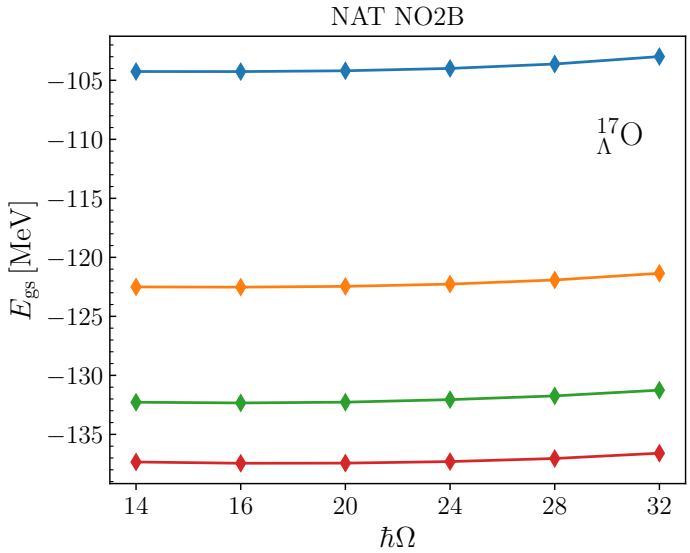

Figure 11.8: NCSM calculations for $^{17}_{\Lambda}$O in NAT basis with NO2B approximation. Same color coding as in Fig. 11.7.

to be equal to the convergence rate for the optimal HO frequency. Finally, we notice that the limits of the HO and NAT calculations differ by about 500 keV, which is contrarily to what would be expected for two complete bases. This is due to the NO2B approximation employed for the NAT basis, and should vanish if the 3N and YNN forces are consistently transformed and fully taken into account.

The systematics we have found for $^{5}_{\Lambda}$He are not limited to light hypernuclei but do also apply to heavier p-shell hypernuclei such as $^{17}_{\Lambda}$O as shown in Fig. 11.8. While we, again, find frequency independent results and a systematic convergence pattern, the convergence rate is not fast enough to obtain converged results. This marks a limit to the reach of the IT-NCSM for hypernuclei and demonstrates the need for developments towards ab initio medium-mass methods for hypernuclei once more.

Overall, we conclude that both the $\hbar\Omega$ independence and the improved convergence rate are beneficial properties making the NAT basis the ideal choice of basis for hypernuclear many-body calculations. However, besides these beneficial features, the NAT basis conflicts with the application of our ANN tool, which is designed to explicitly take advantage of the $\hbar\Omega$ dependence. Hence, the application within the NAT basis would effectively boil down to the $F = 1$ case in Fig. 6.1, that is deficient in accuracy and precision compared to higher numbers of F.

A Hyperon Halo? There is one other aspect that has sparked our interest and that we can investigate further using the natural orbitals. From our study of the NAT single-particle wavefunctions we have found that the Λ orbitals in $^{5}_{\Lambda}$He shown in Fig. 11.4 exhibit long-range tails. Together with the previously discussed gap between the hyperon separation energy of $^{5}_{\Lambda}$He and its neutron separation energy, this suggests that the $^{5}_{\Lambda}$He hypernucleus can be interpreted as a hyperon loosely bound to an α core, which is known as a halo structure. The occurrence of this phenomenon has already been proposed and studied for other hypernuclei [177–182].

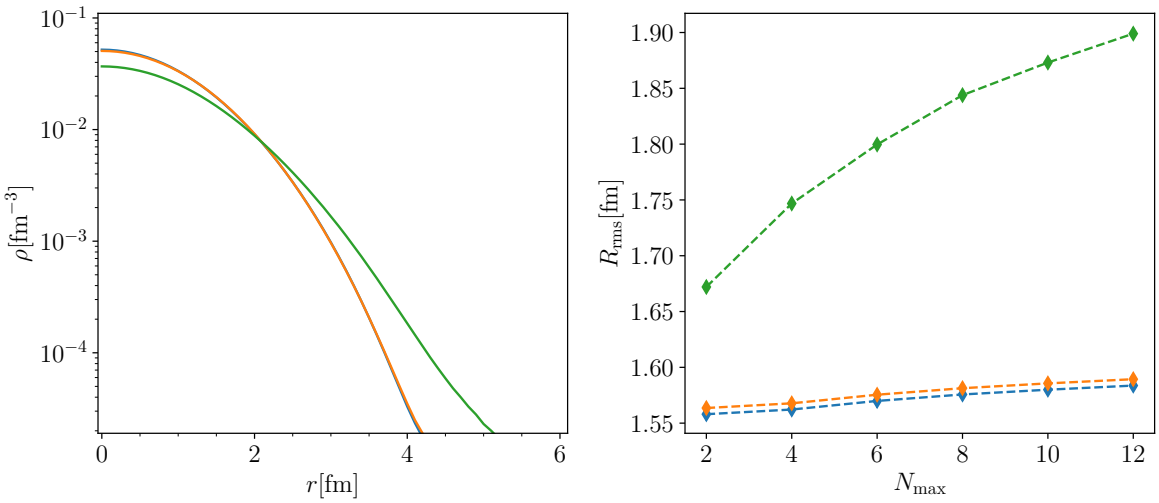

Figure 11.9: One-body densities normalized to one (left-hand panel) and point-particle rms-radii (right-hand panel) for neutrons (blue), protons (orange) and Λ hyperons (green) in ${}^{5}_{\Lambda}\mathrm{He}$. The densities are obtained from a NCSM calculation in NAT basis with NO2B approximation for $\hbar\Omega = 14$ MeV at $N_{\max} = 12$.

In order to investigate this more closely we study the one-body densities in ${}^{5}_{\Lambda}\mathrm{He}$ shown in the left-hand panel of Fig. 11.9, which are calculated with many-body wavefunctions obtained from NCSM calculations at $N_{\max} = 12$. All densities shown here are normalized to 1 instead of the respective particle number to facilitate comparison. As already indicated by the NAT wavefunctions, the Λ density extends significantly beyond the nucleon densities, featuring a long hyperon tail.

In addition to the densities we calculate the point-Λ radius, which should also extend significantly beyond the point-particle radii for the nucleons in the system. According to [42] the operator corresponding to the mean-square radius for any particle species χ is given by

$$\begin{aligned}\hat{R}^2_{\chi,\mathrm{ms}} &= \frac{1}{\langle \hat{n}_\chi \rangle} \sum_{i=1}^{A} \hat{P}^\chi_i (\hat{\mathbf{r}}_i - \hat{\mathbf{R}}_{\mathrm{cm}})^2 \\ &= \frac{1}{\hat{M}\langle \hat{n}_\chi \rangle} \sum_{i<j} \left(\hat{m}_i \hat{P}^\chi_j + \hat{m}_j \hat{P}^\chi_i - n_\chi \frac{\hat{m}_i \hat{m}_j}{\hat{M}} \right) \hat{r}^2_{ij}, \end{aligned} \tag{11.63}$$

where we have introduced a particle projection operator along with a corresponding particle number operator

$$\hat{P}^\chi_i = \begin{cases} 1 & \text{if particle } i \text{ is of species } \chi, \\ 0 & \text{else} \end{cases}, \tag{11.64}$$

$$\hat{n}_\chi = \sum_{i=1}^{A} \hat{P}^\chi_i. \tag{11.65}$$

With this, we can calculate the point-χ rms-radii for the constituents of the ${}^{5}_{\Lambda}\mathrm{He}$ nucleus, which are shown in the right-hand panel of Fig. 11.9. Again, we find that the point-Λ rms-radius is much larger than the point-proton and point-neutron radii. Moreover, it requires significantly

larger model spaces to converge, which is typical for halo nuclei as highly-excited long-range basis states are required to account for the long-range contributions to such extended radii.

Overall, we find a very consistent picture that exhibits strong long-range components in the hyperon channel, which indicate that ${}^{5}_{\Lambda}\mathrm{He}$ is a candidate for a hyperon halo.

Part IV

Towards Ab Initio Theory for Medium-Mass Hypernuclei

12

In-Medium No-Core Shell Model

The IM-NCSM is a powerful ab initio many-body method for medium-mass nuclei [32, 183, 184]. It merges the NCSM with the IM-SRG into a hybrid approach that combines the strengths of both methods, as they complement each other regarding included correlations and truncation schemes. Conceptually, the underlying notion is to pre-diagonalize the Hamiltonian via the IM-SRG to boost the convergence in a subsequent NCSM calculation, thus, extending the reach of the NCSM beyond the p-shell into the medium-mass regime.

Being an "in-medium" method the IM-NCSM is constructed w.r.t. a reference state. For closed-shell systems a single Slater determinant is sufficient, while the application to open-shell nuclei requires more sophisticated reference states. Since we are mainly interested in the extension to hypernuclei, we limit the following discussion to the single-reference IM-SRG for the sake of simplicity. However, the extension to a multi-reference framework has been worked out and we refer the interested reader to [184] for further reading.

12.1 Normal Ordering

The concept of normal ordering is fundamental for the IM-SRG. It is based on the idea of formulating an operator w.r.t. a reference state instead of the vacuum and span the configuration space through particle-hole excitations acting on this state. Consequently, certain parts of many-body operators are reassigned to lower particle ranks. This allows to truncate operators at a given particle rank, while still retaining most of the information from the initial higher-body contributions. In our applications, we will approximate all many-body operators by their normal-ordered counterpart truncated at the two-body level, which is commonly referred to as NO2B approximation.

Since we are focussing on closed-shell systems the reference state $|\Psi_{\mathrm{ref}}\rangle$ that defines the normal-ordering is given by a single Slater determinant $|\Phi_{\mathrm{ref}}\rangle$. Hence, we are working in a particle-hole picture where the states occupied (unoccupied) in $|\Phi_{\mathrm{ref}}\rangle$ resemble holes (particles). A product of creation and annihilation operators $\hat{X}_1 \dots \hat{X}_n$ is then considered to be in normal order if its expectation value w.r.t. $|\Phi_{\mathrm{ref}}\rangle$ vanishes

$$\langle \Phi_{\mathrm{ref}} | \{ \hat{X}_1 \dots \hat{X}_n \} | \Phi_{\mathrm{ref}} \rangle = 0, \tag{12.1}$$

where the normal ordering is denoted with curly brackets. This property in itself is already very useful in many applications, however, in general we are interested in expressing a product of operators in terms of normal-ordered products. This is accomplished via Wick's theorem, which states that any product of operators is equal to its normal-ordered product plus all possible normal-ordered contractions

$$\hat{X}_1 \dots \hat{X}_n = \{\hat{X}_1 \dots \hat{X}_n\} + \sum_{\substack{\text{all}\\ \text{contractions}}} \{\hat{X}_1 \dots \overbrace{\hat{X}_i \dots \hat{X}_j} \dots \hat{X}_n\}, \tag{12.2}$$

where a contraction is defined as

$$\{\hat{X}_1 \dots \overbrace{\hat{X}_i \dots \hat{X}_j} \dots \hat{X}_n\} = \mathrm{sgn}(\pi) \overbrace{\hat{X}_i \hat{X}_j} \{\hat{X}_1 \dots \hat{X}_{i-1} \hat{X}_{i+1} \dots \hat{X}_{j-1} \hat{X}_{j+1} \dots \hat{X}_n\}. \tag{12.3}$$

Here, π refers to the permutation required to move $\hat{X}_i$ and $\hat{X}_j$ next to each other without changing the order of the other operators. We can further evaluate such a contraction, which generally is a complex number, using the product of two operators and applying Wick's theorem, which yields

$$\varrho_{ij} \equiv \langle \Phi_{\mathrm{ref}} | \hat{X}_i \hat{X}_j | \Phi_{\mathrm{ref}} \rangle = \langle \Phi_{\mathrm{ref}} | \left(\{\hat{X}_i \hat{X}_j\} + \overbrace{\hat{X}_i \hat{X}_j} \right) | \Phi_{\mathrm{ref}} \rangle = \overbrace{\hat{X}_i \hat{X}_j}, \tag{12.4}$$

where we have introduced ϱ_{ij} as a shorthand notation as it resembles a one-body density w.r.t. the reference state.

Moreover, Wick's theorem can be generalized to provide insights in products of normal-ordered operator products. It then holds

$$\begin{aligned} \{\hat{X}_1 \dots \hat{X}_n\}\{\hat{Y}_1 \dots \hat{Y}_m\} =& \{\hat{X}_1 \dots \hat{X}_n \hat{Y}_1 \dots \hat{Y}_m\} \\ &+ \sum_{\substack{\text{external}\\ \text{contractions}}} \{\hat{X}_1 \dots \overbrace{\hat{X}_i \dots \hat{X}_n \hat{Y}_1 \dots \hat{Y}_j} \dots \hat{Y}_m\}, \end{aligned} \tag{12.5}$$

where external contractions only include contractions across the two normal-ordered operator products.

One of the most common applications of normal ordering is the aforementioned NO2B approximation, which we have already employed in the context of the natural orbitals in Ch. 11. We will, therefore, discuss the normal ordering of a Hamiltonian with up to three-body contributions as an example. When written in second quantization such a Hamiltonian contains products of creation and annihilation operators up to the 3p3h level (see Eq. (11.12) for comparison), which need to be expressed in terms of their normal-ordered products. Applying Wick's theorem we find

$$\hat{a}^p_q = \{\hat{a}^p_q\} + \varrho^p_q, \tag{12.6}$$

$$\hat{a}^{pq}_{rs} = \{\hat{a}^{pq}_{rs}\} + \varrho^p_r \{\hat{a}^q_s\} - \varrho^p_s \{\hat{a}^q_r\} + \varrho^q_r \{\hat{a}^p_s\} - \varrho^q_s \{\hat{a}^p_r\} + \varrho^p_r \varrho^q_s - \varrho^p_s \varrho^q_r, \tag{12.7}$$

$$\hat{a}^{pqr}_{stu} = \{\hat{a}^{pqr}_{stu}\} + \varrho^p_s \{\hat{a}^{qr}_{tu}\} \pm \mathcal{O}(\varrho) + \varrho^p_s \varrho^q_t \{\hat{a}^r_u\} \pm \mathcal{O}(\varrho^2) + \varrho^p_s \varrho^q_t \varrho^r_u \pm \mathcal{O}(\varrho^3), \tag{12.8}$$

such that the normal-ordered Hamiltonian takes the shape

$$\hat{H}_{\mathrm{NO}} = \tilde{H}^{[0]} + \sum_{pq} \tilde{H}^{[1]}_{p,q}\{\hat{a}^p_q\} + \sum_{pqrs} \tilde{H}^{[2]}_{pq,rs}\{\hat{a}^{pq}_{rs}\} + \sum_{pqrstu} \tilde{H}^{[3]}_{pqr,stu}\{\hat{a}^{pqr}_{stu}\} \tag{12.9}$$

with

$$\tilde{H}^{[0]} = H^{[0]} + \sum_{pq} H^{[1]}_{p,q}\varrho^p_q + \frac{1}{2}\sum_{pqrs} H^{[2]}_{pq,rs}\varrho^p_r\varrho^q_s + \frac{1}{6}\sum_{pqrstu} H^{[3]}_{pqr,stu}\varrho^p_s\varrho^q_t\varrho^r_u, \tag{12.10}$$

$$\tilde{H}^{[1]}_{p,q} = H^{[1]}_{p,q} + \sum_{rs} H^{[2]}_{pr,qs}\varrho^r_s + \frac{1}{2}\sum_{rstu} H^{[3]}_{prs,qtu}\varrho^r_s\varrho^t_u, \tag{12.11}$$

$$\tilde{H}^{[2]}_{pq,rs} = H^{[2]}_{pq,rs} + \sum_{tu} H^{[3]}_{pqt,rsu}\varrho^t_u, \tag{12.12}$$

$$\tilde{H}^{[3]}_{pqr,stu} = H^{[3]}_{pqr,stu}. \tag{12.13}$$

Evidently, at a given n-body level, information from all higher-body contributions enters, while the highest rank remains unchanged. We can employ the aforementioned NO2B approximation by dropping the explicit three-body terms in Eq. (12.9). The NO2B approximated Hamiltonian in its original form, i.e., normal ordered w.r.t. the vacuum state $|0\rangle$ is then obtained via the inverse relations for the normal-ordered operators and one arrives at

$$\begin{aligned}\hat{H}_{\mathrm{NO2B}} = &\left(H^{[0]} + \frac{1}{6}\sum_{pqrstu} H^{[3]}_{pqr,stu}\varrho^p_s\varrho^q_t\varrho^r_u\right) && \text{(zero-body)}\\ &+ \sum_{pq}\left(H^{[1]}_{p,q} - \frac{1}{2}\sum_{rstu} H^{[3]}_{prs,qtu}\varrho^r_t\varrho^s_u\right)\hat{a}^p_q && \text{(one-body)}\\ &+ \frac{1}{4}\sum_{pqrs}\left(H^{[2]}_{pq,rs} + \sum_{ct} H^{[3]}_{pqt,rsu}\varrho^t_u\right)\hat{a}^{pq}_{rs}, && \text{(two-body)}\end{aligned} \tag{12.14}$$

where residual three-body terms enter at every level.

