Doctoral Dissertation

博士論文

Uncertainty quantification in configuration interaction methods for nuclear many-body problems

(核子多体問題のための配置間相互作用法における不定性の評価)

A Dissertation Submitted for the Degree of Doctor of Philosophy

December 2019

令和元年 12 月博士(理学)申請

Department of Physics, Graduate School of Science,

The University of Tokyo

東京大学大学院 理学系研究科 物理学専攻

Sota Yoshida

吉田 聡太

Abstract

One of the most fundamental problems in nuclear structure physics is to understand various static properties of nuclei from nucleon degree of freedom.

In the last decades, we have witnessed two key developments in theoretical nuclear structure physics along this direction. One is the fundamental interaction among nucleons based on the chiral effective field theory, and the other one is *ab initio* methods. Those chiral potentials are becoming more quantitative, and the scope of those *ab initio* methods has been rapidly extended to medium-mass nuclei including doubly open-shell nuclei in the last decade.

We are getting an established path from the fundamental interaction to nuclear many-body problems in a wide range of the nuclear chart. At the same time, however, there is another urgent task to be tackled to understand a wide variety of nuclear phenomena starting from the nuclear potentials. That is to quantify theoretical uncertainties, which is the main theme of this thesis. Uncertainty quantification (UQ) is not just putting error bars on the theoretical predictions, but to know the capability of the model itself, i.e., how much we can know about properties of nuclei using those theoretical methods.

In this thesis, we present new methodologies for UQ by taking configuration interaction (CI) methods as examples.

Firstly, we propose a systematic way to achieve UQ for valence shell-model calculations in Chapter 3. The probability distribution of the effective interactions for a physically motivated model space is obtained by the so-called Laplace approximation, and then the uncertainties in parameters are propagated to the quantities of interest. This gives for the first time systematic uncertainty quantification in e.g., energy spectra in this community. Some important outcomes are discussed. More specifically, our method tell us which states are more likely to have *exotic* structures, i.e. states to be considered beyond the adopted model space, such as α-clustering, intruder states, etc. Such information would give important test ground for the nuclear potentials and *ab initio* methods, and facilitate new experiments.

Secondly, in Chapter 4, we introduce a rather general extrapolation method for *ab initio* full configuration interaction method with quantified uncertainties. In our method, extrapolation is performed not by point estimation of a parametric function, but by probabilistic inference among a huge number of different functions, which are generated by means of Gaussian processes with physics constraints. This may alleviate the overfitting caused by a specific choice of the extrapolation function. We present the application of the method to some published results of the full configuration interaction method, and we found that this method tells us one criterion of when to stop the FCI calculations; it is not obviously inadvisable to carry out FCI

calculations while increasing $N_{\max}$ forever. The point to stop can be deduced from analyzing the behavior of the extrapolated results as a function of the maximum $N_{\max}$ available. Here $N_{\max}$ specifies the maximum number of excitations across major-shell gaps.

The main test ground is limited to the configuration interaction methods, but we expect that the methods and its philosophy can be applied to the others too. If we could achieve proper uncertainty quantifications in any theoretical models, it must be very helpful to facilitate complemental developments of theories and experiments towards deeper understandings of various nuclear properties.

Contents

Title i

Abstract iii

1 Introduction **5**

2 Basics of configuration interaction methods and uncertainty quantifications **11**

2.1 Configuration interaction methods: valence and full CI 11

2.1.1 Eigenvalue problems in CI methods 11

2.1.2 Computational costs 14

2.2 Additional truncations 15

2.2.1 Monte Carlo Shell Model 16

2.2.2 Importance-truncated shell model 18

2.2.3 How to recover the original information? 19

2.3 Inputs of CI methods 19

2.3.1 Nuclear potentials from chiral effective field theory 19

2.3.2 Similarity renormalization group (SRG) 23

2.3.3 Derivations of a valence space Hamiltonian from a free-space interaction 24

2.4 Inference of parameters 29

2.5 How to evaluate posteriors? 31

2.5.1 Stochastic method: Markov Chain Monte Carlo 31

2.5.2 Approximate inference 39

3 Uncertainty quantification in valence shell-model calculations 41

3.1 Inference of shell-model effective interactions 42

3.2 Credible intervals of valence shell model for the 0*p*-shell nuclei 45

3.2.1 Detailed analyses of the results 48

3.3 Scalability 51

3.4 Perspectives and remarks 51

4 Uncertainty quantification of extrapolation problem in CI-type calculations 53

4.1 Extrapolation problems in configuration interaction methods 53

4.2 Gaussian processes 56

4.2.1 Constrained Gaussian Processes (CGP) 59

4.3 Application of CGP to extrapolation problems in full CI method 60

4.4 Results and Discussions 63

4.4.1 Extrapolation of g.s. energies 63

4.4.2 Impact of the constraints 65

4.4.3 Data dependence 66

4.5 Short summary of this chapter 67

5 Summary and future perspectives 69

Acknowledgements 73

A Summary of credible intervals of valence CI calculations 75

B The code and related issues on GPs 87

B.1 Technical issue on Gaussian processes 87

B.1.1 Who may give the *correct* answer?: set A 88

B.1.2 How to deal with non-PSD covariance in the codes?: set B 89

B.1.3 Summary 92

Chapter 1

Introduction

In some research fields of physics, theoretical estimates inevitably rely on highly sophisticated models and computational techniques. This is especially true for modern nuclear structure physics. This is partly due to difficulties as finite number many-body problems. The fact that even adjacent nuclei sometimes exhibit significantly different properties indicates that it is essential for quantitative understandings of nuclear phenomena to take account of nucleon degree of freedom explicitly to some extent. In addition to this, the fundamental interaction among nucleons is still ambiguous due to the highly non-perturbative character of the underlying theory, which is Quantum Chromodynamics (QCD), in the low-energy regime, i.e. the main ground of nucleon degree of freedom. These set nuclear many-body problems apart from others; This makes problems very complicated and, at the same time, very fascinating to tackle.

Theoretical studies in nuclear physics, especially in nuclear structure physics today, are roughly classified into two attributes, which will be called "phenomenological" and "*ab initio*" hereafter. Let us here define these terminologies as follows:

1. phenomenological calculations
   Most parameters in theory are determined to reproduce experimental data of quantities of interest, e.g. binding energies.
   For example, the valence shell model, cluster model, antisymmetric molecular dynamics (AMD), the interactive boson model and mean-field calculations with energy density functionals belong to this class.

2. *ab initio* calculations
   Once a free space nuclear force is given, one does not tune any input parameters in theory to improve the agreement with experimental data. In addition to this, if its accuracy can

be systematically improved by taking account of higher order terms and/or extending its model space[1], we call it *ab initio*.
For example, the Faddeev-Yakubovsky scheme, the hyperspherical harmonics method, the auxiliary field diffusion Monte Carlo method, lattice EFT, No core full configuration interaction (NCFC), Green's Function Monte Carlo, Coupled-Cluster (CC), In-Medium Similarity Renormalization Group (IM-SRG), self-consistent Green's function theory belong to this class.

In the classification above, we put the valence shell model into the phenomenological group. However, we note that the valence shell model is now can be placed in between phenomenological and *ab initio*. This will be briefly reviewed in Sec. 2.3.

In this Ph.D. thesis, we provide the new methodologies to quantify how robust are those calculations and to extract much more information from those calculations. The main test ground is limited to the CI methods, but we expect that the methods and its philosophy can be applied to the others too.

In order to make problems clear, let us quote some parts of the Editorial on Phys. Rev. A [1]:

> *It is not unusual for manuscripts on theoretical work to be submitted without uncertainty estimates for numerical results. [...] It is all too often the case that the numerical results are presented without uncertainty estimates. Authors sometimes say that it is difficult to arrive at error estimates. Should this be considered an adequate reason for omitting them? [...] The comparison with experiment itself provides a test of our theoretical understanding. However, there is a broad class of papers where estimates of theoretical uncertainties can and should be made. Papers presenting the results of theoretical calculations are expected to include uncertainty estimates for the calculations whenever practicable, and especially under the following circumstances:*
>
> 1. *If the authors claim high accuracy, or improvements on the accuracy of previous work.*
> 2. *If the primary motivation for the paper is to make comparisons with present or future high precision experimental measurements.*
> 3. *If the primary motivation is to provide interpolations or extrapolations of known experimental measurements.*
>
> *[...]*

[1] We will define the terminology "model space" later.

These are true not only for manuscripts in Phys. Rev. A covering atomic, molecular, and optical physics and quantum information, but also for manuscripts in many scientific fields. We would like readers to refer to the full text of [1] to prevent misunderstandings due to our partial citation.

In nuclear physics community, the history of uncertainty quantification (UQ) is not so long. Importance of UQ in theoretical estimates in the context of nuclear models is well explained by Refs. [2, 3] and references therein.

As far as we perform calculations using a *model* with the parameters determined by experimental data, we must try to answer quantitatively at least to the following questions:

1. *How reliable and precise are your results?*
2. *What happens if you add new data in the parameter fit?*
3. *What is the origin of the discrepancy between the experimental data and the current prediction?*

These are particularly important for making reliable predictions for proton- or neutron-rich nuclei for which experimental studies are extremely difficult to carry out.

The main theme of this thesis is to make a progress to answer these questions particular in a shell-model calculation, which is generally called configuration interaction (CI) in other fields. The shell-model is superior in terms of its capability; if the computation is feasible, one can obtain ground states, excited states, and transition strengths among those in a consistent manner. For this reason, shell-model results have been widely used to compare with new experimental data and also compare with other theoretical calculations. Moreover, the shell model is no longer just a phenomenological model; it plays a key role to survey the connections between underlying physics and observables of many-nucleons systems owing to the recent developments in modern description of nuclear forces with chiral effective field theory (χEFT), which is consistent with underlying symmetry and its breaking patterns, and in many-body techniques to derive effective nuclear force for valence nucleons. See also Sec. 2.3.1 and Sec. 2.3.3 for more details.

Under these circumstances, uncertainty quantifications in configuration interaction methods, including its *ab initio* counterpart, are of great importance from various viewpoints.

Here is an example. In many literatures, figures like FIG. 1.1 comparing excitation spectra are often shown. We note that the experimental observation errors are too exaggerated and one typically cannot see any error bars in such a scale thanks to the high resolution. An important remark here is that from this kind of plot, there is no information at all about how much reliable

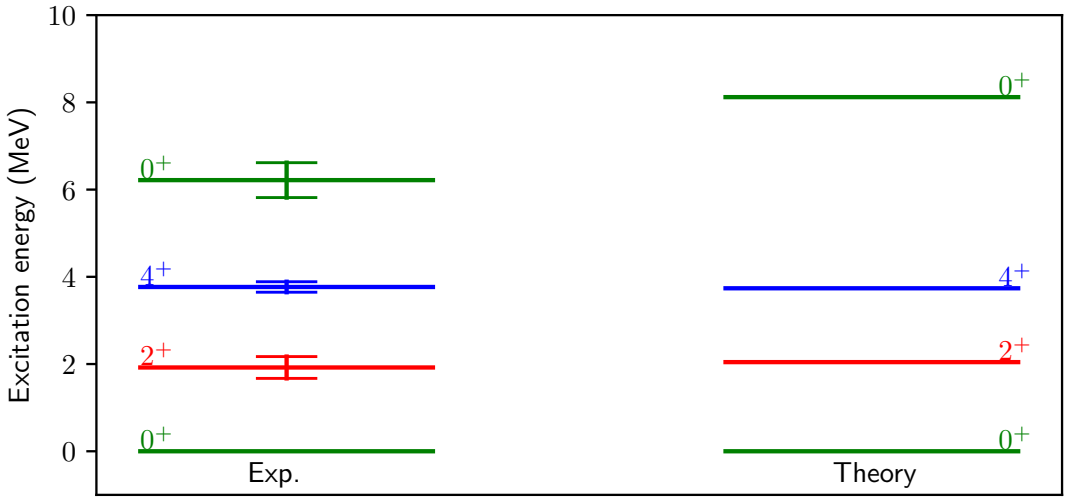

FIG. 1.1: Schematic figure of comparison of excitation spectra between theory and experiments. The unlikely large error bars are drawn to make the *issue* clear.

is this theoretical calculation. One cannot figure out whether or not the agreement in the lowest two excited states and disagreement in the second 0^+ state are accidental unless estimating theoretical uncertainties coming at least from the parameters and many-body method itself. In this thesis, we propose ways to resolve these problematic circumstances.

This thesis is organized as follows: In Chapter 2, some basics about CI methods and uncertainty quantification are introduced. After that, in Chapter 3, we discuss how to make comparisons like FIG. 1.1 more insightful by considering theoretical uncertainties coming from input parameters. In addition to this, we also discuss a novel method to evaluate an intrinsic uncertainty of *ab initio* full configuration interaction method in Chapter 4. Conclusions and some future perspectives are summarized in Chapter 5. Most of the data and codes used in Chapter 3 and Chapter 4 are opened on GitHub [4,5] for reproducibility and to facilitate future studies.

Here are the examples of possible impacts of this study from three viewpoints:

1. The relationships with experimental studies
   In many literatures of experimental studies, *e.g.* gamma-ray spectroscopy, results are shown in comparison with shell-model calculations. The UQ results enable us to quantify the reliability of their statements deduced from comparison with model wave functions. The information of theoretical uncertainty can be used to help to construct level schemes and to design new measurements.

2. The relationships with *ab initio* frameworks
   The UQ in configuration interaction methods may indicate some missing contribution in

current nuclear potentials or many-body techniques to derive effective interactions for a valance space. This would tell us which higher-order components should not be omitted.

3. The impacts on other research fields
   The UQ in shell model may contribute to the understandings in other research fields. For example, beta-decay quantities like half-lives and neutron emission rates have an effect on astrophysical simulations of nucleosynthesis. Moreover, some quantities from nuclear structure calculations play a key role in physics about more fundamental degrees of freedom. For example, nuclear matrix elements of double beta decay are crucial to understand properties of neutrinos.

Through out this thesis, we will introduce some abbreviations. We summarize them in Table 1.1 to prevent readers from confusing.

Table 1.1: Abbreviations and acronyms used in this thesis.

Abbreviations/Acronyms	explanation
A	mass number
ANN	**A**rtificial **N**eural **N**etwork
CC	**C**oupled-**C**luster
cGP	**c**onstrained **G**aussian **P**rocess
CI	**C**onfiguration **I**nteraction
CK	**C**ohen **K**urath
EFT	**E**ffective **F**ield **T**heory
FCI	**F**ull **C**onfiguration **I**nteraction
GP	**G**aussian **P**rocess
g.s.	**g**round **s**tate
HMC	**H**amiltonian **M**onte **C**arlo
ITSM	**I**mportance **T**runcated **S**hell **M**odel
IMSRG	**I**n-**M**edium **S**imilarity **R**enormalization **G**roup
LA	**L**aplace **A**pproximation
LEC(s)	**L**ow-**E**nergy **C**onstant(s)
LS	spin-orbit (force)
MAP	**M**aximum **A** **P**osteriori
MBPT	**M**any-**B**ody **P**erturbation **T**heory
MCMC	**M**arkov **C**hain **M**onte **C**arlo
MCSM	**M**onte **C**arlo **S**hell **M**odel
N	neutron number
NLO	**n**ext-to-**l**eading **o**rder
NNLO, N2LO	**n**ext-to-**n**ext-to-**l**eading **o**rder
N3LO	**n**ext-to-**n**ext-to-**n**ext-to-**l**eading **o**rder
NO2B	**n**ormal **o**rdered **t**wo-**b**ody
QCD	**Q**uantum **C**hromodynamics
RWMH	**R**andom **W**alk **M**etropolis-**H**astings
SPE	**S**ingle **P**article **E**nergy
TBME	**T**wo-**B**ody **M**atrix **E**lement
UQ	**U**ncertainty **Q**uantification
Z	proton number

Chapter 2

Basics of configuration interaction methods and uncertainty quantifications

In this chapter, we introduce basics of configuration interaction methods[1], and of uncertainty quantifications in nuclear many-body problems by taking the CI methods as examples.

In Sec. 2.1, we briefly summarize fundamental topics of CI calculations in nuclear physics, and then we explain the basics of Bayesian inference in Sec. 2.4, sampling schemes in Sec. 2.5.