12.2 In-Medium Similarity Renormalization Group

The IM-SRG [29] is at the heart of the IM-NCSM. In general, it is constructed to drive the Hamiltonian towards a block-diagonal structure w.r.t. a given many-body basis. This is achieved by partitioning the Hamiltonian into diagonal and off-diagonal parts

$$\hat{H} = \hat{H}^{\mathrm{d}} + \hat{H}^{\mathrm{od}} \tag{12.15}$$

with the goal of suppressing the latter via a unitary transformation. The ansatz for this transformation is the same as for the free-space SRG discussed in Sec. 1.2. Yet, we will recall some of the relations here for completeness and to point out some differences in the notation.

Starting point for our discussion of the IM-SRG is a continuous unitary transformation of the Hamiltonian

$$\hat{H}(s) = \hat{U}^\dagger(s)\hat{H}(0)\hat{U}(s), \quad \hat{H}(0) \equiv \hat{H} \tag{12.16}$$

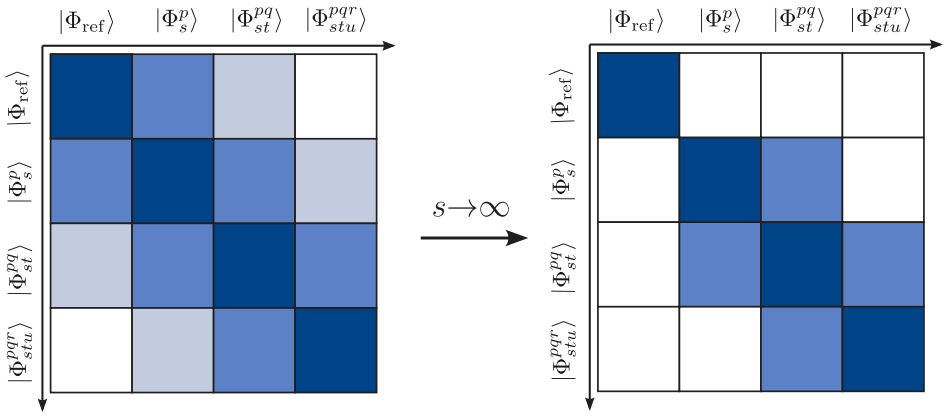

Figure 12.1: Schematic representation of the Hamiltonian before and after the single-reference IM-SRG. For $s \to \infty$ the reference state $|\Phi_{\mathrm{ref}}\rangle$ is decoupled from its particle-hole excitations $|\Phi_{s\ldots}^{p\ldots}\rangle$.

controlled by the flow parameter s. Alternatively, this can be written as a flow equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{H}(s) = \left[\hat{\eta}(s), \hat{H}(s)\right], \tag{12.17}$$

with generator $\hat{\eta}(s)$. Note that this is purely an operator equation which holds the great advantage that, unlike in the NCSM, the many-body basis never needs to be constructed explicitly. Yet, a matrix-element representation is required in order to solve it numerically. Further, any other operators $\hat{O}$ can be transformed analogously

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{O}(s) = \left[\hat{\eta}(s), \hat{O}(s)\right], \quad \hat{O}(0) \equiv \hat{O}. \tag{12.18}$$

The key aspect of the IM-SRG is that it is formulated "in medium", i.e., with regard to a reference state, which in our case is a single Slater determinant in natural orbital basis $|\Phi_{\mathrm{ref}}\rangle$. Therefore, all operators are normal ordered w.r.t. this reference state. For our applications we have chosen the imaginary-time generator [29]. The transformation can, thus, be interpreted as a decoupling of the reference state from its particle-hole excitations as illustrated in Fig. 12.1. This is the main difference to the free-space SRG, where the decoupling is performed for matrix elements in momentum space and not on the level of many-body states.

The computational complexity of the IM-SRG arises during the iterative solution of the flow equation and scales rapidly with higher particle ranks [1]. This is controlled by imposing a truncation on the particle rank and neglecting all higher-body orders. In our case, we employ what is known as IM-SRG(2), hence, we use the NO2B truncation for all operators, either initial or induced during the transformation. With this the right side of the flow equation Eq. (12.17) reduces to the evaluation of commutators between normal-ordered products of creation and annihilation operators which is given by the generalized Wick's theorem Eq. (12.5). Note that, in analogy to the induced many body forces in free-space SRG, these commutators also contain many-body terms beyond the two-body level. In order to stay within the IM-SRG(2) scheme, this requires intermediate truncations to the two-body level, adding to the overall truncation error. Finally, we end up with a system of coupled ordinary differential equations for the matrix elements of the Hamiltonian, which can be solved numerically.

[1] First implementations of the IM-SRG with a complete treatment of three-body contributions are on the fore-front of current developments [185].

Magnus Formalism. So far, we have assumed a general continuous unitary transformation represented by the operator $\hat{U}(s)$ and have concluded that the flow equation has to be solved separately for each observable of interest. Clearly, this is undesirable as it increases the computational effort. We circumvent this by making use of the Magnus formalism [186, 187], where the unitary transformation operator is written as the exponential of an anti-Hermitian operator $\hat{\Omega}(s)$ called Magnus operator

$$\hat{U}(s) = \exp\Big(-\hat{\Omega}(s)\Big), \quad \Omega(0) \equiv 0. \tag{12.19}$$

It can be shown [187] that this operator obeys the following differential equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{\Omega}(s) = \sum_{k=0}^{\infty} \frac{B_k}{k!}\Big[\hat{\Omega}(s), \hat{\eta}(s)\Big]_k, \tag{12.20}$$

where B_k are Bernoulli numbers and $[\ldots\,,\ldots]_k$ denotes k nested commutators, which can be defined recursively via

$$\Big[\hat{X}, \hat{Y}\Big]_k = \Big[\hat{X}, \Big[\hat{X}, \hat{Y}\Big]\Big]_{k-1}, \quad \Big[\hat{X}, \hat{Y}\Big]_0 = \hat{Y}. \tag{12.21}$$

In order to maintain computability, we, again, apply the NO2B truncation to the Magnus operator.

For the computation of $\hat{\Omega}$ we exploit the recursive nature of the nested commutators, where the k-th commutator only depends on the $(k-1)$-th

$$\Big[\hat{\Omega}(s), \hat{\eta}(s)\Big]_k = \Big[\hat{\Omega}(s), \Big[\hat{\Omega}(s), \hat{\eta}(s)\Big]_{k-1}\Big], \tag{12.22}$$

such that the right-hand side of Eq. (12.20) can be evaluated order by order until a convergence criterion

$$\frac{||\hat{\Omega}'_n(s) - \hat{\Omega}'_{n-1}(s)||}{||\hat{\Omega}'_n(s)||} < \epsilon_{\text{Magnus}}\,, \quad \hat{\Omega}'_n(s) := \sum_{k=0}^{n} \frac{B_k}{k!}\Big[\hat{\Omega}(s), \hat{\eta}(s)\Big]_k \tag{12.23}$$

is matched. The big advantage of this formalism is that once $\hat{\Omega}(s)$ is known, any other observable can be transformed via a Baker-Campbell-Hausdorff (BCH) series

$$\hat{O}(s) = \sum_{k=0}^{\infty} \frac{1}{k!}\Big[\hat{\Omega}(s), \hat{O}(0)\Big]_k, \tag{12.24}$$

which directly provides the full transformation up to flow parameter s. Again, the evaluation is carried out until it matches the convergence criterion

$$\frac{||\frac{1}{n!}[\hat{\Omega}_n(s), \hat{O}(0)]_n||}{||\sum_{k=0}^{n} \frac{1}{k!}[\hat{\Omega}_n(s), \hat{O}(0)]_k||} < \epsilon_{\text{BCH}}. \tag{12.25}$$

The BCH series will become essential, when we discuss extensions to hypernuclei in Ch. 13.

Commutator Evaluation. The computationally most demanding part is the evaluation of commutators. This has been worked out in much detail in various other works (see e.g. [29, 30]). We will, therefore, not provide the full derivation but only present the resulting expressions.

Within the IM-SRG(2) a commutator $\hat{C}$ of two scalar operators $\hat{A}$ and $\hat{B}$ has the form

$$\hat{C} = [\hat{A}, \hat{B}] \approx C^{[0]} + \sum_{ar} C^{[1]}_{pq}\{\hat{a}^p_q\} + \frac{1}{4}\sum_{pqrs} C^{[2]}_{pq,rs}\{\hat{a}^{pq}_{rs}\}, \tag{12.26}$$

where the last step resembles an approximation due to the truncation to the NO2B level. With occupation numbers

$$n_p = \langle \Psi_{\text{ref}}|\hat{a}^\dagger_p \hat{a}_p|\Psi_{\text{ref}}\rangle = \begin{cases} 1 & \text{if } |p\rangle \text{ is occupied} \\ 0 & \text{if } |p\rangle \text{ is unoccupied} \end{cases}, \tag{12.27}$$

$$\bar{n}_p = 1 - n_p \tag{12.28}$$

the expressions for the many-body matrix elements according to [29] are given by

$$\begin{aligned} C^{[0]} = &\sum_{pq} A^{[1]}_{p,q} B^{[1]}_{q,p}(n_p - n_q) \\ &+ \frac{1}{4}\sum_{pqrs}(A^{[2]}_{pq,rs}B^{[2]}_{rs,pq} - B^{[2]}_{pq,rs}A^{[2]}_{rs,pq})n_p n_q \bar{n}_r \bar{n}_s, \end{aligned} \tag{12.29}$$

$$\begin{aligned} C^{[1]}_{p,q} = \hat{\Xi}^{[1]}_{pq}\Bigg[&2\sum_r A^{[1]}_{p,r}B^{[1]}_{r,q} \\ &+ \sum_{rs}\left(A^{[1]}_{r,s}B^{[2]}_{pr,qs} - B^{[1]}_{r,s}A^{[2]}_{ps,qr}\right)(n_r - n_s) \\ &+ \sum_{rst} A^{[2]}_{pr,st}B^{[2]}_{st,qr}(n_r\bar{n}_s\bar{n}_t + \bar{n}_r n_s n_t)\Bigg], \end{aligned} \tag{12.30}$$

$$\begin{aligned} C^{[2]}_{pq,rs} = \hat{\Xi}^{[2]}_{pqrs}\Bigg[&4\sum_t\left(A^{[1]}_{p,t}B^{[2]}_{tq,rs} - B^{[1]}_{p,t}A^{[2]}_{tq,rs}\right) \\ &+ \sum_{tu}\left(A^{[2]}_{pq,tu}B^{[2]}_{tu,rs}(1 - n_t - n_u) + 4\,A^{[2]}_{pt,ru}B^{[2]}_{qu,st}(n_t - n_u)\right)\Bigg], \end{aligned} \tag{12.31}$$

where $\hat{\Xi}^{[1]}_{pq}$ and $\hat{\Xi}^{[2]}_{pqrs}$ are index symmetrizers that read

$$\hat{\Xi}^{[1]}_{pq} = \frac{1}{2}(1 \pm P_{pq}), \tag{12.32}$$

$$\hat{\Xi}^{[2]}_{pqrs} = \frac{1}{8}(1 - P_{pq})(1 - P_{rs})(1 \pm P_{pr}P_{qs}). \tag{12.33}$$

Here, "+" is used for a Hermitian and "−" for an anti-Hermitian $\hat{C}$. If $\hat{A}$ and $\hat{B}$ are both hermitian or anti-hermitian $\hat{C}$ is anti-hermitian, otherwise $\hat{C}$ is hermitian. In an actual implementation one resorts to an angular-momentum-coupled versions of these expressions as they allow for a more efficient computation. These spherical results can be found in, e.g., [183, 184].

12.3 In-Medium No-Core Shell Model

Throughout the Chs. 2, 11, and 12 we have discussed all ingredients that enter an IM-NCSM calculation and can now combine them into a hybrid method. The IM-NCSM is a four step process (five steps in the multi-reference framework) as schematically depicted in Fig. 12.2.

The starting point is the construction of an optimized reference state, which is a first approximation for the state of the system we are interested in. For single-reference applications the HF Slater determinant provides such a reference state that can further be optimized via natural orbitals. As we have seen, this holds the additional advantage of being independent of the underlying HO oscillator frequency. For the construction of the reference state we employ a nuclear Hamiltonian containing realistic NN and 3N interactions.

In a next step all operators need to be normal ordered w.r.t. the reference state. In preparation for the subsequent IM-SRG calculation they are truncated at the two-body level and higher orders are discarded.

We then perform an IM-SRG(2) calculation in Magnus formalism, during which the reference state is decoupled from its particle-hole excitations. We end up with a transformed Hamiltonian that is optimized through the inclusion of many-body correlations. All other operators are transformed consistently via the BCH series.

As a final step, a subsequent NCSM calculation is performed with the evolved Hamiltonian. For a fully decoupled reference state ($s \to \infty$) the NCSM calculation will yield converged results for the ground-state in $N_{\max} = 0$ already. The IM-SRG can, therefore, be understood as a convergence accelerator for the subsequent NCSM calculation, enabling the computation of nuclei that are beyond the reach of the standard (IT-)NCSM. The subsequent NCSM calculation further gives access to excited states that are structurally similar to the ground state, thus, overcoming the limitation of the IM-SRG to a single state.

For completeness, we note that the IM-NCSM can be extended to open-shell nuclei [32]. The equal-filling approximation allows for the construction of natural orbitals for open-shell systems and a meaningful reference state $|\Psi_{\text{ref}}\rangle$, resembled by a superposition of Slater determinants, is obtained through a NCSM calculation in a small, e.g., $N_{\max} = 0, 2$ model space. In a multi-reference IM-SRG calculation the whole subspace spanned by the configurations in $|\Psi_{\text{ref}}\rangle$ is being decoupled from its particle-hole excitations. A subsequent NCSM calculation in this subspace provides the ground-state energy along with the excitation energies for structurally similar excited states. However, the IM-SRG is typically restricted to a scalar flow equation, which transfers to a limitation of the IM-NCSM to even-A nuclei.

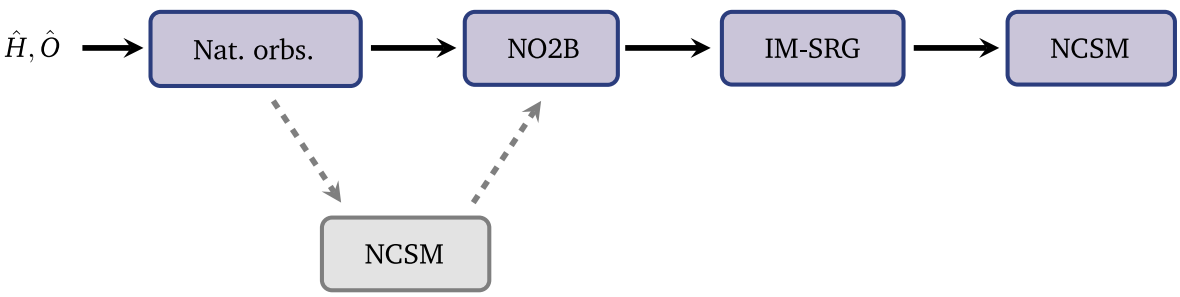

Figure 12.2: Schematic representation of the multi-step process in the IM-NCSM. The intermediate NCSM calculation depicted in gray only applies to the multi-reference case. See text for details.

12.4 Particle-Attached/-Removed Extension

The particle-attached/particle-removed extension introduced in [183, 184] is designed to tackle the limitation to closed-shell nuclei imposed by the single-reference IM-SRG. It is based on the assumption that the reference state constructed for a closed-shell nucleus provides a reasonably good approximation of a structurally similar reference state for its isotopic or isobaric neighbor, which we will refer to as target nucleus. In particular, the idea is to construct the reference state and perform the IM-SRG calculation for the closed-shell nucleus, while the subsequent NCSM calculation is carried out for the open-shell target nucleus. Hence, the decoupling of the closed-shell reference state is assumed to yield a partial decoupling of the respective state in the target system, thus, accelerating the convergence in the subsequent NCSM calculation, which also accounts for the missing correlations.

However, there is one more adjustment required. The NCSM calculation is performed for the target nucleus for which

$$A_{\text{target}} = \begin{cases} A_{\text{closed-shell}} + 1 & \text{(particle-attached)} \\ A_{\text{closed-shell}} - 1 & \text{(particle-removed)} \end{cases}. \tag{12.34}$$

Hence, all inital operators entering the IM-NCSM calculation should be constructed w.r.t. A_{target}. This affects the kinetic energy given in Eq. (1.7), which, with emphasis on the target nucleus, reads

$$\hat{T}_{\text{int}} = \frac{2}{A_{\text{target}}} \sum_{i<j}^{A} \frac{\hat{q}_{ij}^2}{m_{\text{N}}}. \tag{12.35}$$

As a final remark we point out that the particle-attached/-removed approach can also lift the limitation to even-A nuclei in the multi-reference framework. Considering results for multiple different neighbors then allows for a consistency check and potentially an uncertainty estimation of this extension (see e.g. Fig. 6.14 in [184]).

13

Hypernuclei in the IM-NCSM

To date, medium-mass and heavier systems with strangeness have only been computed within phenomenological approaches [188] along with some first applications of HF-MBPT and Brueckner-Hartree-Fock [42]. In the nucleonic sector, however, developments of medium-mass methods have made impressive progress in both reach and precision of ab initio calculations. Clearly, our understanding of hypernuclear structure and the underlying baryonic interaction would greatly benefit from these methodological advances, if these ab initio methods could be transferred to the hypernuclear domain.

In the previous chapter we have discussed the IM-NCSM as a powerful ab initio tool for the description of medium-mass nuclei. Since we already have an implementation of the hypernuclear NCSM at hand, it provides the ideal starting point for developments towards ab initio calculations of hypernuclei in the medium-mass sector. A direct expansion of the IM-SRG to hypernuclei is challenging as, both, the analytical expressions for the commutator evaluation as well as their implementation, require major revisions. In addition, singly-strange hypernuclei naturally are open-shell systems, hence, they require a multi-reference treatment, which further complicates the implementation.

Considering that the binding of the hyperon to the other nucleons is governed by a relatively weak YN interaction, it seems reasonable to start exploring the hypernuclear medium-mass sector in a simplified manner, where we consider a purely nucleonic parent nucleus, to which we introduce a hyperon. Right away, this is akin to the particle-attached concept we have discussed in Sec. 12.4.

In the following, we implement this in form of a hyperon-attached formalism as a first step towards a hypernuclear IM-NCSM.

13.1 Hyperon-Attached IM-NCSM

As we have already mentioned, the hyperon-attached formalism is a straight-forward extension of the particle-attached IM-NCSM concept. A schematic representation of the different steps is shown in Fig. 13.1.

We start out by constructing the nuclear Hamiltonian for the nucleonic parent of the hypernucleus we are interested in. While the initial nucleonic Hamiltonian in the particle-attached formalism had to be adapted to the target nucleus, this is not required here, since the

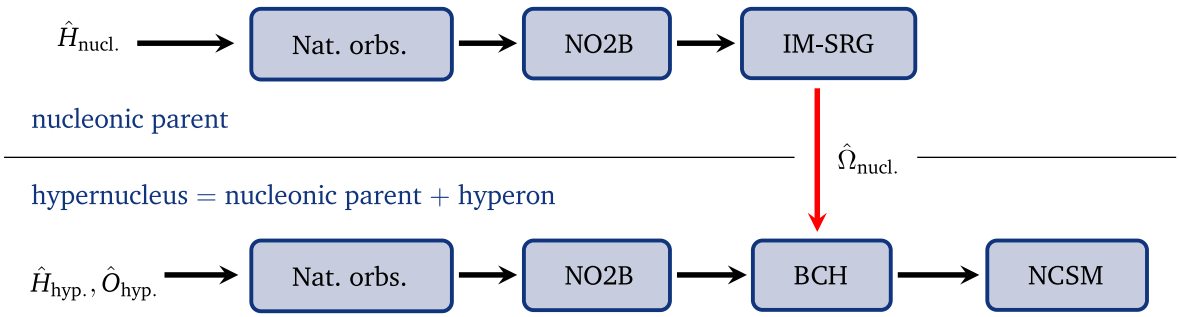

Figure 13.1: Schematic representation of the single-reference hyperon-attached IM-NCSM. The hypernucleus and its nucleonic parent are initially treated separately before the information from the nucleonic parent enters the calculation of the hypernucleus in form of the Magnus operator. See text for details.

nucleonic Hamiltonian will never enter the final NCSM calculation. Still, one might expect a better approximation of the target system, if the initial nucleonic Hamiltonian already carries the proper kinetic energy. This, however, is not as straight-forward as in the nucleonic case due to the hyperon mass being larger than the nucleon mass. Alternatively Eq. (1.5) can be approximated by choosing the total mass operator to evaluate to $A_{\text{target}} - 1$ times the nucleon mass m_{N} plus the mass of a Λ hyperon as the Λ-Σ mixing is comparably small. With this, one can write

$$\hat{T}_{\text{int}} \approx \frac{4}{A_{\text{target}} + \frac{m_\Lambda - m_{\text{N}}}{m_{\text{N}}}} \sum_{i<j}^{A} \frac{\hat{q}_{ij}^2}{2m_{\text{N}}}, \tag{13.1}$$

where we have used the assumption of equal nucleon masses.

In a next step the nuclear Hamiltonian is converted to natural orbitals and normal ordered before it enters an IM-SRG calculation. Since the latter is performed in the Magnus formalism we obtain a Magnus operator $\hat{\Omega}_{\text{nucl.}}(s)$ that carries the information about the decoupled state in the nucleonic parent. Note that this allows us to perform a standard IM-SRG calculation for non-strange nuclei, which renders further changes of the implementation obsolete.

Key to the hyperon-attached formalism is now to embed $\hat{\Omega}_{\text{nucl.}}(s)$ in the hypernuclear Hilbert space of the target hypernucleus and use the BCH series to transform the hypernuclear Hamiltonian $\hat{H}_{\text{hyp.}}$ and any other operators for hypernuclear observables $\hat{O}_{\text{hyp.}}$ accordingly, such that

$$\hat{O}_{\text{hyp.}}(s) = \sum_{k=0}^{\infty} \frac{1}{k!} \left[\hat{\Omega}_{\text{nucl.}}(s), \hat{O}_{\text{hyp.}}(0) \right]_k. \tag{13.2}$$

This can be understood as an inclusion of many-body correlations into the nucleonic components of $\hat{H}_{\text{hyp.}}$, while the hyperonic components remain unchanged. For this embedding we complement the Magnus operator by a unity operator in the hyperonic parts. The transformed operators then enter a hypernuclear NCSM calculation that accounts for the missing YN correlations. Note that these NCSM application apply the same center-of-mass treatment as discussed for the NAT basis with $\lambda_{\text{cm}} = 0.3$.

Before we look at applications to medium-mass hypernuclei, which the hyperon-attached IM-NCSM is meant for, we will first discuss results for ${}^{5}_{\Lambda}\text{He}$ as it allows for a comparison of

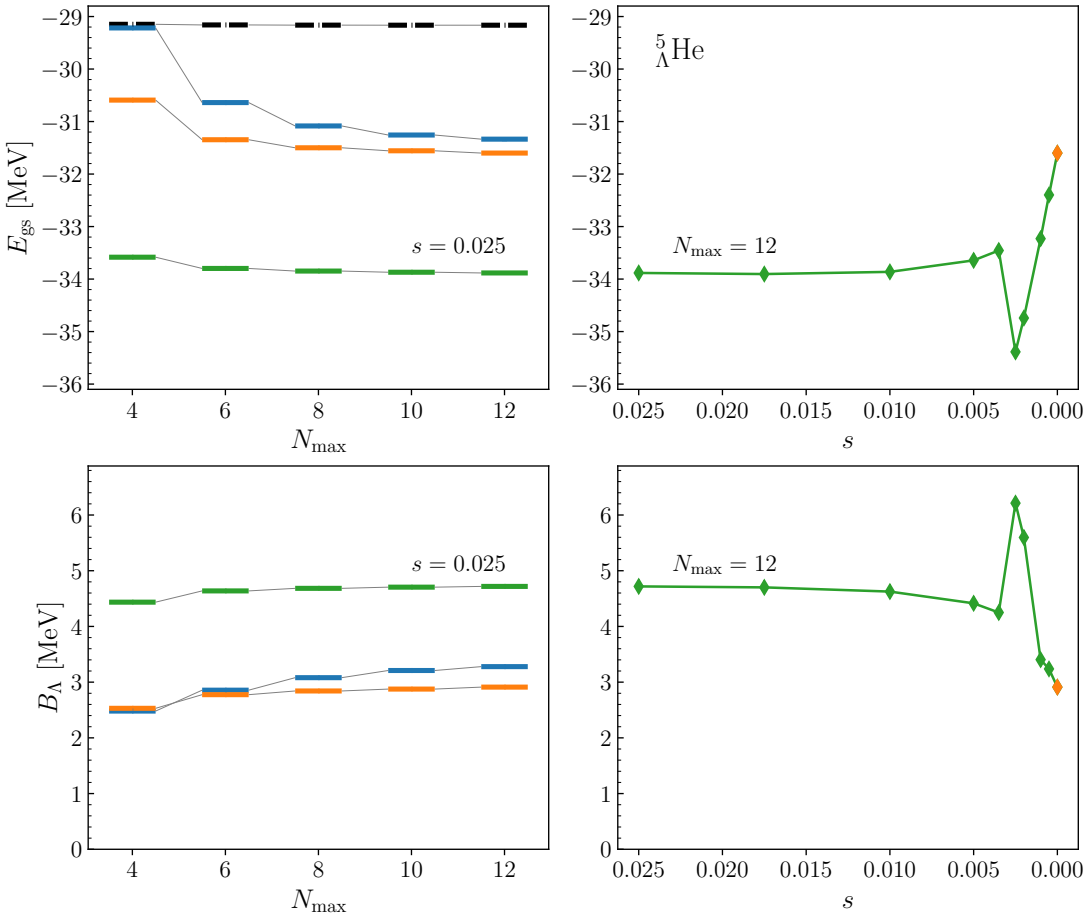

Figure 13.2: Ground-state energies (upper panels) and hyperon separation energies (lower panels) of $^5_\Lambda$He in HO basis with $\hbar\Omega = 20$ MeV including full three-body forces (—), in NAT basis with NO2B approximation (—), and from the hyperon-attached IM-NCSM (—). IM-NCSM results for the nucleonic parent ^{4}He (– –) are given for comparison. The left-hand panels show the convergence with N_{max}, while the right-hand panels illustrate the s dependence for the IM-NCSM calculations.

different methods and calculations can be converged w.r.t. N_{max}. Figure 13.2 shows such NCSM calculations of the ground-state energy and hyperon separation energy of $^5_\Lambda$He for the same Hamiltonian in HO basis, in NAT basis with NO2B truncation, and IM-SRG transformed (left-hand panels). As anticipated, we find that the IM-NCSM calculations converge even faster than for the HO or NAT bases. However, we find that the converged results do not agree with each other. These discrepancies arise from the different approximations that have been introduced. The difference between the HO and NAT calculations are, as discussed previously, a result of the NO2B approximation used for the NAT calculations, while the HO calculations are performed with full NN+3N interactions. The cause for the larger deviations for the IM-NCSM result is less evident. Studying results for different flow parameters s as shown in the right-hand panels of Fig. 13.2 this deviation seems to increase continuously with increasing s except for some non-monotonous behavior for very small flow parameters. The occurring peak is most likely an artifact of the employed generator and since it is remedied for larger s it is of little concern for our applications. For the IM-SRG(2) it is known that the repetitive NO2B approximation accumulates to an error of about 1-2% [189, 190]. This, however, does not fully translate to the Magnus formalism we are employing here, as the exponentiation

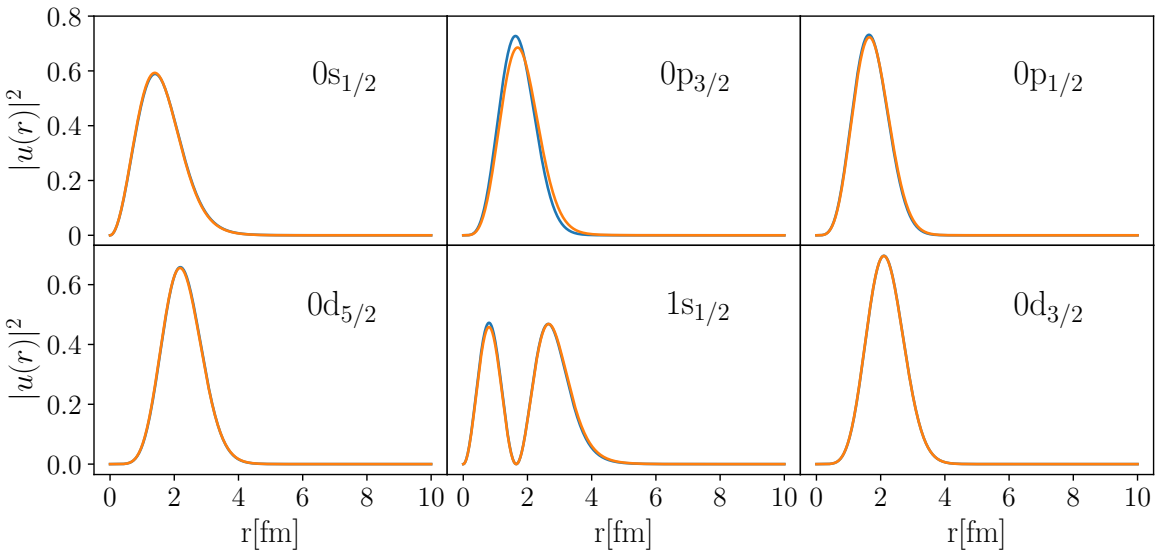

Figure 13.3: NAT squared radial wavefunctions for the neutron in ^{4}He(—) and $^5_\Lambda$He (—).

ensures unitarity of the transformation independent of the quality of the Magnus operator. But, the hypernuclear Hamiltonian and the subsequent transformation of via the BCH series are also truncated at an NO2B level. It is, therefore, likely that this accounts for the major part of the deviation.