2.1 Configuration interaction methods: valence and full CI

2.1.1 Eigenvalue problems in CI methods

The basic question in the nuclear many-body problems is how to solve the non-relativistic Schrödinger equation for identical fermions:

$$\hat{H}\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ...) = E\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ...), \tag{2.1}$$

$$\hat{H} \equiv \sum_i^A \left(-\frac{\hbar^2}{2m}\nabla^2 \right) + \sum_{i<j} V(\boldsymbol{r}_i - \boldsymbol{r}_j) \tag{2.2}$$

[1]The reason why we use the plural form is that we distinguish full configuration interaction from one for a valence space.

where we assumed the system consists of A nucleons with mass m. The A nucleons in a given nucleus interact with each other through $V(\boldsymbol{r}_i - \boldsymbol{r}_j)$. We omitted an external potential $U(\boldsymbol{r}_i)$ in the equation above.

The idea of configuration interaction methods is to expand the wavefunction Ψ in a convenient many-body basis to handle and the common choice for that is Slater determinants:

$$\Psi_n = \sum_j c_j^{(n)} \Phi_j, \tag{2.3}$$

$$\Phi_j \propto \det \begin{bmatrix} \varphi_{j_1}(1) & \cdots & \varphi_{j_1}(A) \\ \vdots & \ddots & \vdots \\ \varphi_{j_A}(1) & \cdots & \varphi_{j_A}(A) \end{bmatrix}, \tag{2.4}$$

where we consider the eigenstate of the Hamiltonian labeled by the index n, and the labels $\{j_1, j_2, ..., j_A\}$ specify many-body configurations under the given single-particle basis. This naturally incorporates Pauli principle and antisymmetric property of fermionic wave functions.

We move to Dirac's bra-ket notation, and then, in harmonic oscillator single-particle basis, the many-body state Φ_j is written in a second quantized form

$$|\Phi_j\rangle = a_{j_1}^\dagger a_{j_2}^\dagger \cdots a_{j_A}^\dagger |vacuum\rangle, \tag{2.5}$$

where $a_j^\dagger$ is the operator that creates the single fermion state j.

We here define two terminologies, *valence CI* and *full CI*, by the difference in the choice of the vacuum state. The schematic pictures of full CI and valence CI are shown in FIG. 2.1. In full CI, $|vacuum\rangle$ is a fermion vacuum. On the other hand, $|vacuum\rangle$ in valence CI is an inert core, which is usually taken as the nearest doubly closed-shell nuclei.

Now, the eigenvalue problem can be written as:

$$\sum_j \langle \Phi_k | H | \Phi_j \rangle c_j^{(n)} = E_n c_k^{(n)}. \tag{2.6}$$

In many practical situations, $\langle \Phi | H | \Phi \rangle$ is the large sparse matrix and one is interested in tens of low-lying states. Therefore, the Lanczos method [6] becomes a method of choice for those eigenvalue problems. One of the major advantages of CI methods is that one can evaluate ground state energy, excitation spectra, and other observables in a consistent manner: Once one obtained eigenstates, any observable of the target nucleus can be obtained by evaluating expectation values for the operators of interest.

In general, we work with rotationally invariant Hamiltonians. Therefore, the eigenstates of

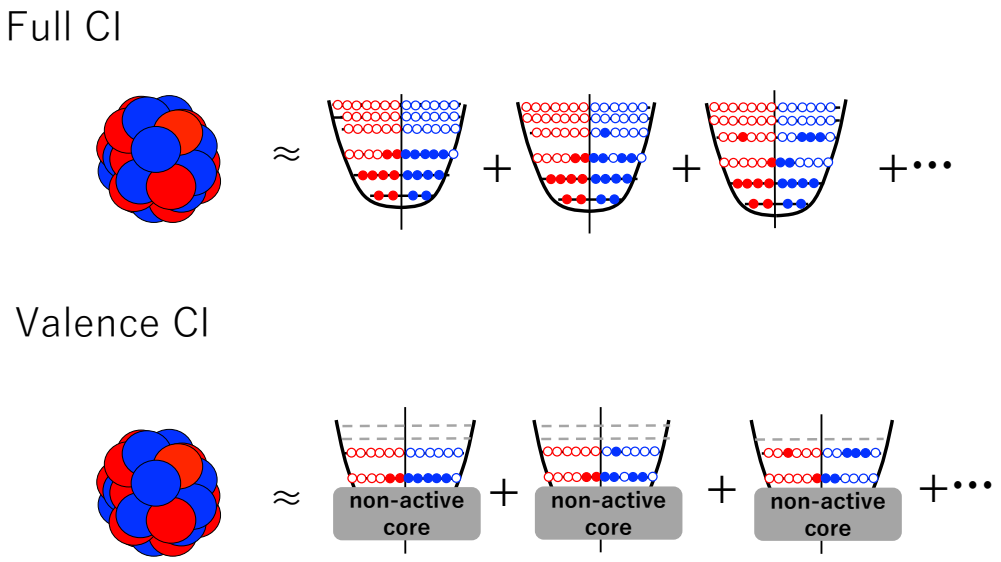

FIG. 2.1: The schematic figure of full and valence CI.

the Hamiltonian have good angular momenta.

$$H|\Psi\rangle = E|\Psi\rangle, \tag{2.7}$$
$$J^2|\Psi\rangle = J(J+1)|\Psi\rangle, \tag{2.8}$$
$$J_z|\Psi\rangle = M|\Psi\rangle, \tag{2.9}$$

where we followed the convention to take z-axis as the quantization axis, and M can be $-J, -J+1, \cdots, J$. The value of E does not depend on a value of $M = \sum_i^A m_i$, where m_i is j_z value of i-th single particle wave function. If we work with all states of a given M under the spherically symmetric Hamiltonians, the eigenstates is guaranteed to have a good quantum number J. This is the so-called M-scheme and $M = 0$ and $M = 1/2$ are used for even and odd nuclei, respectively. Many shell-model codes in M-scheme are available: ANTOINE [7], BIGSTICK [8], MFDn [9–11], and KSHELL [12, 13].

One can also work with the so-called J-scheme, i.e. working with the basis which has a definite J value. It is known to be more costly to achieve calculations in J-scheme although the matrix size is usually smaller than that in M-scheme by an order of magnitude. There are also J-scheme codes on the market such as NuShellX [14].

Here, to avoid confusion, we explicitly introduce the definition of a terminology *model space* in this thesis as follows: the all activated single-particle orbitals and many-body configurations allowed. The latter one is simply called configurations from now on.

In full CI (FCI) calculations, all nucleons are activated under the given single-particle basis and wave functions are typically represented in a truncated harmonic oscillator basis. This

involves two basis space parameters, $\hbar\Omega$ and $N_{\max}$. The $\hbar\Omega$ is the expansion parameter of the basis states and the spacing between oscillator major shells. The $N_{\max}$ defines the maximum number of harmonic oscillator quanta[2] $N = 2n+\ell$ allowed in the many-body basis states above the lowest configurations for a target nucleus, where n and ℓ are the principal quantum number and the azimuthal quantum number respectively. The naive filling configurations correspond to $N = 0$. In the case of ^{6}Li, for example, the configurations with one proton excitation from the $0p$ shells ($n = 0, \ell = 1$) to the $1s$ shell ($n = 1, \ell = 0$) and two neutron excitations from the $0s$ shell ($n = 0, \ell = 0$) to the $0d$ shells are included in the configurations with $N_{\max} \geq N = 5$.

2.1.2 Computational costs

Table 2.1: Rough estimates of M-scheme dimensions for FCI.

	M-scheme dim.			
N_{max}	^{4}He	^{6}Li	^{12}C	^{16}O
2	5.9×10^1	8.0×10^2	1.8×10^4	1.3×10^3
4	9.5×10^2	1.7×10^4	1.1×10^6	3.5×10^5
6	7.9×10^3	2.0×10^5	3.3×10^7	2.7×10^7
8	4.5×10^4	1.6×10^6	5.9×10^8	1.0×10^9
10	2.0×10^5	9.7×10^6	7.8×10^9	2.4×10^{10}
12	7.1×10^5	4.9×10^7	8.1×10^{10}	4.1×10^{11}
14	2.2×10^6	2.1×10^8	6.9×10^{11}	5.4×10^{12}

Table 2.2: Estimates of M-scheme dimensions for valence CI.

nucleus	inert core	model space	M-scheme dim.
^{6}Li	^{4}He	$0p$ (full)	1.0×10^1
^{12}C	^{4}He	$0p$ (full)	5.0×10^1
^{28}Si	^{16}O	$1s, 0d$ (full)	9.4×10^4
^{28}Si	^{16}O	$1s, 0d, 1p, 0f$ ($2\hbar\omega$)	3.9×10^7
^{28}Si	^{16}O	$1s, 0d, 1p, 0f$ (full)	3.7×10^{10}
^{56}Ni	^{40}Ca	$1p, 0f$ (full)	1.1×10^9

In valence CI calculations, we assume the existence of an inert core, i.e. we divide the A nucleons into two ingredients, an inert core and valence particles. If we naively compare between full CI and valence CI methods for the same nucleus, the computational cost is significantly reduced by assuming the existence of the inert core. This enables us to extend

[2]Other than this paragraph, N denotes the neutron number.

our scope to medium-mass nuclei. Of course, as described later in the context of effective interactions, it does not necessarily mean that it is easier to obtain the accurate results with valence CI over the full CI method.