Another potential source of error again lies within the choice of $|\Phi_{\text{ref}}\rangle$ since the Magnus operator and the hypernuclear Hamiltonian are normal ordered w.r.t. different reference states. In particular, the Magnus operator is normal ordered w.r.t. a NAT Slater determinant for ^{4}He, while the reference state for the hypernuclear Hamiltonian is a NAT Slater determinant for $^5_\Lambda$He. As demonstrated in Fig. 13.3, the NAT wavefunctions for the nucleons in ^{4}He and $^5_\Lambda$He are nearly identical across all orbitals with only minor deviation in the $0\text{p}_{3/2}$ orbital. Thus, in good approximation, the hypernuclear reference Slater determinant $|\Phi^{\text{hyp}}_{\text{ref}}\rangle$ and the reference state of the nucleonic parent $|\Phi^{\text{nuc}}_{\text{ref}}\rangle$ relate like

$$|\Phi^{\text{nuc}}_{\text{ref}}\rangle \approx \hat{c}_\Lambda |\Phi^{\text{hyp}}_{\text{ref}}\rangle, \tag{13.3}$$

where $\hat{c}_\Lambda$ is the annihilation operator corresponding to the Λ single-particle state in $|\Phi^{\text{hyp}}_{\text{ref}}\rangle$. Since the Magnus operator is already normal ordered w.r.t. $|\Phi^{\text{nuc}}_{\text{ref}}\rangle$ we can write

$$\begin{aligned}
\langle\Phi^{\text{hyp}}_{\text{ref}}|\hat{\Omega}(s)|\Phi^{\text{hyp}}_{\text{ref}}\rangle &= \langle\Phi^{\text{hyp}}_{\text{ref}}|\hat{c}^\dagger_\Lambda \hat{c}_\Lambda \hat{\Omega}(s)|\Phi^{\text{hyp}}_{\text{ref}}\rangle \\
&= \langle\Phi^{\text{hyp}}_{\text{ref}}|\hat{c}^\dagger_\Lambda \hat{\Omega}(s) \hat{c}_\Lambda|\Phi^{\text{hyp}}_{\text{ref}}\rangle \\
&\approx \langle\Phi^{\text{nuc}}_{\text{ref}}|\hat{\Omega}(s)|\Phi^{\text{nuc}}_{\text{ref}}\rangle \\
&= 0,
\end{aligned} \tag{13.4}$$

from which we can conclude that the Magnus operator is also normal ordered w.r.t. the hypernuclear reference state. Hence, we expect the resulting error to be negligible.

13.2 Application to Medium-Mass Hypernuclei

Finally, we investigate calculations of hypernuclei that are out of range for the NCSM. Let us first discuss ${}^{17}_{\Lambda}\mathrm{O}$, which is still light enough so that we can provide NCSM calculations with natural orbitals as comparison. Both, the results for the hyperon-attached IM-NCSM and the NCSM calculations for the ground-state energy and hyperon separation energy for ${}^{17}_{\Lambda}\mathrm{O}$ are shown in Fig. 13.4. For the ground-state energy we find a significantly increased convergence rate for the IM-NCSM calculation, which converges to a precision better than 1% at $N_{\max} = 8$. Turning to the hyperon separation energy we observe an overbinding of the hyperon in the IM-NCSM relative to the NCSM analogously to our findings for ${}^{5}_{\Lambda}\mathrm{He}$. Moreover, the convergence of B_Λ is significantly slower than for the NCSM calculations due to the rapid convergence of the nucleonic parent as indicated by the black dashed lines. For, both, ${}^{4}\mathrm{He}$ and ${}^{16}\mathrm{O}$ the IM-NCSM calculations are essentially converged at $N_{\max} = 2$, thus, the convergence rate of B_Λ in the corresponding hypernuclei is the same as for the ground-state energy. Therefore, quantifying the overbinding of the hyperon in ${}^{17}_{\Lambda}\mathrm{O}$ is difficult but we estimate it to be of similar size as for ${}^{5}_{\Lambda}\mathrm{He}$.

With the hyperon-attached IM-NCSM we are able to access even larger hypernuclei providing the first ab initio results for ${}^{41}_{\Lambda}\mathrm{Ca}$, which are presented in Fig. 13.5. A comparison with IT-NCSM calculations is hardly possible as calculations above $N_{\max} = 4$ require excessive amounts of computational resources or unphysical IT truncations. Nevertheless, we find that the ground-state energy converges even faster than for ${}^{17}_{\Lambda}\mathrm{O}$, yielding nearly converged results at $N_{\max} = 4$ already. Moreover, the convergence rate for the hyperon-separation energy is comparable to the NCSM calculation with natural orbitals, yet it still suffers from the rapid convergence of the nucleonic parent and seems to exhibit a similar overbinding as we have found for ${}^{5}_{\Lambda}\mathrm{He}$ and ${}^{17}_{\Lambda}\mathrm{O}$.

Overall, we find that the hyperon-attached IM-NCSM increases the reach of the NCSM

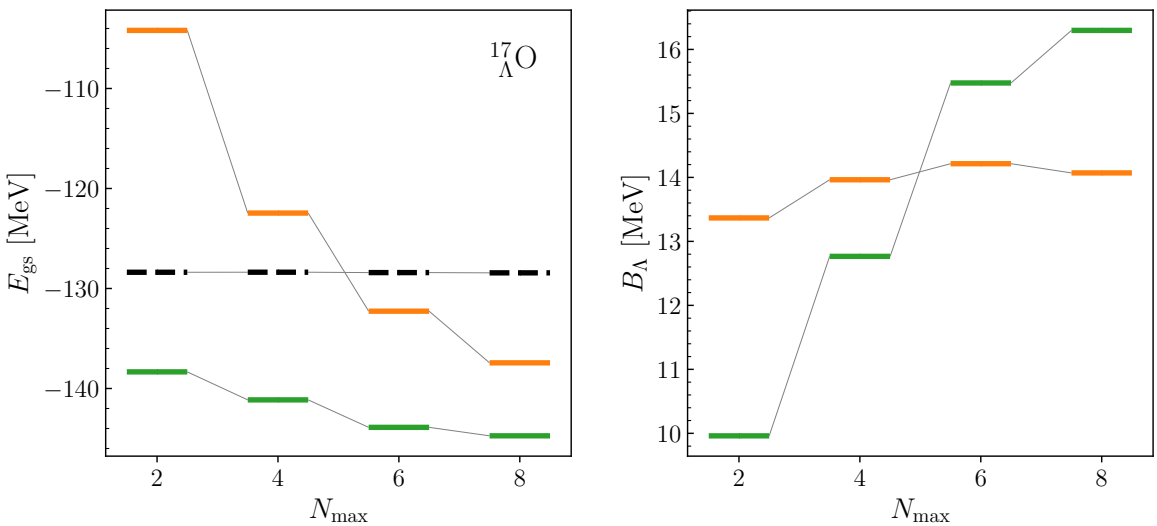

Figure 13.4: Ground-state energies (left panel) and hyperon separation energies (right panel) of ${}^{17}_{\Lambda}\mathrm{O}$ in NAT basis with NO2B approximation (—) and from the hyperon-attached IM-NCSM (—). IM-NCSM results for the nucleonic parent ${}^{16}\mathrm{O}$ (– –) are given for comparison.

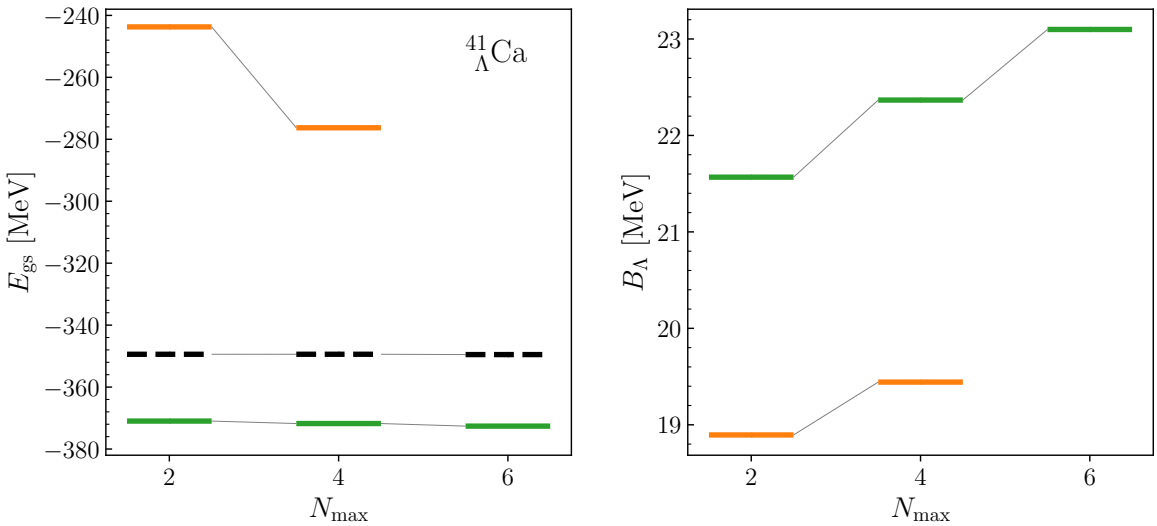

Figure 13.5: Same as Fig. 13.4 but for $^{41}_{\Lambda}$Ca and ^{40}Ca.

into the medium-mass regime and provides a great starting point for extensions to open-shell hypernuclei and further development of hypernuclear in-medium methods.

On a final note, we emphasize that with the current single-reference implementation we are only able to study hypernuclei that exhibit a closed-shell nucleonic parent. As we have shown, these tend to show a rapid convergence in the IM-NCSM. Open-shell nuclei, on the other hand, exhibit a slower convergence, which would benefit the convergence rate of the hyperon separation energy as it is a difference-based observable. However, a quantitative investigation of this is not possible until the multi-reference extension for the hyperon-attached IM-NCSM is implemented.

Summary and Outlook

In this work we have presented progress on different frontiers of hypernuclear structure theory. Building on recent extensions of the NCSM to hypernuclei together with realistic YN interactions from chiral EFT we were able to identify and address three important challenges for ab initio hypernuclear theory: the poorly constrained YN interaction, precision calculations of p-shell hypernuclei with many-body uncertainty estimation, and the push to ab initio calculations for medium-mass hypernuclei.

Precision with ANN Extrapolations

We have started out by developing a machine-learning tool supplemental to NCSM calculations, for both nuclei and hypernuclei, with the goal of overcoming the limitation to finite model spaces and the lack of realistic uncertainty estimates on the many-body level. Inspired by previous work along these lines, we have constructed ANNs that are capable of predicting observables in the infinite Hilbert space based on sequences of NCSM calculations in accessible model spaces. By learning the observable-specific convergence patterns from calculations in few-body systems without being tied to a single nucleus, interaction, or state, we were able to convert this extrapolation problem into an interpolation problem, for which ML methods are known to excel. We also showed that the convergence patterns are universal across a wide range of nuclei. It is this universality that allows the direct application of the ANNs to hypernuclei, which is remarkable considering that the networks have never been informed with hypernuclear data. Moreover, through a statistical evaluation of multiple ANNs and systematic data sampling we can extract meaningful many-body uncertainties that are crucial for precision calculations of (hyper)nuclear properties.

We have demonstrated that these ANNs can reliably predict ground-state energies and mass rms-radii for various p-shell (hyper)nuclei, exceeding the capabilities of classical extrapolation schemes. Through an extension to difference-based observables the ANNs for ground-state energies can also be used to precisely predict excitation energies as well as hyperon separation energies, again building on the universality of the networks. Clearly, this ANN tool is a powerful asset that extends the reach and precision of the NCSM and its design allows for a straight-forward adaption to other observables or even many-body methods, thus, providing a great foundation for future developments of extrapolation schemes and many-body uncertainty estimates. Work on extensions to electromagnetic observables is already in progress.

YN Interaction Constrained on p-Shell Hypernuclei

With this extended NCSM framework at hand, we have turned to the YN interaction, which is poorly constrained by the few experimental hypernuclear scattering data available. After identifying the most relevant degrees of freedom in the LO interaction we constructed an

optimized YN interaction using ground-state spectroscopic data for p-shell hypernuclei as additional constraints for the LECs. We have shown that these adjustments alleviate the previously found systematic overbinding of the hyperon in p-shell hypernuclei, while maintaining a good description of the available scattering data. Especially for light hypernuclei we have shown exceptional agreement with experimental data. Yet, we have also observed that the hyperon separation energies show an increasing dependence on the nucleonic interactions for larger particle numbers A, which needs to be taken into account for future optimizations of YN interactions.

Considering that these improvements have been possible with a LO interaction, this demonstrates the great potential of ground-state and spectroscopic data as additional constraints on hypernuclear interactions. However, the restriction to a LO interaction comes with its limitations and some remaining tensions cannot be resolved by the identified degrees of freedom. This and the lack of uncertainty estimations for a LO interaction necessitate the extension of this optimization scheme to NLO and N^{2}LO, where an analogous optimization procedure is possible, which would perfectly complement other efforts on the uncertainty estimation for YN interactions [95, 166]. Due to the increasing number of LECs in higher chiral orders, this becomes high-dimensional optimization problem. Hence, emulators based on eigenvector continuation can be used to reduce the computational effort [191, 192].

Natural Orbitals – Optimized Basis and Diagnostic Tool

After addressing both, the deficiencies of the YN interaction and the limitations of the NCSM as a many-body method for p-shell hypernuclei, we have shifted our focus to extending the reach to heavier systems. In a first step we have adapted the concept of natural orbitals to the hypernuclear sector and have been able to show that its outstanding properties, i.e., the independence on the underlying HO frequency and the increased convergence rate can directly be transferred to hypernuclei. The improved convergence rate extends the reach of the NCSM.

Moreover, we have found that the radial wavefunctions of the natural orbitals themselves can provide great insight into the structure of hypernuclei. By looking at their radial extend we have been able to identify ${}^{5}_{\Lambda}\mathrm{He}$ as a promising candidate for a hyperon halo, which has previously been unnoticed. This finding is supported by a long tail in the Λ density distribution, a Λ rms-radius that exceeds the nucleon radius by at least 30%, and a strong separation of scales for the hyperon and nucleon separation energies. Hence, natural orbital wavefunctions are a valuable diagnostic tool especially when it comes to different particle species.

IM-NCSM for Medium-Mass Hypernuclei

In addition to its applications and benefits within the NCSM, natural orbitals as an optimized single-particle basis mark an ideal starting point for in-medium methods as they allow for the construction of optimized reference states. This brings us to the final part of this work, which is centered around the ab initio description of medium-mass hypernuclei. Connecting to the great advances of IM-SRG based methods in the nucleonic regime, we have developed a hyperon-attached IM-NCSM that allows for a partially separate treatment of the nucleonic parent and the additional hyperon. Akin to previously developed particle-attached/particle-removed approaches, we have shown that it is possible to incorporate all nucleonic correlations in an IM-SRG calculation and include the hyperon and the missing

hyperon-nucleon correlations in a subsequent NCSM calculation. This has led to a greatly improved convergence rate for ground-state energies of hypernuclei allowing us to report the first ab initio calculation for ${}^{41}_{\Lambda}\mathrm{Ca}$, which marks a major step towards an ab initio framework for medium-mass hypernuclei.

As for all novel methods there remain some discrepancies that require further investigation and improvement. For now, applications of the hyperon-attached IM-NCSM are limited to hypernuclei with closed-shell nucleonic parents, however, extensions to open-shell systems are conceptually straight-forward, since they have been worked out in detail in the nucleonic regime.

What's next?

In the near future, hypernuclear structure theory will follow its nucleonic counterpart into a precision era. Accessing additional sources of hypernuclear data such as data for p-shell hypernuclei will remain a vital aspect of further developments of realistic YN interactions, especially when incorporating higher chiral orders. The work presented here has shown the great potential that lies within this additional data. Furthermore, ML applications like our ANN tool resemble a major step towards precision calculations as many-body uncertainties for not just ground-state energies but also hyperon separation energies, excitation energies and radii are now well under control and can be quantified. As a next step, a full uncertainty quantification including, both, many-body and interaction uncertainties should be the focus of future work. However, with the achievements of this work calculations of p-shell hypernuclei are already precise enough to provide guidance for experimental efforts that are planned for the near future.

Beyond precision calculations of p-shell hypernuclei, ab initio calculations have entered the regime of medium-mass hypernuclei. Our developments on hypernuclear natural orbitals and the hyperon-attached IM-NCSM provide a great foundation for future developments of IM methods with strangeness. A natural next step is the extension to a multi-reference framework which will then give access to hypernuclei with open-shell nucleonic parents. From there, one will eventually want to incorporate the strangeness degree of freedom into the flow equations of the IM-SRG to gain access to the full medium-mass regime of the hypernuclear chart.

Finally, systematic extensions of ab initio methods to multi-strange systems and hypernuclear matter are on the horizon. These, however, require major efforts on both, the many-body and interaction side, where the latter further requires sufficient experimental data in the hyperon-hyperon sector, which will be difficult to obtain for another while.

Appendices

A

Numerical Values for ANN Predictions

A.1 Ground-State Energies and Spectra

E_{gs} [MeV]	$\mathcal{N}_{\max}$ 8	10	12	$\mathcal{N}_{\max}$ 8	10	12
		^{2}H			^{3}H	
ABS	−2.164(69)	−2.225(49)	−2.194(28)	−8.518(70)	−8.539(47)	−8.504(29)
DIFF	−2.218(32)	−2.226(27)	−2.199(13)	−8.498(45)	−8.500(22)	−8.488(11)
MINMAX	−2.195(47)	−2.223(27)	−2.195(14)	−8.482(65)	−8.509(33)	−8.487(12)
class. extrap.	−2.013(106)	−2.147(*)	−2.183(106)	−8.434(110)	−8.480(109)	−8.473(24)
conv. result	−2.200	−2.200	−2.200	−8.481	−8.481	−8.481
		^{4}He			^{6}Li	
ABS	−28.604(75)	−28.580(56)	−28.558(45)	−31.649(271)	−31.822(158)	−31.994(100)
DIFF	−28.527(26)	−28.526(13)	−28.527(7)	−31.625(126)	−31.824(84)	−32.011(45)
MINMAX	−28.513(71)	−28.526(25)	−28.526(8)	−31.951(198)	−32.026(122)	−32.143(96)
class. extrap.	−28.503(48)	−28.528(24)	−28.523(9)	−31.718(271)	−31.906(189)	−32.008(362)
conv. result	−28.524	−28.524	−28.524	–	–	–
		^{7}Li			^{8}Li	
MINMAX	−39.460(192)	−39.392(106)	−39.397(58)	−41.348(206)	−41.358(137)	–
class. extrap.	−39.118(336)	−39.236(133)	−39.392(68)	−41.138(474)	−41.227(156)	–
		^{9}Be			^{12}C	
ABS	−58.210(308)	−58.498(207)	–	−98.038(245)	−97.693(156)	–
DIFF	−58.371(256)	−58.624(108)	–	−98.139(229)	−97.757(62)	–
MINMAX	−58.911(259)	−58.951(130)	–	−98.987(313)	−97.917(130)	–
class. extrap.	−58.743(517)	−58.815(209)	–	−97.78(116)	−97.639(578)	–

Table A.1: Numerical values for ground-state energies for various p-shell nuclei obtained with the ANN tool corresponding to Figs. 6.2, 6.3, and 6.6. Classical extrapolations according to Sec. 3.2 and the converged values, where available, are given for comparison. Uncertainties denoted as (*) mark unreasonably large values due to a breakdown of the extrapolation method.

E_{ex} [MeV]	$\mathcal{N}_{\max}$ 8	10	12	$\mathcal{N}_{\max}$ 8	10	12
		^{6}Li			class. extrap.	
3^+	2.435(115)	2.421(50)	2.397(36)	2.381(354)	2.414(189)	2.342(362)
0^+	3.910(99)	3.863(62)	3.786(58)	3.918(346)	3.758(189)	3.652(362)
2^+	4.741(145)	4.562(96)	4.458(68)	4.673(593)	4.384(189)	4.435(473)
		^{7}Li			class. extrap.	
$\frac{1}{2}^-$	0.367(52)	0.332(29)	0.323(14)	0.344(336)	0.334(152)	0.323(68)
$\frac{5}{2}^-$	5.013(129)	4.927(47)	4.881(32)	4.864(462)	4.910(133)	4.959(106)
$\frac{5}{2}^-$	6.947(133)	6.807(101)	6.708(69)	6.933(554)	6.805(245)	6.650(118)
$\frac{7}{2}^-$	8.135(28)	7.951(59)	7.879(32)	8.046(525)	7.938(225)	7.843(108)
		^{8}Li			class. extrap.	
1^+	1.215(91)	1.102(69)	–	1.175(504)	1.107(194)	–
3^+	2.530(75)	2.544(4)	–	2.511(522)	2.538(156)	–
4^+	7.051(173)	6.794(152)	–	6.856(879)	6.796(242)	–

Table A.2: Numerical values for excitation energies for the lowest-lying natural parity states in $^{6}_{\Lambda}$Li, $^{7}_{\Lambda}$Li, and $^{8}_{\Lambda}$Li obtained with the ANN tool corresponding to Fig. 6.6. Classical extrapolations according to Sec. 3.2 and the converged values, where available, are given for comparison.

A.2 Radii

R_{rms} [MeV]	$\mathcal{N}_{\max}$ 8	10	12	$\mathcal{N}_{\max}$ 8	10	12
		^{2}H			^{3}H	
ABS	1.935(31)	1.953(38)	1.931(32)	1.739(47)	1.754(19)	1.732(17)
DIFF	2.128(45)	2.104(35)	2.075(38)	1.794(31)	1.766(27)	1.733(15)
MINMAX	1.995(46)	1.966(23)	1.979(25)	1.726(20)	1.712(14)	1.703(08)
conv. result	1.979	1.979	1.979	1.696	1.696	1.696
		^{4}He				
ABS	1.440(25)	1.462(13)	1.462(9)			
DIFF	1.492(15)	1.472(10)	1.466(5)			
MINMAX	1.461(10)	1.461(4)	1.460(2)			
conv. result	1.460	1.460	1.460			
		^{6}Li			^{7}Li	
MINMAX	2.235(18)	2.284(22)	2.291(18)	2.285(15)	2.307(15)	2.325(14)
		^{8}Li			^{9}Be	
MINMAX	2.308(16)	2.327(17)	–	2.308(16)	2.327(17)	–

Table A.3: Numerical values for mass rms-radii for various p-shell nuclei obtained with the ANN tool corresponding to Figs. 7.1 and 7.2. Converged values, where available, are given for comparison.

B

ANN Predictions for Hypernuclear Observables with $\mathrm{YN}_{\mathrm{opt}}$

This appendix provides supplemental material to Ch. 10 consisting of figures that give insight in the quality of the employed ANN predictions and a collection of numerical values for the results shown in Figs. 10.8 to 10.10. Figures of the ANN predictions are only shown for the $\mathrm{NN}_{\mathrm{EMN}} + 3\mathrm{N}_{\mathrm{H}} + \mathrm{YN}_{\mathrm{opt}}$ interaction.

Note that an assessment of the predictions becomes more difficult with increasing particle number A due to the limited $N_{\max}$ reach. Therefore, the main criteria for a good, plausible prediction are the consistency of predictions over multiple $\mathcal{N}_{\max}$ and whether a human practitioner would make a similar estimation.

B.1 Predictions for Hyperon Separation Energies

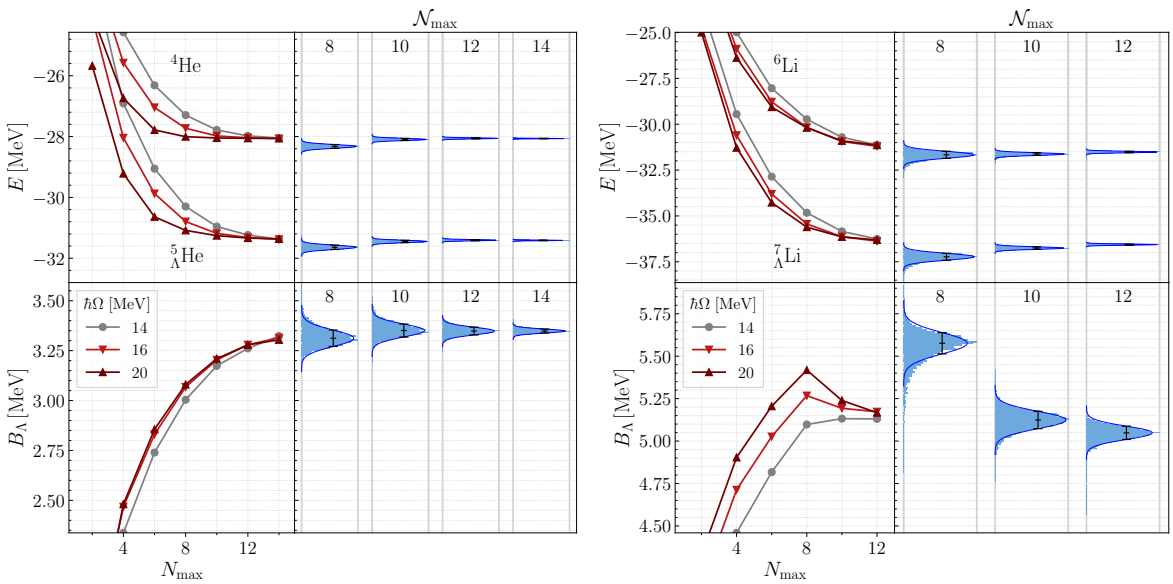

Figure B.1: Evaluation data along with predictions of ground-state energies (upper panels) and hyperon separation energies (lower panels) of the two showcase hypernuclei $^5_\Lambda\mathrm{He}$ and $^7_\Lambda\mathrm{Li}$ obtained with the same 1000 ANNs in MINMAX mode as discussed in Ch. 6.

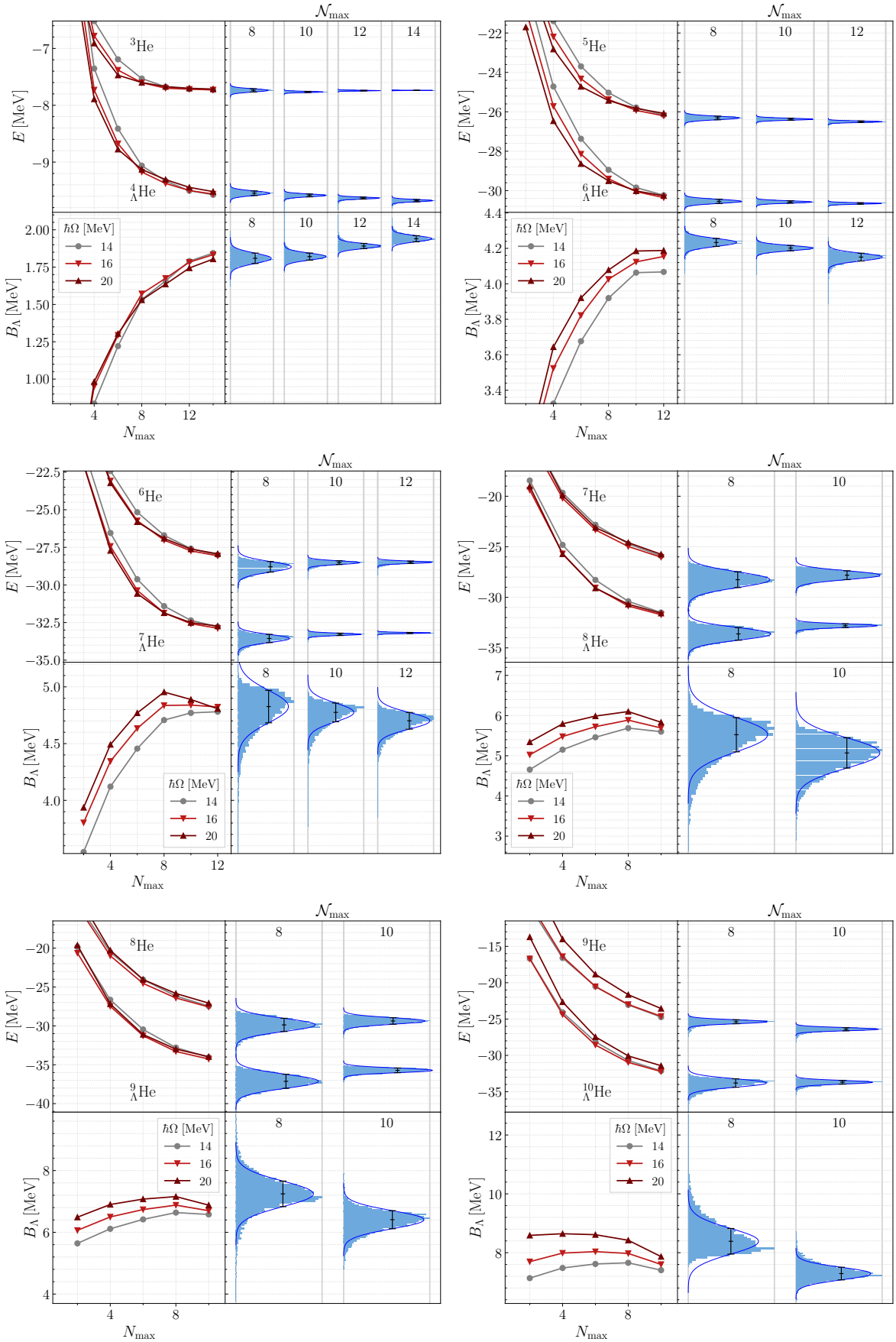

Figure B.2: Same as Fig. B.1 but for $^4_\Lambda$He, $^6_\Lambda$He, $^7_\Lambda$He, $^8_\Lambda$He, $^9_\Lambda$He, and $^{10}_\Lambda$He.

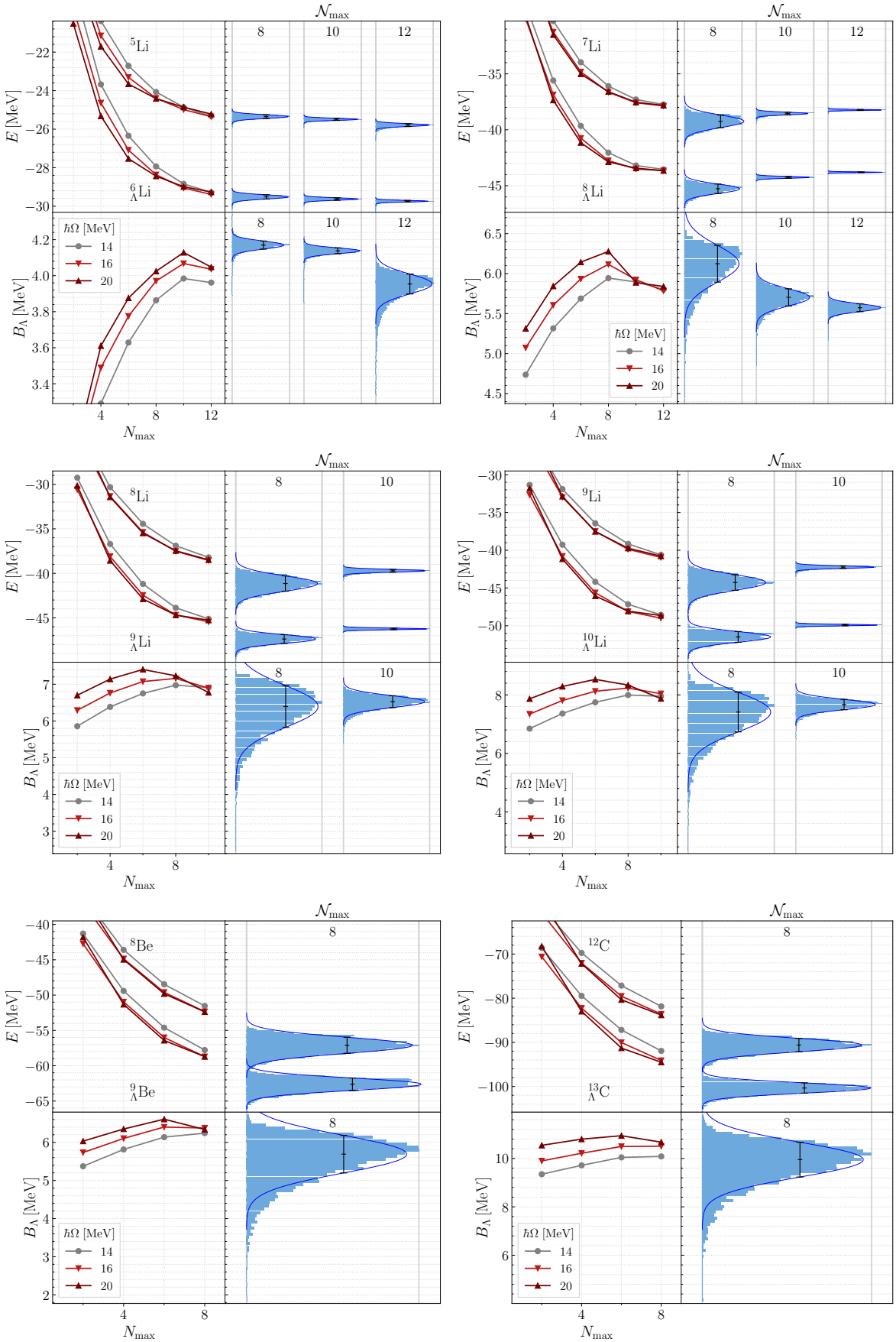

Figure B.3: Same as Fig. B.1 but for $^{6}_{\Lambda}$Li, $^{8}_{\Lambda}$Li, $^{9}_{\Lambda}$Li, $^{10}_{\Lambda}$Li, $^{9}_{\Lambda}$Be, and $^{13}_{\Lambda}$C.