Now we explore how much computational costs are typically required in FCI and valence CI calculations. In Tables 2.1 and 2.2, we summarized M-scheme dimensions of some nuclei with positive parity and $M = 0$ in both cases, full and valence CI[3]. The M-scheme dimensions are evaluated with KSHELL code [12]. One can see that M-scheme dimensions increases as the model size is extended and, especially in the FCI case, this easily exceeds the current limitation, which is said to be around 10^{11}. As will be shown in Chap. 4, the values at $N_{\max}$ around $10 - 20$ is still far from exact values. Regarding the valence CI case, "(full)" means that all possible excitations in the model space are allowed and $X\hbar\omega$ express the typical way to specify the truncation of the configurations. In the case of ^{28}Si in the $sd - pf$ shells, $2\hbar\omega$ means that only two excitations across the major shell gap are allowed, while all excitations among the orbitals in the same major shell are taken into account. There are more alternatives to specify the truncation such as one based on the occupation numbers of each orbital.

2.2 Additional truncations

In the last subsection, we showed rapid increases in the M-scheme dimension as one goes to heavier nuclei. Even if we introduce the truncation of the number of particle-hole excitations, it reduces only a few order of magnitudes. This reduction of the dimension does alleviate the problem, however, it can be easily imagined that this does not help any more in heavier systems or more neutron- or proton-rich nuclei where more than one-major shell to be taken into account.

One needs more efficient representations of the many-body wave functions while keeping original information as much as possible. To this end, we here introduce some additional truncation techniques which are used in frontier studies using CI methods in nuclear structure physics.

Now we introduce the terminology *importance truncation* that gives an efficient representation of the many-body wave functions. We will introduce two different methods to incorporate the idea of importance truncation, but the essential idea is common to both. The general idea of importance truncation is selecting a small subset of the many-body basis states which is physically more relevant to the observables of interest. If one could find a subset of many-body states that dominates most of the exact wave function, the usage of that subset gives a very

[3]The estimates for FCI are done under the $e_{\max} = 9$, which is the maximum number of the harmonic oscillator quanta for the single-particle states.

efficient evaluation of the observables and significantly reduces computational costs. This is because, in that case, it is expected that the rest part, which makes the discrepancy between the exact one and CI estimates, can be regarded as non-informative component of nuclear structure and then can be restored by some extrapolation method.

In the rest of this section, we will briefly introduce the Monte Carlo Shell Model [15, 16] and Importance-Truncated Shell Model [26]. We note that there are other alternatives to work with a smaller basis size to approximate wave functions using density matrix renormalization group (DMRG) [17, 18] or the hybrid multi-determinant (HMD) method [19, 20].

2.2.1 Monte Carlo Shell Model

In the Monte Carlo Shell Model (MCSM) [15, 16], a wave function is represented as linear combination of the projected deformed Slater determinants:

$$|\Psi^{MCSM}_{N_\#}\rangle = \sum_{n=1}^{N_\#} \sum_{K=J}^{J} f^{J\pi}_{n,K} P^{J\pi}_{MK} |\tilde{\Phi}_n\rangle, \tag{2.10}$$

where $N_\#$ is the number of basis and the $P^{J\pi}_{MK}$ is the projection operator in terms of both the total angular momentum J and the parity π. This is given as

$$P^{J\pi}_{MK} = \frac{1+P(\pi)}{2}\frac{2J+1}{8\pi^2}\int d\Omega D^{J*}_{MK}(\Omega) e^{i\alpha\hat{J}_z} e^{i\beta\hat{J}_y} e^{i\gamma\hat{J}_z}. \tag{2.11}$$

Here $P(\pi)$ takes $\pm P$[4] for $\pi = \pm$ with double-sign in same order, $D^{J*}_{MK}(\Omega)$ is the Wigner D-matrix, and $\Omega \equiv (\alpha, \beta, \gamma)$ are the Euler angles. The $|\tilde{\Phi}_n\rangle$ is deformed Slater determinant:

$$|\tilde{\Phi}_n\rangle = \prod_{\alpha=1}^{A} \left(\sum_{j=1}^{N_{sp}} M_{j\alpha} a^\dagger_j \right) |vacuum\rangle, \tag{2.12}$$

where A is the number of nucleons activated, and $N_{sp}(\geq A)$ is the number of the single-particle basis states. Imposing the normalization condition on $|\tilde{\Phi}_n\rangle$, we obtain the relation

$$\sum_j (M_{j\alpha})^\dagger M_{j\alpha} = 1, \tag{2.13}$$

for any α.

Since the superposition of those deformed Slater does not have a good angular momentum in general, we need projections shown above to restore the original rotational symmetry.

[4]The P is space inversion operator.

One advantage of the deformed Slater is that one can obtain an intuitive understandings of deformation of the target nucleus. Taking a close look into the overlap between the calculated wave function and the deformed Slater determinants on the potential energy surface, we can visually understand the deformation of the target nucleus. We refer the readers to, e.g. [21] introducing the so-called T-plot.

Now let us go back to the eigenvalue problem with these MCSM wave functions. The problem of interest is the following generalized eigenvalue problem:

$$\sum_{nK}\langle\tilde{\Phi}_m|H f^{J\pi}_{n,K}P^{J\pi}_{MK}|\tilde{\Phi}_n\rangle = \mathcal{E}\sum_{nK}\langle\tilde{\Phi}_m|f^{J\pi}_{n,K}P^{J\pi}_{MK}|\tilde{\Phi}_n\rangle, \tag{2.14}$$

where the $\mathcal{E}$ is the energy eigenvalue and the H is the adopted Hamiltonian. See the Sec. 2.3 for more details on H. Then, we obtain the energy eigenvalue and the coefficients $\{f^{J\pi}_{n,K}\}$. The main task in MCSM is to approximate well the exact wave function, which is generally not feasible to evaluate, with less number of basis $N_\#$. To this end, we need to optimize the $\{M_{j\alpha}\}$. Let $M^{(n)}$ denote the transformation matrix between spherical and deformed Slater determinants for n-th MCSM basis state.

The first procedure is to generate many candidates for the $M^{(1)}$ and then to choose the best one to minimize the energy. Those samples are generated in a stochastic way using the auxiliary-field Monte Carlo technique, see Ref. [15,16] for more details. The second procedure is the optimization of $M^{(1)}$ so as to minimize the energy eigenvalue using the conjugate gradient method.

These steps are iterated by increasing $N_\#$ by one until the energy eigenvalue becomes rather *stable*. It is of course the tradeoff between convergence behavior and computational costs. We note that the optimization of $M^{(i)}$ is performed while fixing $M^{(1)}, M^{(2)}, \ldots, M^{(i-1)}$. This guarantees the variational feature of the energy eigenvalue with respect to $N_\#$. This iterative procedure is called the sequential conjugate gradient (SCG) method in Refs. [22–24].

We also note that there is an alternative way to construct the MCSM basis states: one can choose N_s and optimize the $M^{(1)}, M^{(2)}, \ldots, M^{(N_s)}$ simultaneously using e.g. the conjugate gradient method.

Typically, one can obtain plausible estimates of the quantity of interest with an oder of 100 MCSM basis states. For this reason, the MCSM technique have significantly extended our scope with CI methods to medium-mass nuclei. Representative examples can be found in Ref. [25]: the original dimension of the basis vector is beyond 10^{31}.

By definition, the calculated lowest energy eigenvalue is always an upper bound to the exact ground state energy under the given interaction. This feature is useful to analyze the convergence pattern of the calculation with respect to the number of MCSM basis states.

2.2.2 Importance-truncated shell model

We here introduce an alternative approach to approximate the wave function with less number of basis states. For the applications to the valence CI case, the method is called in different names such as the importance-truncated shell model (ITSM) or the importance-truncated valence-space shell model (IT-VSSM). Its counterparts for FCI is known as importance-truncated no-core shell model (IT-NCSM). The primary studies using IT-SM focused on FCI calculations [26–30] but then naturally applied to the valence CI case, too [31]. The essential idea is common among them, so we simply call the method ITSM for the sake of simplicity.

In the ITSM, one starts with a *reference state*

$$|\Psi_{\rm ref}\rangle = \sum_{j\in\mathcal{M}_{\rm ref}} c_{j,{\rm ref}}|\Phi_j\rangle \tag{2.15}$$

which represents the initial approximation of the exact wave function using less number of basis. In the case of valence CI, for example, the wave function calculated with the $0\hbar\omega$ truncation in which no excitation across a major-shell gap is allowed. Then, one evaluates the importance of the contributions from basis states outside the model space, $|\Phi_i\rangle \notin \mathcal{M}_{\rm ref}$:

$$\kappa_i = -\frac{\langle\Phi_i|H|\Psi_{\rm ref}\rangle}{\epsilon_j - \epsilon_{\rm ref}} = -\sum_{j\in\mathcal{M}_{\rm ref}} c_{j,{\rm ref}}\frac{\langle\Phi_i|H|\Phi_j\rangle}{\epsilon_j - \epsilon_{\rm ref}}, \tag{2.16}$$

where $\epsilon_j - \epsilon_{\rm ref}$ corresponds to the excitation energy which is evaluated by the unperturbed single-particle energies. The key idea of ITSM is that one chooses a fixed importance threshold $\kappa_{\rm min}$ and then include the all basis states whose absolute values of importance measure, $|\kappa|$, larger than $\kappa_{\rm min}$. The basis states are now iteratively constructed as a function of $\kappa_{\rm min}$. Let $\mathcal{M}_{IT}$ denotes the set of basis states included by the iterations. After the iterations, wave functions in IT-SM are represented as

$$|\Psi^{IT}_{\kappa_{\rm min}}\rangle = \sum_{j\in\mathcal{M}_{IT}} c_{j,{\rm ref}}|\Phi_j\rangle. \tag{2.17}$$

Since the energy eigenvalues obtained by ITSM are approximations to the exact values and these are computed with a subset of basis states of the original valence space, the energy eigenvalues are variational:

$$E_{\rm ref} = E(\infty) \geq E(\kappa_{\rm min}) \geq E(0) = E_{\rm exact}. \tag{2.18}$$

2.2.3 How to recover the original information?

We have discussed so far how to represent CI wave functions in an effective way. However, it is naturally expected that there still remains a finite discrepancy between the MCSM/ITSM output and the full model-space calculation. In general, one has to extrapolate the sequence of results by MCSM/ITSM to estimate the exact value of the observables of interest. That is typically done by the chi-squared minimization with low-rank polynomials. There is an extrapolation method using energy variances that makes the extrapolation in MCSM more stable [22, 32]. However, that requires additional computational costs to calculate expectation values of H^2, $\langle \Psi^{MCSM} | H^2 | \Psi^{MCSM} \rangle$. The efficiency is determined by the tradeoff between the additional cost to evaluate $\langle H^2 \rangle$ and the reduction of the original dimension size.

We will revisit these extrapolation issues in Chapter 4.

2.3 Inputs of CI methods

In this section, we will give a brief overview of recent progresses in inputs of nuclear many-body methods including CI-type calculations. In Sec. 2.3.1 and Sec. 2.3.2, we briefly explain recent progresses in describing free-space nucleon-nucleon interactions and higher many-body forces by chiral effective field theory [33–35, 54] and how convergence of many-body methods could be fastened by the so-called similarity renormalization group method [99, 100]. Then, in Sec. 2.3.3, we introduce some a way to link those free-space nucleon potentials to the so-called effective interactions for a valence space.

2.3.1 Nuclear potentials from chiral effective field theory

One of the most important developments in recent theoretical nuclear physics is the description of nuclear potentials by chiral effective field theory (chiral EFT).

The chiral EFT is a low-energy effective field theory of Quantum Chromodynamics (QCD) and has information of the symmetry and its breaking pattern of QCD. The chiral EFT gives a connection between nucleon degree of freedom and quarks and gluons. The history of chiral EFT goes back to the S. Weinberg's seminal papers [33,34] about the chiral perturbation theory. Hereafter, we will focus on the application of chiral perturbation theory to the description of nuclear potentials.

The general idea of effective field theory is the separation of scales. When one is interested in some system, one has to properly choose the relevant degrees of freedom to capture the

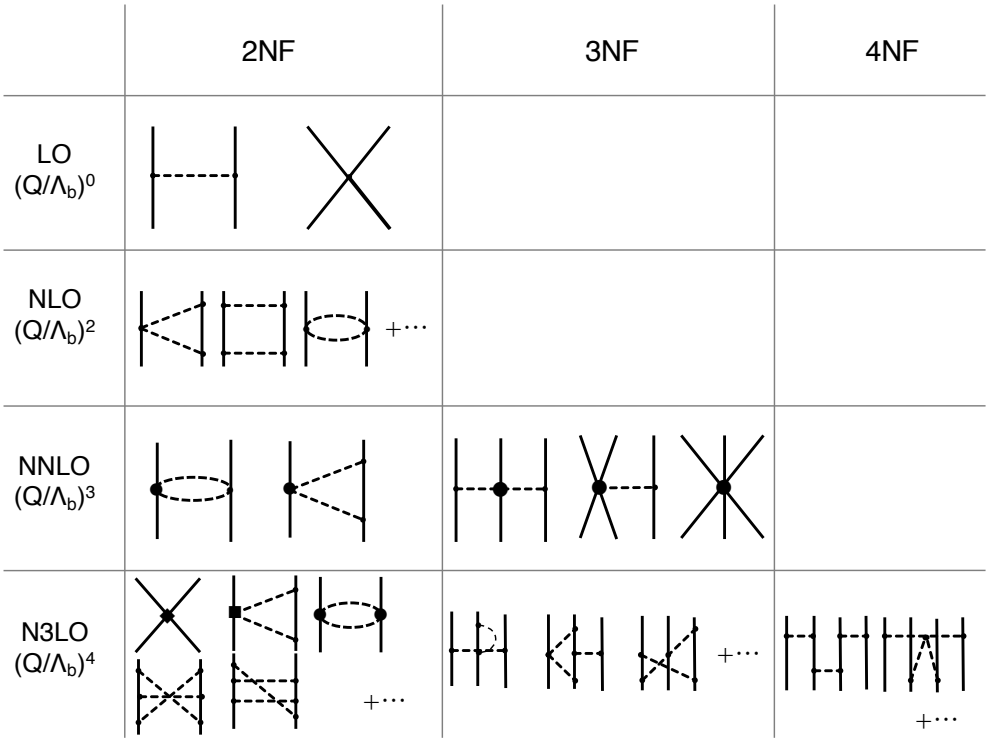

FIG. 2.2: Hierarchy of nuclear forces from chiral effective field theory under the conventional Weinberg's power counting. Solid lines and dashed lines are nucleon and pion lines, respectively. Vertices with small dots, large dots, solid squares, and diamonds correspond to, and $\Delta_i = 0, 1, 2,$ and 4, respectively.

dominant part of the system. This is achieved by introducing a typical momenta Q and the breakdown scale of the effective field theory Λ_b. Then, each contribution is classified by perturbative expansion in terms of $(Q/\Lambda_b) << 1$, and its hierarchy is expressed by the power ν. The effective Lagrangian is given by

$$\mathcal{L}_{eff.} = \sum_{\nu} \left(\frac{Q}{\Lambda_b}\right)^{\nu} \mathcal{F}_{\nu}(Q, \{c_i\}), \tag{2.19}$$

where $\mathcal{F}_{\nu}$ is a function of Q and $\{c_i\}$ and at the given order ν. The the $\{c_i\}$ is a set of free parameters in theory, which is the so-called low-energy constants (LECs).

By choosing the typical low-momentum scale Q as nucleon momenta of the order of the pion mass and the breakdown scale to be $\Lambda_b \sim 1$ GeV, the relevant degrees of freedom are only nucleons and pions, and heavy mesons and nucleon resonances are integrated out, and the effects of those degrees of freedom are implicitly included through LECs. These LECs are usually determined by fitting to nucleon-nucleon (NN) or pion-nucleon (πN) scattering data.