B.2 Numerical Values

Obs	$NN_{EM} + 3N_N$	$NN_{EMN} + 3N_H$	Exp	$NN_{EM} + 3N_N$	$NN_{EMN} + 3N_H$		Exp
				YN_P	YN_P	YN_{opt}	
	^{3}He			$^{4}_{\Lambda}$He			
E_{gs}	−7.72	−7.74(1)	−7.72	−10.15(6)	−10.38(1)	−9.68(2)	−10.11(3)
$E_{ex}(0^+)$	–	–	–	2.31	1.38(2)	1.21(2)	1.15
B_Λ	–	–	–	2.43(6)	2.64(1)	1.94(2)	2.39(3)
	^{4}He			$^{5}_{\Lambda}$He			
E_{gs}	−28.49	−28.07(1)	−28.30	−32.91(1)	−32.79(1)	−31.41(1)	−31.42(2)
B_Λ	–	–	–	4.42(4)	4.73(1)	3.35(1)	3.12(2)
	^{5}He			$^{6}_{\Lambda}$He			
E_{gs}	−26.83(8)	−26.50(4)	−27.56(1)	−32.2(1)	−32.01(3)	−30.65(4)	−31.74(10)
$E_{ex}(\frac{1}{2}^-\|2^-)$	1.96(23)	1.23(19)	1.27	0.15	0.15(1)	0.17(1)	–
$E_{ex}(0^-)$	–	–	–	2.41(17)	1.82(21)	1.72(21)	–
$E_{ex}(1^-)$	–	–	–	2.70(18)	1.90(26)	1.90(21)	–
B_Λ	–	–	–	5.43(4)	5.51(3)	4.15(2)	4.18(10)
	^{5}Li			$^{6}_{\Lambda}$Li			
E_{gs}	–	−25.79(8)	−26.330	–	−31.12(4)	−29.74(4)	−30.83(50)
$E_{ex}(\frac{1}{2}^-\|2^-)$	–	0.95(21)	–	–	0.13(1)	0.14(1)	–
$E_{ex}(\frac{3}{2}^-\|0^-)$	–	5.01(37)	–	–	1.49(25)	1.44(23)	–
$E_{ex}(\frac{1}{2}^-\|1^-)$	–	7.04(36)	–	–	1.61(26)	1.58(23)	–
B_Λ	–	–	–	–	5.32(7)	3.95(6)	4.50(50)
	^{6}He			$^{7}_{\Lambda}$He			
E_{gs}	−28.96(13)	−28.49(8)	−29.27(5)	−35.33(6)	−34.59(4)	−33.19(4)	−34.95(25)
$E_{ex}(2^+\|\frac{3}{2}^+)$	2.21(26)	2.03(12)	1.80(1)	2.20(7)	2.20(2)	2.20(2)	–
$E_{ex}(2^+\|\frac{5}{2}^+)$	–	4.36(20)	–	2.28(7)	2.27(2)	2.27(2)	–
$E_{ex}(1^+\|\frac{3}{2}^+)$	–	4.74(20)	–	5.63(15)	5.05(15)	4.88(18)	–
B_Λ	–	–	–	6.24(11)	6.13(8)	4.70(7)	5.68(25)
	^{6}Li			$^{7}_{\Lambda}$Li			
E_{gs}	−32.36(4)	−31.52(4)	−31.99	−39.25(4)	−37.97(3)	−36.56(3)	−37.57(3)
$E_{ex}(3^+\|\frac{3}{2}^+)$	2.60(4)	2.72(1)	2.14	0.59(1)	0.57(1)	0.58(1)	0.69
$E_{ex}(0^+\|\frac{5}{2}^+)$	4.12(3)	3.87(9)	3.51	2.45(5)	2.57(1)	2.63(1)	2.05
$E_{ex}(2^+\|\frac{7}{2}^+)$	–	4.25(16)	4.26	2.93(4)	3.03(1)	3.09(1)	2.52
B_Λ	–	–	–	6.89(6)	6.51(4)	5.05(4)	5.58(3)
	^{7}He			$^{8}_{\Lambda}$He			
E_{gs}	−27.97(11)	−27.81(44)	−28.86(1)	−35.43(1)	−34.35(12)	−32.82(20)	−36.02(70)
$E_{ex}(\frac{1}{2}^-\|1^-)$	2.19(26)	2.40(44)	–	0.02	0.02(1)	0.02(2)	–
$E_{ex}(\frac{5}{2}^-\|0^-)$	–	3.69(43)	2.92(9)	2.85(14)	1.74(39)	1.96(30)	–
$E_{ex}(\frac{3}{2}^-\|1^-)$	–	4.17(47)	–	3.29(16)	2.35(37)	2.53(27)	–
B_Λ	–	–	–	7.46(10)	6.69(41)	5.07(38)	7.16(70)

Table B.1: Numerical values for ground-state energies, excitation energies, and hyperon separation energies for various p-shell hypernuclei and their nucleonic parents corresponding to Figs. 10.8 to 10.10. All values are given in MeV and errors are many-body errors obtained with the ANN extrapolation tool. Experimental values are taken from [7, 11, 146, 162–165, 167] and correspond to the black markers for the hyperon separation energies in the aforementioned figures.

Obs	$NN_{EM}+3N_N$	$NN_{EMN}+3N_H$	Exp	$NN_{EM}+3N_N$ / YN_P	$NN_{EMN}+3N_H$ / YN_P	$NN_{EMN}+3N_H$ / YN_{opt}	Exp
	^{7}Li			$^{8}_{\Lambda}$Li			
E_{gs}	$-39.61(24)$	$-38.23(5)$	-39.245	$-47.27(25)$	$-45.46(4)$	$-43.80(4)$	$-46.02(5)$
$E_{ex}(\frac{1}{2}^-\|0^-)$	$0.31(1)$	$0.07(1)$	0.48	$0.44(1)$	$0.38(1)$	$0.36(1)$	–
$E_{ex}(\frac{7}{2}^-\|2^-)$	$4.93(1)$	$5.08(2)$	$4.63(1)$	$0.59(1)$	$0.46(1)$	$0.46(1)$	–
$E_{ex}(\frac{5}{2}^-\|1^-)$	$7.09(8)$	$6.17(13)$	$6.68(5)$	$0.90(1)$	$0.71(1)$	$0.67(1)$	–
B_Λ	–	–	–	$7.67(16)$	$7.23(5)$	$5.57(5)$	$6.77(5)$
	^{8}He			$^{9}_{\Lambda}$He			
E_{gs}	$-30.23(8)$	$-29.37(42)$	$-31.40(9)$	$-39.01(2)$	$-37.19(18)$	$-35.72(30)$	–
$E_{ex}(2^+\|\frac{3}{2}^+)$	$4.15(26)$	$4.13(43)$	$3.1(5)$	$4.82(16)$	$3.99(30)$	$4.29(27)$	–
$E_{ex}(1^+\|\frac{5}{2}^+)$	–	$5.43(46)$	–	$5.17(19)$	$4.40(27)$	$4.64(29)$	–
B_Λ	–	–	–	$8.78(9)$	$7.89(38)$	$6.41(29)$	–
	^{8}Li			$^{9}_{\Lambda}$Li			
E_{gs}	$-41.72(35)$	$-39.72(16)$	-41.28	$-50.56(35)$	$-48.02(9)$	$-46.25(9)$	$-49.65(7)$
$E_{ex}(1^+\|\frac{5}{2}^+)$	$0.94(3)$	$0.54(17)$	0.98	$0.55(1)$	$0.48(1)$	$0.49(1)$	$0.57(12)$
$E_{ex}(-\|\frac{3}{2}^+)$	–	–	–	$1.40(2)$	$1.07(3)$	$1.05(3)$	$1.49(9)$
$E_{ex}(-\|\frac{1}{2}^+)$	–	–	–	$1.73(3)$	$1.34(4)$	$1.28(5)$	–
B_Λ	–	–	–	$8.29(21)$	$8.29(17)$	$6.53(16)$	$8.37(7)$
	^{8}Be			$^{9}_{\Lambda}$Be			
E_{gs}	$-56.24(29)$	$-57.11(114)$	-56.50	$-64.70(40)$	$-64.00(65)$	$-62.62(88)$	$-63.09(15)$
$E_{ex}(2^+\|\frac{3}{2}^+)$	$3.27(5)$	$3.45(11)$	3.03	$3.10(5)$	$3.35(15)$	$3.42(17)$	$3.02(1)$
$E_{ex}(4^+\|\frac{5}{2}^+)$	–	$11.90(34)$	11.35	$3.30(5)$	$3.44(14)$	$3.49(16)$	$3.07(1)$
$E_{ex}(2^+\|\frac{7}{2}^+)$	–	$16.68(29)$	16.63	–	–	–	–
B_Λ	–	–	–	$8.50(50)$	$7.11(77)$	$5.69(49)$	$6.59(15)$
	^{9}He			$^{10}_{\Lambda}$He			
E_{gs}	$-28.21(22)$	$-26.40(22)$	$-28.91(11)$	$-37.67(12)$	$-35.70(36)$	$-33.70(26)$	–
$E_{ex}(\frac{3}{2}^-\|1^-)$	$3.66(43)$	$2.94(20)$	$1.99(10)$	$0.63(17)$	$0.70(10)$	$0.65(7)$	–
$E_{ex}(-\|1^-)$	–	–	–	$5.39(41)$	$4.87(34)$	$4.28(25)$	–
$E_{ex}(-\|2^-)$	–	–	–	$5.46(39)$	$4.89(35)$	$4.32(25)$	–
B_Λ	–	–	–	$8.78(9)$	$7.89(38)$	$6.41(29)$	–
	^{9}Li			$^{10}_{\Lambda}$Li			
E_{gs}	$-45.45(35)$	$-42.24(18)$	-45.34	$-55.83(45)$	$-51.83(10)$	$-49.91(10)$	–
$E_{ex}(\frac{1}{2}^-\|2^-)$	$1.87(4)$	$0.89(15)$	$2.69(1)$	$0.42(1)$	$0.37(1)$	$0.38(1)$	–
$E_{ex}(-\|0^-)$	–	–	–	$1.84(1)$	$0.91(2)$	$0.88(2)$	–
$E_{ex}(-\|1^-)$	–	–	–	$2.85(1)$	$1.70(4)$	$1.62(5)$	–
B_Λ	–	–	–	$10.37(21)$	$9.59(19)$	$7.67(18)$	–
	^{12}C			$^{13}_{\Lambda}$C			
E_{gs}	$-98.7, 0.80$	$-90.61(152)$	-92.16	$-113.10(80)$	$-102.02(91)$	$-100.34(117)$	$-103.85(12)$
$E_{ex}(2^+\|\frac{3}{2}^+)$	$3.97(5)$	$3.29(18)$	4.44	$3.95(6)$	$3.71(20)$	$3.61(21)$	$4.85(7)$
$E_{ex}(1^+\|\frac{5}{2}^+)$	–	$10.08(127)$	$12.71(1)$	$4.62(6)$	$4.16(24)$	$4.03(26)$	–
$E_{ex}(4^+\|\frac{3}{2}^+)$	–	$11.38(66)$	$13.30(20)$	–	$8.51(78)$	$8.98(101)$	–
B_Λ	–	–	–	$14.50(110)$	$12.08(102)$	$9.95(72)$	$11.69(12)$

Table B.2: Continuation of Tab. B.1.

List of Acronyms

3N three-nucleon.

ANN artificial neural network.

BCH Baker-Campbell-Hausdorff.

BHF Brueckner-Hartree-Fock.

CC coupled-cluster.

CI configuration interaction.

CNN convolutional neural networks.

EC eigenvector continuation.

EFT effective field theory.

EMN Entem-Machleidt-Nosyk.

FCFF fully-connected feed-forward.

FY Faddeev-Yakubovski.

GEM Gaussian expansion methods.

HF Hartree-Fock.

HF-MBPT Hartree-Fock many-body perturbation theory.

HO harmonic oscillator.

HPC high-performance computing.

IM-NCSM in-medium no-core shell model.

IM-SRG in-medium similarity renormalization group.

IR infrared.

IT importance truncation.

IT-NCSM importance-truncated no-core shell model.

LECs low-energy constants.

LO leading order.

MBPT many-body perturbation theory.

ML machine learning.

MSE mean-square error.

N^{2}LO next-to-next-to-leading order.

N^{3}LO next-to-next-to-next-to-leading order.

NAT natural orbitals.

NCSM no-core shell model.

NEAT neuroevolution of augmented topologies.

NLO next-to-leading order.

NN nucleon-nucleon.

NO2B normal-ordered two-body.

QCD quantum chromodynamics.

QMC quantum Monte Carlo.

ReLU rectified linear unit.

rms root-mean-square.

RNN recurrent neural networks.

SCGF self-consistent Green's function.

SGD stochastic gradient descent.

SRG similarity renormalization group.

UV ultraviolet.

YN hyperon-nucleon.

YNN hyperon-nucleon-nucleon.

Bibliography

[1] E. Epelbaum, H.-W. Hammer, and U.-G. Meißner, "Modern theory of nuclear forces", Rev. Mod. Phys. **81**, 1773 (2009).

[2] R. Machleidt and D. Entem, "Chiral effective field theory and nuclear forces", Physics Reports **503**, 1 (2011).

[3] S. D. Głazek and K. G. Wilson, "Renormalization of Hamiltonians", Phys. Rev. D **48**, 5863 (1993).

[4] F. Wegner, "Flow Equations for Hamiltonians", Annalen der Physik (Leipzig) **3**, 77 (1994).

[5] S. K. Bogner, R. J. Furnstahl, and R. J. Perry, "Similarity renormalization group for nucleon-nucleon interactions", Phys. Rev. C **75**, 061001 (2007).

[6] M. Danysz and J. Pniewski, "Delayed disintegration of a heavy nuclear fragment: I", The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **44**, 348 (1953).

[7] D. Davis, "50 years of hypernuclear physics: I. The early experiments", Nucl. Phys. A **754**, Proceedings of the Eighth International Conference on Hypernuclear and Strange Particle Physics, 3 (2005).

[8] H. Takahashi et al., "Observation of a ${}^{6}_{\Lambda\Lambda}\mathrm{He}$ Double Hypernucleus", Phys. Rev. Lett. **87**, 212502 (2001).

[9] P. K. Saha et al., "Production of the Neutron-Rich Hypernucleus ${}^{10}_{\Lambda}\mathrm{Li}$ in the (π^-, K^+) Double Charge-Exchange Reaction", Phys. Rev. Lett. **94**, 052502 (2005).

[10] F. Cusanno et al., "High-Resolution Spectroscopy of ${}^{16}_{\Lambda}\mathbf{N}$ by Electroproduction", Phys. Rev. Lett. **103**, 202501 (2009).

[11] S. N. Nakamura et al., "Observation of the ${}^{7}_{\Lambda}\mathrm{He}$ Hypernucleus by the $(e, e' K^+)$ Reaction", Phys. Rev. Lett. **110**, 012502 (2013).