In FIG. 2.2, we show a schematic figure of nuclear forces in chiral effective field theory.

Under a given power counting scheme, one can classify all contributions like this figure. Here we followed the convention, the power counting introduced in Ref. [35]. In this power counting scheme, the power ν of the chiral expansion is given as

$$\nu = -2 + 2A - 2C + 2L + \sum_i \Delta_i, \tag{2.20}$$

$$\Delta_i \equiv d_i + \frac{n_i}{2} - 2, \tag{2.21}$$

where A is nucleon numbers involved, C denotes the number of separately connected pieces, L is the number of loops in the diagram, d_i is the number of derivatives or pion-mass insertions, and n_i denotes the number of nucleon legs in vertex i. Here the sum runs over all vertices i in the diagram. This definition of ν is slightly different from Weinberg's original power counting by $3A - 6$. By this, ν is essentially free from A-dependence that is not suitable for $A \geq 3$ systems [35]. We refer the readers interested in the power counting issue and related ones to Refs. [36–41].

The following points are often cited as excellent points of chiral EFT potentials:

- The accuracy of the perturbative expansion can be systematically improved.
- This enables us to take account of many-body (three-body, four-body, ...) forces in a systematic way.
- The derivation of many-body currents can be achieved in a consistent way.

As can be seen from FIG. 2.2, more and more diagrams appear in higher orders. At the leading order (LO), we only have one-pion exchange term (the left diagram) and contact term (the right one) in two nucleon force (2NF). The next order contribution, next-to-leading order (NLO), is not $\nu = 1$, but $\nu = 2$. This is due to the symmetry. Three different two-pion exchange contributions at NLO are shown in the figure. At next-to-next-to-leading order (NNLO), the 2NF contributions with different coupling constants appear, which have $\Delta_i = 1$, and the three-nucleon forces firstly emerge. From left to right, two-pion exchange (TPE) term, one-pion exchange (OPE) term, and contact (CON) term, respectively. While the LECs associated with TPE (c_1, c_3, and c_4) are common with 2NF, the LECs with OPE (c_D) and CON (c_E) firstly appear in three-nucleon force at NNLO. These c_D and c_E are determined by fitting to some quantities which are relevant to many-body systems larger than two-body. In the same manner, four-body, five-body, and higher many-body forces can be systematically derived in principle. In what follows, let us consider up to three-body forces for practical purposes. We note that many-diagrams for 3NF appear at N3LO, but there is no new LEC at this order, i.e. only the LECs already determined in the NN-sector appear.

Many quantitative chiral forces are available now and those are classified into the two families: EM(EMN) and EGM(EKM). These are the acronyms of the authors of original and following studies. Let EM denote the chiral EFT interactions derived by D. R. Entem and R. Machleidt [35, 42–47]. These are sometimes called Idaho potentials. We note that in some recent studies [45–47], N. Kaiser and/or Y. Nosyk are also included as authors, and after Ref. [47] the acronym EMN is now widely used for the up-to-date NN interaction. On the other hand, let EGM/EKM denote ones by E. Epelbaum, Ulf-G. Meißner and their collaborators [37, 40,41,48–62]. The interactions which belong to the latter family are also called Bochum-Jülich potentials. One of the major differences among them lie in their regularization schemes.

We have many options to determine LECs in 2NF and 3NF. We classified them into the following three: (A) Extrapolative, (B) Interpolative, and (C) Order-by-oder.

For (A) and (B), the way to determine LECs appeared in 2NF is common, which is fit to scattering data, and the difference lies only in the determination of the LECs in 3NF. In type-(A), c_D and c_E are determined by fitting to observables of few-body systems such as the triton binding energy and the nd doublet scattering length ${}^2a_{nd}$ [48], the binding energies of ^{3}H and ^{4}He [63], the binding energy of ^{3}H and the charge radius of ^{4}He [64], and the properties of p-shell nuclei [65]. An alternative approach is using weak processes [66–70]. A common thing among them is that those LECs are determined by few-body systems, so the usage of those LECs in heavier nuclei can be called extrapolative.

Some studies using type-(A) interactions have exhibited that chiral EFT potentials which are accurate for light nuclei vary considerable in heavier nuclei [71] and in saturation point for symmetric nuclear matter [64]. One of the chiral interaction, the so-called EM1.8/2.0, in Ref. [64] is known to give rather plausible saturation point and systematic description of binding energies of light to medium-mass nuclei [71, 72]. This indicates that it is essential for quantitative chiral interactions to be consistent with both few-body and matter properties. We refer to the interactions in which empirical saturation points are considered as type-(B) interactions. There are not so many studies along this line, but this is one possible choice of future generations of chiral interactions, see Ref. [73, 74]. The usage of this type of interaction in nuclear structure calculation can be regarded as interpolative in terms of mass region.

Alternative approach is simultaneous fit of all LECs to scattering data (and, in some cases, many-body observables of medium mass nuclei). This approach is used in various interactions such as NNLO$_{\rm opt}$ [75] and NNLO$_{\rm sat}$ [76], and NNLO$_{\rm sep}$/NNLO$_{\rm sim}$ [77]. The successful description of binding energies and charge radii from light to medium-mass nuclei with NNLO$_{\rm sat}$ interaction using normal-ordered two-body approximation indicates that effects of 3NFs can be more or less absorbed into two-body part as effective NN force(, see also Figure 11 in Ref. [78]).

We note that there are some studies of effective NN interaction without explicit 3NFs such as a family of JISP (J-matrix inverse scattering potentia) [79] (JISP6 [80] and JISP16 [81]), and Daejeon16 [82, 83]. These numbers 6 and 16 mean that data up to $A \leq 6$ and $A \leq 16$ is used to construct the effective interactions.

We also note that there are several works on chiral interactions from different aspects. In Ref. [84–88], 3NFs are included as a density-dependent effective nucleon-nucleon interaction which is obtained by summing up the third nucleon leg over the states in the Fermi sea. Here are some examples of applications for nuclear matter [89] and reaction theory [90, 91].

So far we have introduced the many chiral EFT interactions on the market and shown that there are many options to construct chiral interactions especially in the context of LECs. Another important topic on the input interaction is uncertainty in LECs and its propagation to physical quantities. The UQ become one of the hottest topic of (*ab initio*) nuclear theory. We should mention that there are pioneering studies along this line about the uncertainty analysis in the determination of LECs and uncertainty propagation [77, 92–97]. Those studies could drive us to see some link between fundamental interaction and nuclear many-body systems via LECs.

2.3.2 Similarity renormalization group (SRG)

The realistic potentials including chiral forces have strong repulsion in their short-range component, and this repulsion deteriorates the convergence of many-body calculations significantly. Note that chiral interactions are relatively softer than other realistic interactions such as Argonne potential [98]. In general, however, it is needed to soften the interaction, in other words to decouple low and high momentum components in a free-space interaction, while keeping its original information as much as possible. It is the current common choice for that purpose to evolve nuclear Hamiltonians using the renormalization group method which is known as similarity renormalization group (SRG).

The SRG was introduced independently by Głazek and Wilson [99] and Wegner [100]. It is based on unitary transformations which drive the initial Hamiltonian toward a more band- or block-diagonal form. The SRG transformation can be achieved in an energy-independent way and it preserves all observables as far as one keeps induced many-body forces in all orders.

The SRG flow is expressed by a unitary transformation $U(s)$ as,

$$H(s) = U(s)H(0)U^{\dagger}(s) \equiv T_{\rm rel} + V(s), \tag{2.22}$$

where $T_{\rm rel}$ is the relative kinetic energy to be independent of the flow parameter s and $H(s = 0) = T_{\rm rel} + V(s = 0)$ is the initial Hamiltonian, and $V(s)$ is the evolved potential. Taking the

derivative of above equation with respect to s, it follows

$$\frac{d}{ds}H(s) = [\eta(s), H(s)], \tag{2.23}$$

where $\eta(s)$ is anti-Hermitian generator defined as

$$\eta(s) \equiv \frac{dU(s)}{ds}U^{\dagger}(s) = -\eta^{\dagger}(s). \tag{2.24}$$

The unitary transformation above can be written as an s-ordered exponential [102]:

$$U(s) = \mathcal{T}_s \exp\left(\int_0^s ds'\eta(s')\right). \tag{2.25}$$

Here we take the simplest choice of the generator, which was firstly suggested by Wegner,

$$\eta_s = [H^d(s), H(s)] = [H^d(s), H^{od}(s)]. \tag{2.26}$$

where H^d and H^{od} are diagonal and off-diagonal part of the Hamiltonian. In this particular choice of the generator, it can be analytically shown that the off-diagonal part is exponentially driven to zero [100]. There are other alternatives for the generator. For example, White's generator [101] is known to be convenient for a specific problem, i.e. it is the common choice for in-medium SRG[5]. We refer the interested readers to Refs. [102–104] for more details on the generator choice. We note that it is a common convention to use $\lambda \equiv s^{-1/4}$, which has units of fm^{-1}, instead of s.

2.3.3 Derivations of a valence space Hamiltonian from a free-space interaction

In this subsection, we give a quick overview of derivations of effective interaction for valence CI calculations.

The history of deriving effective shell-model Hamiltonians from a free-space nucleon-nucleon interaction goes back to around 60's [105–109].

While phenomenological (data-driven) approaches to construct shell-model Hamiltonians have provided systematic descriptions of medium-mass nuclei with great success [110–118][6], many progresses have been achieved along the direction to derive it from a free-space NN

[5]Modified versions of White's generator are sometimes used.

[6]Some used the so-called G-matrix theory as its starting point and then modified it so as to improve the agreement with data.

force using the many-body perturbation theory (MBPT) [88, 119–132], the Coupled-Cluster method [133–135], the valence space in-medium similarity renormalization group (VS-IMSRG) [71, 136–138]. In the rest of this subsection, we will briefly introduce derivations of effective interactions by taking the VS-IMSRG as an example.

The in-medium similarity renormalization group (IM-SRG) was originally developed as an *ab intio* method [139] for ground state properties of (sub-shell) closed nuclei like the Coupled-Cluster [140, 141] and the unitary-model operator approach [142–144]. Early applications of these the so-called post Hartree-Fock (HF) methods were limited to ground state properties of (sub-shell) closed nuclei. However, these were extended to deal with other quantities such as Gamow-Teller transition strength and excited states [145–147]. Soon after the first publication, the IM-SRG was also applied to open-shell nuclei [148]. For these methods, several review articles are available now for the CC [149] and the IM-SRG [78, 102].

The key concept of the IM-SRG is the applications of the SRG to the nuclear many-body medium. As mentioned above, the SRG decouples, for example, low- and high-momentum components of the Hamiltonian. On the other hand, the IM-SRG decouples the ground state and the other excited states. To explain that, let us define some notations. In what follows, we work with a general form of A-body Hamiltonians truncated at three-body level:

$$H = \left(1 - \frac{1}{A}\right) T^{(1)} + T^{(2)} + V^{(2)} + V^{(3)}, \tag{2.27}$$

where $V^{(n)}$ is n-body force and kinetic terms are given as

$$T^{(1)} \equiv \sum_i \frac{\boldsymbol{p}_i^2}{2m}, \tag{2.28}$$

$$T^{(2)} \equiv -\sum_{i<j} \frac{\boldsymbol{p}_i \cdot \boldsymbol{p}_j}{mA}. \tag{2.29}$$

We can rewrite the Hamiltonian in second-quantized form:

$$H = \left(1 - \frac{1}{A}\right) \sum_{ij} T^{(1)}_{ij} a_i^\dagger a_j + \frac{1}{4} \sum_{ijkl} (T^{(2)}_{ijkl} + V^{(2)}_{ijkl}) a_i^\dagger a_j^\dagger a_k a_l + \frac{1}{36} \sum_{ijklmn} V^{(3)}_{ijklmn} a_i^\dagger a_j^\dagger a_k^\dagger a_l a_m a_n. \tag{2.30}$$

The operators above can be normal ordered with respect to a reference state $|\Phi\rangle$, which is obtained by e.g. the Hartree-Fock state:

$$H = E + \sum_{ij} f_{ij} : a_i^\dagger a_j : + \sum_{ijkl} \Gamma_{ijkl} : a_i^\dagger a_j^\dagger a_k a_l : + \sum_{ijklmn} W_{ijklmn} : a_i^\dagger a_j^\dagger a_k^\dagger a_l a_m a_n :, \tag{2.31}$$

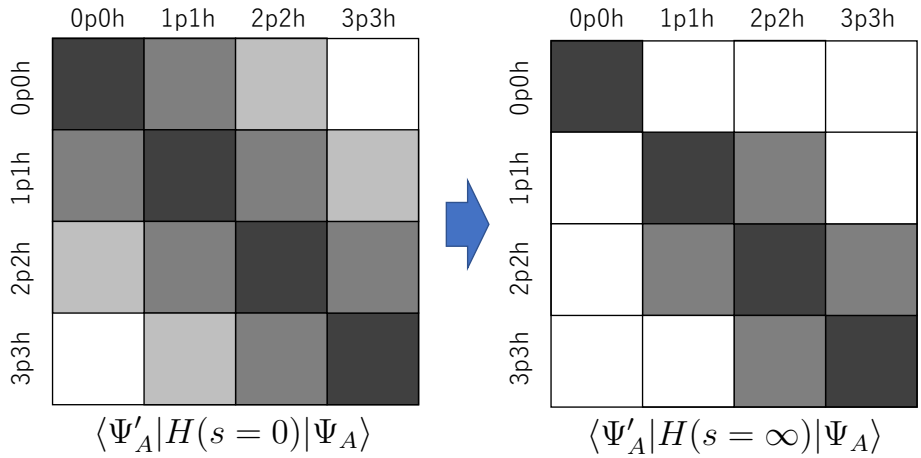

FIG. 2.3: Schematic figure of the IM-SRG flow. See the main text for more details.

where most of the normal ordered operators give $\langle\Phi| : a_i^\dagger \cdots a_j : |\Phi\rangle = 0$ and the coefficients of zero- to three-body terms are given in a simplified form:

$$E = \left(1 - \frac{1}{A}\right) \sum_i \langle i|T^{(1)}|i\rangle n_i + \frac{1}{2}\sum_{ij}\langle ij|T^{(2)} + V^{(2)}|ij\rangle n_i n_j + \frac{1}{6}\sum_{ijk}\langle ijk|V^{(3)}|ijk\rangle n_i n_j n_k, \tag{2.32}$$

$$f_{ij} = \left(1 - \frac{1}{A}\right) \langle i|T^{(1)}|j\rangle + \sum_k \langle ik|T^{(2)} + V^{(2)}|jk\rangle n_k + \frac{1}{2}\sum_{kl}\langle ikl|V^{(3)}|jkl\rangle n_k n_l, \tag{2.33}$$

$$\Gamma_{ijkl} = \langle ij|T^{(2)} + V^{(2)}|kl\rangle + \sum_m \langle ijm|V^{(3)}|klm\rangle n_m, \tag{2.34}$$

$$W_{ijklmn} = \langle ijk|V^{(3)}|lmn\rangle, \tag{2.35}$$

using occupation numbers in the reference state, $n_p = \theta(\epsilon_p - \epsilon_F)$. Here ϵ_F denotes the Fermi energy of the target nucleus. When the contribution of this term with W is relatively small, it is validated to neglect the third term in Eq. (2.31), which is known as the normal-ordered two-body (NO2B) approximation. Putting the normal-ordered Hamiltonian Eq. (2.31) into the IM-SRG flow equation,

$$\frac{dH(s)}{ds} = [\eta(s), H(s)] \tag{2.36}$$

with $\eta(s) = \eta^{(1)} + \eta^{(2)}$, which are one-body and two-body generators, we obtain the set of flow equations for E, f, Γ, and W. As in the free-space SRG, this flow equation is formulated as a commutator of operators, so this induces higher-order operators than two-body. It is the common choice to keep up to two-body level in the IM-SRG flow. The flow equation for