[12] A. Esser et al., "Observation of ${}^{4}_{\Lambda}\mathrm{H}$ Hyperhydrogen by Decay-Pion Spectroscopy in Electron Scattering", Phys. Rev. Lett. **114**, 232501 (2015).

[13] H. Fujioka et al., "Extension of the J-PARC Hadron Experimental Facility - summary report -", 2017, arXiv:1706.07916 [nucl-ex].

[14] H. H. U. Association et al., "Write-ups for workshop on the project for the hadron experimental facility of J-PARC, Partial collection of LOIs at the extended hadron hall and the related topics", 2019, arXiv:1906.02357 [nucl-ex].

[15] K. Aoki et al., "Extension of the J-PARC Hadron Experimental Facility: Third White Paper", 2021, arXiv:2110.04462 [nucl-ex].

[16] I. Bombaci, "The hyperon puzzle in neutron stars", in Proceedings of the 12th International Conference on Hypernuclear and Strange Particle Physics (HYP2015) (2017), p. 101002.

[17] D. Lonardoni et al., "Hyperon puzzle: hints from quantum Monte Carlo calculations", Phys. Rev. Lett. **114**, 092301 (2015).

[18] A. Gal, E. V. Hungerford, and D. J. Millener, "Strangeness in nuclear physics", Rev. Mod. Phys. **88**, 035004 (2016).

[19] H. Polinder, J. Haidenbauer, and U.-G. Meißner, "Hyperon–nucleon interactions—a chiral effective field theory approach", Nucl. Phys. A **779**, 244 (2006).

[20] J. Haidenbauer et al., "Hyperon–nucleon interaction at next-to-leading order in chiral effective field theory", Nucl. Phys. A **915**, 24 (2013).

[21] J Haidenbauer, U.-G. Meißner, and A. Nogga, "Hyperon–nucleon interaction within chiral effective field theory revisited", Eur. Phys. J. A **56**, 91 (2020).

[22] J. Haidenbauer et al., "Hyperon–nucleon interaction in chiral effective field theory at next-to-next-to-leading order", Eur. Phys. J. A **59**, 63 (2023).

[23] B. R. Barrett, P. Navrátil, and J. P. Vary, "Ab initio no core shell model", Prog. Part. Nucl. Phys. **69**, 131 (2013).

[24] P. Navrátil et al., "Recent developments in no-core shell-model calculations", J. Phys. G **36**, 083101 (2009).

[25] R. Roth, "Importance truncation for large-scale configuration interaction approaches", Phys. Rev. C **79**, 064324 (2009).

[26] D. C. Zheng et al., "Microscopic calculations of the spectra of light nuclei", Physical Review C **48**, 1083 (1993).

[27] K. Kowalski et al., "Coupled Cluster Calculations of Ground and Excited States of Nuclei", Phys. Rev. Lett. **92**, 132501 (2004).

[28] W. Dickhoff and C. Barbieri, "Self-consistent Green's function method for nuclei and nuclear matter", Progress in Particle and Nuclear Physics **52**, 377 (2004).

[29] H. Hergert et al., "The In-Medium Similarity Renormalization Group: A novel ab initio method for nuclei", Physics Reports **621**, Memorial Volume in Honor of Gerald E. Brown, 165 (2016).

[30] H Hergert, "In-medium similarity renormalization group for closed and open-shell nuclei", Physica Scripta **92**, 023002 (2016).

[31] J. Carlson et al., "Quantum Monte Carlo methods for nuclear physics", Rev. Mod. Phys. **87**, 1067 (2015).

[32] E. Gebrerufael et al., "Ab Initio Description of Open-Shell Nuclei: Merging No-Core Shell Model and In-Medium Similarity Renormalization Group", Phys. Rev. Lett. **118**, 152503 (2017).

[33] J. A. Melendez, S. Wesolowski, and R. J. Furnstahl, "Bayesian truncation errors in chiral effective field theory: Nucleon-nucleon observables", Phys. Rev. C **96**, 024003 (2017).

[34] J. A. Melendez et al., "Quantifying correlated truncation errors in effective field theory", Phys. Rev. C **100**, 044001 (2019).

[35] I. Svensson, A. Ekström, and C. Forssén, "Bayesian estimation of the low-energy constants up to fourth order in the nucleon-nucleon sector of chiral effective field theory", Phys. Rev. C **107**, 014001 (2023).

[36] A Nogga, H. Kamada, and W. Gloeckle, "The Hypernuclei ${}^{4}_{\Lambda}\mathrm{He}$ and ${}^{4}_{\Lambda}\mathrm{H}$: Challenges for Modern Hyperon-Nucleon Forces", Phys. Rev. Lett. **88**, 172501 (2002).

[37] A. Nogga, "Light hypernuclei based on chiral interactions at next-to-leading order", Few-body Syst. **55**, 757 (2014).

[38] E. Hiyama and T. Yamada, "Structure of light hypernuclei", Prog. Part. Nucl. Phys. **63**, 339 (2009).

[39] D. Lonardoni, S. Gandolfi, and F. Pederiva, "Effects of the two-body and three-body hyperon-nucleon interactions in Λ hypernuclei", Phys. Rev. C **87**, 041303 (2013).

[40] D. Lonardoni, F. Pederiva, and S. Gandolfi, "Auxiliary Field Diffusion Monte Carlo study of the hyperon–nucleon interaction in Λ-hypernuclei", Nucl. Phys. A **914**, XI International Conference on Hypernuclear and Strange Particle Physics (HYP2012), 243 (2013).

[41] R. Wirth et al., "Ab Initio Description of p-Shell Hypernuclei", Phys. Rev. Lett. **113**, 192502 (2014).

[42] R. Wirth, "Ab-Initio Approach to Hypernuclei" (Technische Universität Darmstadt, 2018).

[43] R. Wirth et al., "Hypernuclear no-core shell model", Phys. Rev. C **97**, 064315 (2018).

[44] H. Le et al., "Jacobi no-core shell model for p-shell hypernuclei", Eur. Phys. J. A **56**, 1 (2020).

[45] D. Frame et al., "Impurity lattice Monte Carlo for hypernuclei", Eur. Phys. J. A **56**, 248 (2020).

[46] E Epelbaum, H Krebs, and U. G. Meißner, "Improved chiral nucleon-nucleon potential up to next-to-next-to-next-to-leading order", The European Physical Journal A **51**, 1 (2015).

[47] R. J. Furnstahl et al., "Quantifying truncation errors in effective field theory", Phys. Rev. C **92**, 024005 (2015).

[48] P. Maris et al., "Uncertainties in ab initio nuclear structure calculations with chiral interactions", Frontiers in Physics **11**, `10.3389/fphy.2023.1098262` (2023).

[49] K. A. Wendt et al., "Infrared length scale and extrapolations for the no-core shell model", Phys. Rev. C **91**, 061301 (2015).

[50] S. Bogner et al., "Convergence in the no-core shell model with low-momentum two-nucleon interactions", Nuclear Physics A **801**, 21 (2008).

[51] P. Maris, J. P. Vary, and A. M. Shirokov, "Ab initio no-core full configuration calculations of light nuclei", Phys. Rev. C **79**, 014308 (2009).

[52] P Maris et al., "Light nuclei with semilocal momentum-space regularized chiral interactions up to third order", Physical Review C **103**, 054001 (2021).

[53] R. J. Furnstahl, S. N. More, and T. Papenbrock, "Systematic expansion for infrared oscillator basis extrapolations", Phys. Rev. C **89**, 044301 (2014).

[54] J. W. Clark, T. Lindenau, and M. L. Ristig, "Scientific Applications of Neural Nets", Scientific Applications of Neural Nets **522**, 10.1007/BFb0104276 (1999).

[55] A. Boehnlein et al., "Colloquium: Machine learning in nuclear physics", Reviews of Modern Physics **94**, 031003 (2022).

[56] G. A. Negoita et al., "Deep learning: Extrapolation tool for ab initio nuclear theory", Phys. Rev. C **99**, 054308 (2019).

[57] W. G. Jiang, G. Hagen, and T. Papenbrock, "Extrapolation of nuclear structure observables with artificial neural networks", Phys. Rev. C **100**, 054326 (2019).

[58] I. Vidaña et al., "Hypernuclear structure with the new Nijmegen potentials", Phys. Rev. C **64**, 044301 (2001).

[59] H. Hergert, "A guided tour of ab initio nuclear many-body theory", Frontiers in Physics **8**, 379 (2020).

[60] K. Hebeler et al., "Normal ordering of three-nucleon interactions for ab initio calculations of heavy nuclei", Phys. Rev. C **107**, 024310 (2023).

[61] P. Demol et al., "Improved many-body expansions from eigenvector continuation", Phys. Rev. C **101**, 041302 (2020).

[62] P. Demol et al., "Bogoliubov many-body perturbation theory under constraint", Annals of Physics **424**, 168358 (2021).

[63] Q. Yuan, J. Li, and H. Li, "Ab initio calculations for well deformed nuclei: 40Mg and 42Si", Physics Letters B **848**, 138331 (2024).

[64] S. Schulz, "Four-Nucleon Forces in Ab Initio Nuclear Structure" (Technische Universität Darmstadt, 2018).

[65] F. Hoyle, "On Nuclear Reactions Occuring in Very Hot STARS. I. the Synthesis of Elements from Carbon to Nickel.", Astrophysical Journal Supplement, vol. 1, p. 121 (1954) **1**, 121 (1954).

[66] E. Epelbaum et al., "Ab initio calculation of the Hoyle state", Physical Review Letters **106**, 192501 (2011).

[67] R. B. Wiringa, V. G. J. Stoks, and R. Schiavilla, "Accurate nucleon-nucleon potential with charge-independence breaking", Phys. Rev. C **51**, 38 (1995).

[68] R. Machleidt, "High-precision, charge-dependent Bonn nucleon-nucleon potential", Phys. Rev. C **63**, 024001 (2001).

[69] A. Shirokov et al., "N3LO NN interaction adjusted to light nuclei in ab exitu approach", Physics Letters B **761**, 87 (2016).

[70] E. Epelbaum, "Nuclear forces from chiral effective field theory: a primer", 2010, arXiv:1001.3229 [nucl-th].

[71] S. Scherer, "A primer for chiral perturbation theory" (Springer, 2012).

[72] S. Weinberg, "Phenomenological Lagrangians", Physica A: Statistical Mechanics and its Applications **96**, 327 (1979).

[73] S. Weinberg, "Nuclear forces from chiral Lagrangians", Physics Letters B **251**, 288 (1990).

[74] S. Weinberg, "Effective chiral Lagrangians for nucleon-pion interactions and nuclear forces", Nuclear Physics B **363**, 3 (1991).

[75] S. Weinberg, "Three-body interactions among nucleons and pions", Physics Letters B **295**, 114 (1992).

[76] K. Hebeler et al., "Improved nuclear matter calculations from chiral low-momentum interactions", Phys. Rev. C **83**, 031301 (2011).

[77] K. Hebeler, "Three-nucleon forces: Implementation and applications to atomic nuclei and dense matter", Physics Reports **890**, Three-nucleon forces: Implementation and applications to atomic nuclei and dense matter, 1 (2021).

[78] T. Hüther et al., "Family of chiral two- plus three-nucleon interactions for accurate nuclear structure studies", Phys. Lett. B **808**, 135651 (2020).

[79] P. Maris et al., "Nuclear properties with semilocal momentum-space regularized chiral interactions beyond $\mathrm{N^2LO}$", Phys. Rev. C **106**, 064002 (2022).

[80] A. Ekström et al., "Δ isobars and nuclear saturation", Phys. Rev. C **97**, 024332 (2018).

[81] W. G. Jiang et al., "Accurate bulk properties of nuclei from $A = 2$ to ∞ from potentials with Δ isobars", Phys. Rev. C **102**, 054301 (2020).

[82] J. Hoppe et al., "Probing chiral interactions up to next-to-next-to-next-to-leading order in medium-mass nuclei", Phys. Rev. C **100**, 024318 (2019).

[83] C. Drischler, K. Hebeler, and A. Schwenk, "Chiral Interactions up to Next-to-Next-to-Next-to-Leading Order and Nuclear Saturation", Phys. Rev. Lett. **122**, 042501 (2019).

[84] V. Somà et al., "Novel chiral Hamiltonian and observables in light and medium-mass nuclei", Phys. Rev. C **101**, 014318 (2020).

[85] D. R. Entem, R. Machleidt, and Y. Nosyk, "High-quality two-nucleon potentials up to fifth order of the chiral expansion", Phys. Rev. C **96**, 024004 (2017).

[86] R. Roth, T. Neff, and H. Feldmeier, "Nuclear structure in the framework of the Unitary Correlation Operator Method", Prog. Part. Nucl. Phys. **65**, 50 (2010).

[87] R. Roth et al., "Similarity-Transformed Chiral $NN + 3N$ Interactions for the Ab Initio Description of ^{12}**C** and ^{16}**O**", Phys. Rev. Lett. **107**, 072501 (2011).

[88] R. J. Furnstahl and K Hebeler, "New applications of renormalization group methods in nuclear physics", Rept. Prog. Phys. **76**, 126301 (2013).

[89] R. Roth et al., "Evolved chiral $NN + 3N$ Hamiltonians for ab initio nuclear structure calculations", Phys. Rev. C **90**, 024325 (2014).

[90] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", United States Governm. Press Office Los Angeles, CA, `10.6028/jres.045.026` (1950).

[91] S Liebig, U.-G. Meißner, and A Nogga, "Jacobi no-core shell model for p-shell nuclei", The European Physical Journal A **52**, 1 (2016).

[92] S. Binder et al., "Few-nucleon systems with state-of-the-art chiral nucleon-nucleon forces", Phys. Rev. C **93**, 044002 (2016).

[93] J. Melendez, "Effective field theory truncation errors and why they matter" (The Ohio State University, 2020).

[94] A. Ekström et al., "Accurate nuclear radii and binding energies from a chiral interaction", Phys. Rev. C **91**, 051301 (2015).

[95] D. Gazda, T. Yadanar Htun, and C. Forssén, "Nuclear physics uncertainties in light hypernuclei", Phys. Rev. C **106**, 054001 (2022).

[96] S. König et al., "Eigenvector continuation as an efficient and accurate emulator for uncertainty quantification", Physics Letters B **810**, 135814 (2020).

[97] T. Djärv et al., "Bayesian predictions for $A = 6$ nuclei using eigenvector continuation emulators", Phys. Rev. C **105**, 014005 (2022).

[98] R. J. Furnstahl, G. Hagen, and T. Papenbrock, "Corrections to nuclear energies and radii in finite oscillator spaces", Phys. Rev. C **86**, 031301 (2012).

[99] S. A. Coon et al., "Convergence properties of ab initio calculations of light nuclei in a harmonic oscillator basis", Phys. Rev. C **86**, 054002 (2012).

[100] S. N. More et al., "Universal properties of infrared oscillator basis extrapolations", Phys. Rev. C **87**, 044326 (2013).

[101] S. Yoshida, "Nonparametric Bayesian approach to extrapolation problems in configuration interaction methods", Physical Review C **102**, 024305 (2020).

[102] C. Adams et al., "Variational Monte Carlo calculations of A < 4 nuclei with an artificial neural-network correlator ansatz", Physical Review Letters **127**, 022502 (2021).

[103] R. Utama, W.-C. Chen, and J. Piekarewicz, "Nuclear charge radii: density functional theory meets Bayesian neural networks", Journal of Physics G: Nuclear and Particle Physics **43**, 114002 (2016).

[104] L. Neufcourt et al., "Bayesian approach to model-based extrapolation of nuclear observables", Phys. Rev. C **98**, 034318 (2018).

[105] A. Ekström et al., "Bayesian optimization in ab initio nuclear physics", Journal of Physics G: Nuclear and Particle Physics **46**, 095101 (2019).

[106] S Athanassopoulos et al., "Nuclear mass systematics using neural networks", Nuclear Physics A **743**, 222 (2004).

[107] S. Akkoyun et al., "An artificial neural network application on nuclear charge radii", Journal of Physics G: Nuclear and Particle Physics **40**, 055106 (2013).

[108] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", Neural Networks **2**, 359 (1989).

[109] C. M. Bishop and N. M. Nasrabadi, "Pattern recognition and machine learning", Vol. 4, 4 (Springer, 2006).

[110] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", nature **521**, 436 (2015).

[111] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning" (MIT press, 2016).

[112] J. Gu et al., "Recent advances in convolutional neural networks", Pattern Recognition **77**, 354 (2018).

[113] Z. Li et al., "A survey of convolutional neural networks: analysis, applications, and prospects", IEEE transactions on neural networks and learning systems **33**, 6999 (2021).

[114] A. Vaswani et al., "Attention is all you need", Advances in neural information processing systems **30** (2017).

[115] K. O. Stanley and R. Miikkulainen, "Evolving Neural Networks through Augmenting Topologies", Evolutionary Computation **10**, 99 (2002).

[116] T. Brown et al., "Language models are few-shot learners", Advances in neural information processing systems **33**, 1877 (2020).

[117] R. Rojas, "The backpropagation algorithm", in *Neural networks* (Springer, 1996), pp. 149–182.

[118] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization.", Journal of machine learning research **12** (2011).

[119] T. Tieleman, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude", COURSERA: Neural networks for machine learning **4**, 26 (2012).

[120] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", 2017, arXiv:1412.6980 [nucl-th].

[121] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond", 2019, arXiv:1904.09237 [nucl-th].

[122] M. Knöll et al., "Machine learning for the prediction of converged energies from ab initio nuclear structure calculations", Physics Letters B **839**, 137781 (2023).

[123] T. Wolfgruber, M. Knöll, and R. Roth, "Precise neural network predictions of energies and radii from the no-core shell model", 2023, arXiv:2310.05256 [nucl-th].

[124] G. Cybenko, "Approximation by superpositions of a sigmoidal function", Mathematics of control, signals and systems **2**, 303 (1989).

[125] K. Xu et al., "How neural networks extrapolate: From feedforward to graph neural networks", 2020, arXiv:2009.11848 [nucl-th].

[126] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines", in Proceedings of the 27th international conference on machine learning (ICML-10) (2010), pp. 807–814.

[127] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization", 2017, arXiv:1711.05101 [nucl-th].

[128] P Reinert, H. Krebs, and E. Epelbaum, "Semilocal momentum-space regularized chiral two-nucleon potentials up to fifth order", The European Physical Journal A **54**, 1 (2018).

[129] P Maris et al., "Light nuclei with semilocal momentum-space regularized chiral interactions up to third order", Physical Review C **103**, 054001 (2021).

[130] S. R. Beane et al., "Hyperon–nucleon scattering from fully-dynamical lattice QCD", Nuclear Physics A **794**, 62 (2007).

[131] H. Nemura et al., "Hyperon–nucleon force from lattice QCD", Physics Letters B **673**, 136 (2009).

[132] H. Nemura et al., "Lambda-Nucleon and Sigma-Nucleon interactions from lattice QCD with physical masses", PoS **LATTICE2016**, 101 (2017).

[133] P. M. M. Maessen, T. A. Rijken, and J. J. de Swart, "Soft-core baryon-baryon one-boson-exchange models. II. Hyperon-nucleon potential", Phys. Rev. C **40**, 2226 (1989).

[134] T. A. Rijken, V. G. J. Stoks, and Y. Yamamoto, "Soft-core hyperon-nucleon potentials", Phys. Rev. C **59**, 21 (1999).

[135] T. A. Rijken and Y. Yamamoto, "Extended-soft-core baryon-baryon model. II. Hyperon-nucleon interaction", Phys. Rev. C **73**, 044008 (2006).

[136] A. Reuber et al., "Correlated $\pi\pi$ and KK exchange in the baryon-baryon interaction", Nuclear Physics A **608**, 243 (1996).

[137] J. Haidenbauer and U.-G. Meißner, "Jülich hyperon-nucleon model revisited", Phys. Rev. C **72**, 044005 (2005).

[138] E. Hiyama et al., "$\Lambda - \Sigma$ conversion in ${}^{4}_{\Lambda}\mathrm{He}$ and ${}^{4}_{\Lambda}\mathrm{H}$ based on a four-body calculation", Phys. Rev. C **65**, 011301 (2001).

[139] H. Nemura, Y. Akaishi, and Y. Suzuki, "Ab initio Approach to s-Shell Hypernuclei ${}^{3}_{\Lambda}\mathrm{H}$, ${}^{4}_{\Lambda}\mathrm{H}$, ${}^{4}_{\Lambda}\mathrm{He}$, and ${}^{5}_{\Lambda}\mathrm{He}$ with a $\Lambda N - \Sigma N$ Interaction", Phys. Rev. Lett. **89**, 142504 (2002).

[140] A. Nogga, H. Kamada, and W. Glöckle, "The Hypernuclei ${}^{4}_{\Lambda}\mathrm{He}$ and ${}^{4}_{\Lambda}\mathrm{H}$: Challenges for Modern Hyperon-Nucleon Forces", Phys. Rev. Lett. **88**, 172501 (2002).

[141] B. F. Gibson et al., "Importance of Baryon-Baryon Coupling in Hypernuclei", Progress of Theoretical Physics Supplement **117**, 339 (1994).

[142] S. Petschauer et al., "Leading three-baryon forces from SU (3) chiral effective field theory", Phys. Rev. C **93**, 014001 (2016).

[143] R. Wirth and R. Roth, "Similarity renormalization group evolution of hypernuclear Hamiltonians", Phys. Rev. C **100**, 044313 (2019).

[144] R. Wirth and R. Roth, "Induced Hyperon-Nucleon-Nucleon Interactions and the Hyperon Puzzle", Phys. Rev. Lett. **117**, 182501 (2016).

[145] R. Wirth and R. Roth, "Light neutron-rich hypernuclei from the importance-truncated no-core shell model", Phys. Lett. B **779**, 336 (2018).

[146] O. Hashimoto and H. Tamura, "Spectroscopy of Λ hypernuclei", Progress in Particle and Nuclear Physics **57**, 564 (2006).

[147] B. Sechi-Zorn et al., "Low-Energy Λ-Proton Elastic Scattering", Phys. Rev. **175**, 1735 (1968).

[148] G. Alexander et al., "Study of the $\Lambda - N$ System in Low-Energy $\Lambda - p$ Elastic Scattering", Phys. Rev. **173**, 1452 (1968).

[149] R. Engelmann et al., "Inelastic Σ-p-interactions at low momenta", Physics Letters **21**, 587 (1966).

[150] F. Eisele et al., "Elastic $\Sigma \pm p$ scattering at low energies", Physics Letters B **37**, 204 (1971).

[151] K. Miwa et al., "Measurement of the differential cross sections of the $\Sigma^- p$ elastic scattering in momentum range 470 to 850 MeV/c", Phys. Rev. C **104**, 045204 (2021).

[152] T Nanamura et al., "Measurement of differential cross sections for Σ^+p elastic scattering in the momentum range 0.44-0.80 GeV/c", Progress of Theoretical and Experimental Physics **2022**, 093D01 (2022), eprint: https://academic.oup.com/ptep/article-pdf/2022/9/093D01/45684562/ptac101.pdf.

[153] K. Miwa et al., "Precise Measurement of Differential Cross Sections of the $\Sigma^- p \to \Lambda n$ Reaction in Momentum Range $470 - 650\ \mathrm{MeV}/c$", Phys. Rev. Lett. **128**, 072501 (2022).

[154] J. A. Rowley, "Improved Λp Elastic Scattering Cross Sections Between 0.9 and 2.0 GeV/c and Connections to the Neutron Stars" (Ohio University, 2022).

[155] M. Knöll and R. Roth, "Hyperon-nucleon interaction constrained by light hypernuclei", Physics Letters B **846**, 138258 (2023).

[156] P. Eckert, P. Achenbach, et al., "Chart of Hypernuclides — Hypernuclear Structure and Decay Data", hypernuclei.kph.uni-mainz.de, ver. 2022.12.12, 2021.

[157] J. Haidenbauer, personal communication.

[158] J. Kadyk et al., "Λp interactions in momentum range 300 to 1500 MeV/c", Nuclear Physics B **27**, 13 (1971).

[159] J. Hauptman, J. Kadyk, and G. Trilling, "Experimental study of Λp and $\Xi^0 p$ interactions in the range 1 – 10 GeV/c", Nuclear Physics B **125**, 29 (1977).

[160] D. Entem and R Machleidt, "Accurate charge-dependent nucleon-nucleon potential at fourth order of chiral perturbation theory", Phys. Rev. C **68**, 041001 (2003).

[161] P. Navratil, "Local three-nucleon interaction from chiral effective field theory", Few-Body Syst. **41**, 117 (2007).

[162] M. Wang et al., "The Ame2012 atomic mass evaluation", Chin. Phys. C **36**, 1603 (2012).

[163] D. Tilley et al., "Energy levels of light nuclei A=5, 6, 7", Nucl. Phys. A **708**, 3 (2002).

[164] D. Tilley et al., "Energy levels of light nuclei A=8,9,10", Nucl. Phys. A **745**, 155 (2004).

[165] F. Ajzenberg-Selove, "Energy levels of light nuclei $A = 11 - 12$", Nucl. Phys. A **506**, 1 (1990).

[166] H. Le et al., "Separation energies of light Λ hypernuclei and their theoretical uncertainties", The European Physical Journal A **60**, 3 (2024).

[167] J. Kelley et al., "Energy levels of light nuclei A=11", Nuclear Physics A **880**, 88 (2012).

[168] A. Tichai et al., "Natural orbitals for ab initio no-core shell model calculations", Phys. Rev. C **99**, 034321 (2019).

[169] J. Hoppe et al., "Natural orbitals for many-body expansion methods", Phys. Rev. C **103**, 014321 (2021).

[170] R. Roth et al., "Hartree-Fock and many body perturbation theory with correlated realistic NN interactions", Phys. Rev. C **73**, 044312 (2006).

[171] I. Shavitt and R. J. Bartlett, "Many-body methods in chemistry and physics: MBPT and coupled-cluster theory" (Cambridge university press, 2009).

[172] A. Tichai et al., "Hartree–Fock many-body perturbation theory for nuclear ground-states", Physics Letters B **756**, 283 (2016).

[173] S. Perez-Martin and L. M. Robledo, "Microscopic justification of the equal filling approximation", Phys. Rev. C **78**, 014304 (2008).

[174] E. Schrödinger, "Quantisierung als Eigenwertproblem", Annalen der Physik **385**, 437 (1926).

[175] A. Szabo and N. S. Ostlund, "Modern quantum chemistry: introduction to advanced electronic structure theory" (Courier Corporation, 2012).

[176] M. Wang et al., "The AME 2020 atomic mass evaluation (II). Tables, graphs and references*", Chinese Physics C **45**, 030003 (2021).

[177] K. Miyagawa et al., "Properties of the bound $\Lambda(\Sigma)$NN system and hyperon-nucleon interactions", Phys. Rev. C **51**, 2905 (1995).

[178] E. Hiyama et al., "Three-body model study of A=6–7 hypernuclei: Halo and skin structures", Phys. Rev. C **53**, 2075 (1996).

[179] S.-I. Ando, "Hypernuclei in halo/cluster effective field theory", International Journal of Modern Physics E **25**, 1641005 (2016).

[180] Y. Zhang, H. Sagawa, and E. Hiyama, "Hyperon halo structure of C and B isotopes", Phys. Rev. C **103**, 034321 (2021).

[181] H.-T. Xue et al., "Deformation and hyperon halo in hypernuclei", Phys. Rev. C **106**, 044306 (2022).

[182] S. Velardita et al., "Method to evidence hypernuclear halos from a two-target interaction cross section measurement", The European Physical Journal A **59**, 139 (2023).

[183] E. Gebrerufael, "In-Medium No-Core Shell Model for Ab Initio Nuclear Structure Calculations" (Technische Universität Darmstadt, 2017).

[184] K. Vobig, "Electromagnetic Observables and Open-Shell Nuclei from the In-Medium No-Core Shell Model" (2020).

[185] M. Heinz et al., "In-medium similarity renormalization group with three-body operators", Phys. Rev. C **103**, 044318 (2021).

[186] W. Magnus, "On the exponential solution of differential equations for a linear operator", Communications on Pure and Applied Mathematics **7**, 649 (1954).

[187] T. D. Morris, N. M. Parzuchowski, and S. K. Bogner, "Magnus expansion and in-medium similarity renormalization group", Phys. Rev. C **92**, 034331 (2015).

[188] C Samanta, P. R. Chowdhury, and D. N. Basu, "Generalized mass formula for non-strange and hypernuclei with SU(6) symmetry breaking", Journal of Physics G: Nuclear and Particle Physics **32**, 363 (2006).

[189] R. Roth et al., "Medium-Mass Nuclei with Normal-Ordered Chiral $NN+3N$ Interactions", Phys. Rev. Lett. **109**, 052501 (2012).

[190] E. Gebrerufael, A. Calci, and R. Roth, "Open-shell nuclei and excited states from multireference normal-ordered Hamiltonians", Phys. Rev. C **93**, 031301 (2016).

[191] D. Frame et al., "Eigenvector Continuation with Subspace Learning", Phys. Rev. Lett. **121**, 032501 (2018).

[192] T. Duguet et al., "Eigenvector Continuation and Projection-Based Emulators", 2024, arXiv:2310.19419 [nucl-th].

Acknowledgements

Working on this thesis, for the most part during a global pandemic, has been an adventure with a manifold of challenges, which I certainly would not have been able to overcome without the support of the great people around me.

First and foremost, I am deeply grateful for my supervisor Prof. Robert Roth. Not only for the never-ending support of my work and the freedom to pursue my own ideas, but also for exciting astronomy and fun times in Spain. And thank you for the sometimes much-needed pep talks. To this day, it remains a mystery to me how you manage to spark so much motivation and optimism with only a few words.

Second, I want to thank Prof. Achim Schwenk for agreeing to be the second advisor of this work. I will make sure to get you some more cookies.

Next, a big THANK YOU goes out to the people I had the pleasure working with in the TNP++ group throughout the past seven years: Alex, Carl, Cedric, Christina, Julius, Katrin, Klaus, Laura, Lisa, Luis, Marc, Michelle, Patrick, Roland, Stefan, Thomas, Tobi, Tobi, Tobi, Vanessa, and anyone I might have missed here. In particular, I want to thank Laura for all the mega walks, talks, birds, and laughter, and Tobi for several thousand hours on Discord, several thousand kilometers on the road, and hopefully less than a thousand hours of IT support with infinite patience.

Moreover, I want to thank the best office mates I could have asked for: Hannah, Isak, Jan, Lars, Rodric, Yannick, and, to some extend, Matthias. Work is much more fun with you around.

Last but not least, there is a bunch of amazing people, who do not really care about physics but are responsible for the success of this work in their own unique ways: Cathrin, Christina, Dennis, Heiko, Jana, Kim, Lucas, Ole, Robert, and, of course, my family. Thank you for your ongoing support and simply for being part of my life.

Academic CV

Education

2014	**Abitur** Albert-Einstein-Schule, Groß Bieberau
2014 - 2018	**Bachelor of Science, Physics** Technische Universität Darmstadt
2018 - 2019	**Master of Science, Physics** Technische Universität Darmstadt
2020 - 2024	**Doctoral Studies, Physics** Technische Universität Darmstadt

Publications

M. Knöll et al., "Machine learning for the prediction of converged energies from ab initio nuclear structure calculations", Physics Letters B **839**, 137781 (2023)

M. Knöll and R. Roth, "Hyperon-nucleon interaction constrained by light hypernuclei", Physics Letters B **846**, 138258 (2023)

T. Wolfgruber, M. Knöll, and R. Roth, "Precise neural network predictions of energies and radii from the no-core shell model", 2023, arXiv:2310.05256 [nucl-th]

Erklärung gemäß §9 Promotionsordnung

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, den 5. Juli 2024

Marco Knöll