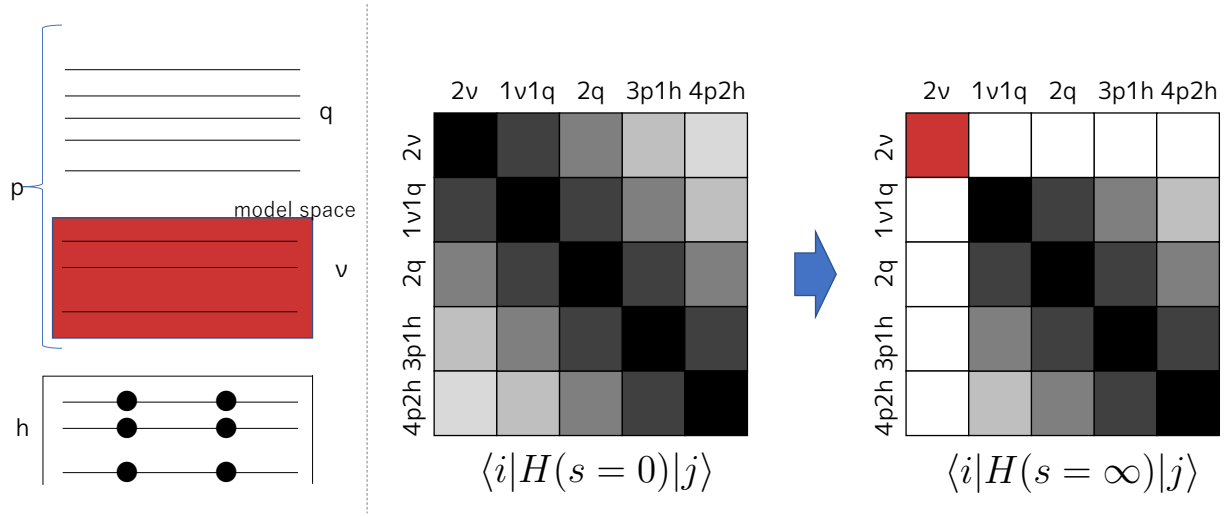

FIG. 2.4: Schematic figure of VS-IMSRG flow. The leftmost panel shows the separation of the single-particle space with hole (h), valence particle (ν), and non-valence particle (q).

zero-body term is written as

$$\frac{dE(s)}{ds} = \sum_{ij}(n_i - n_j)\eta^{(1)}_{ij} f_{ij} + \frac{1}{2}\sum_{ijkl}\eta^{(2)}_{ijkl}\Gamma_{ijkl} n_i n_j \bar{n}_k \bar{n}_l, \tag{2.37}$$

where $\bar{n}_k \equiv 1 - n_k$. As a consequence of the IM-SRG flow, the ground state is decoupled from other states in a non-perturbative way and the 0-body term gives the interacting ground state energy. The schematic figure is shown in FIG. 2.3. The grayscale figure shows the coupling among reference state (which is defined as the zero particle zero hole ($0p0h$) configuration), $1p1h$, $2p2h$, and $3p3h$. Under the given Hamiltonian truncated up to two-body level, the many-body states specified as $npnh$ can only couple to $(n \pm 2)$ particle-hole excitation. This is why the coupling between $0p0h$ and $3p3h$ is colored in white. As a sequence of the IM-SRG flow, the off-diagonal couplings between $npnh$ and $(n \pm 2)p(n \pm 2)h$ are driven to zero.

The expectation value of $H(s)$ gives better estimate of the ground state energy than HF:

$$E^{\mathrm{IMSRG}}_{g.s.} \equiv \lim_{s\to\infty}\langle\Phi|H(s)|\Phi\rangle \leq \langle\Phi|H(s=0)|\Phi\rangle \equiv E_{HF}. \tag{2.38}$$

The idea of the IM-SRG is naturally applied to decouple the valence space and others, and this enlarges our microscopic scope to doubly-open shell nuclei. We can receive a benefit of valence CI methods, i.e. one can systematically obtain ground state properties, excited states, electromagnetic transitions, and so on. Of course there still remains the problem of computational scaling of the valence CI method.

Let $|i\rangle$ and $|j\rangle$ denote some two-particle states in A-body system. Under this, we can draw

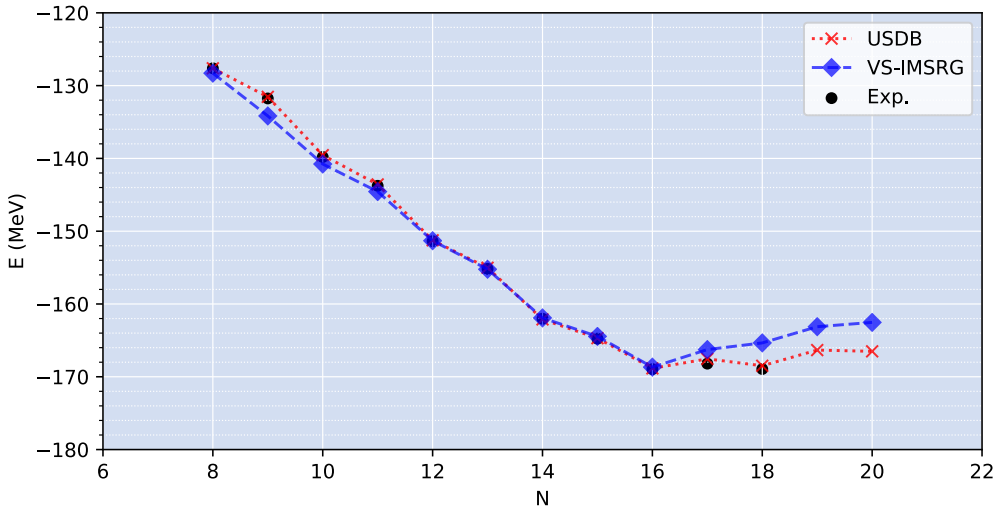

FIG. 2.5: Oxygen ground-state energies with the USDB interaction [115] and the VS-IMSRG interactions [138].

the schematic picture shown in FIG. 2.4. The initial Hamiltonian truncated up to two-body level can couple the two-body configurations 2ν, $1\nu 1q$, $2q$, $3p1h$, $4p2h$ to each other. Here h, ν, and q denote the hole, valence particle (ν), and non-valence particle (q), respectively. This includes the coupling between the model space, outer space, and hole states. The IM-SRG flow can be used to decouple these couplings and then to renormalize those effect into the effective Hamiltonian for the model space in a non-perturbative way. This is schematically shown in the rightmost panel of FIG. 2.4.

In FIG. 2.5, we show the ground-state energies of oxygen-isotopes as a function of neutron number N. The black dot points are experimental data, and the lines with cross and diamond symbols are calculated results using phenomenological USDB [115] and the VS-IMSRG interactions[7] [138] respectively. We must note that in the VS-IMSRG approach all contributions (from zero-body to two-body) to the energy eigenvalues are fully calculated with the IM-SRG calculation and the shell model while, in phenomenological shell-model calculations, the ground-state energy of the inert core is usually taken from a database. In that sense, the comparison of the agreement between them and data is not fair, and we must note that the VS-IMSRG results are remarkable and encouraging to go along these bottom-up approaches.

We have so far discussed ways to construct valence-space Hamiltonians using many-body methods. An alternative approach is the one discussed in Ref. [150], which directly link the chiral EFT potentials to the effective interactions for a valence space. In this work, they

[7]The reason why we used the plural form is that effective interactions are derived for each nucleus in the VS-IMSRG approach.

construct chiral interactions for a valence space, which is the *sd*-shell in the work, with center-of-mass-dependent operators that arise due to the Galilean invariance breaking by assuming the existence of the inert core.

All these are complemental studies to see the connection between the nuclear force and effective interactions for a valence space.

2.4 Inference of parameters

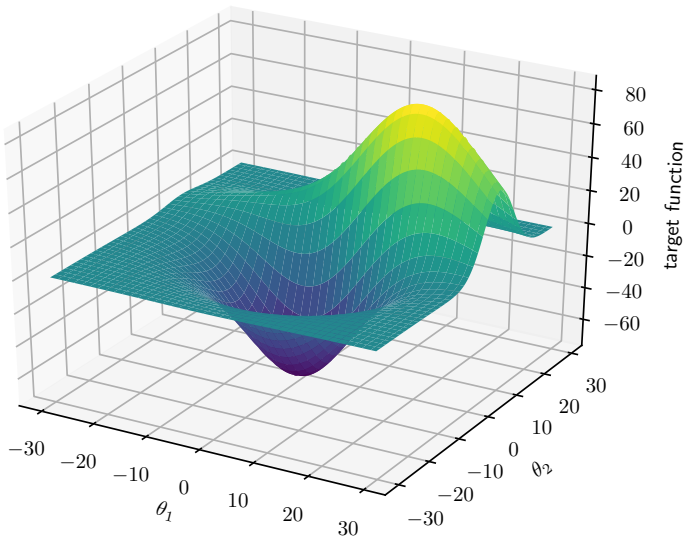

FIG. 2.6: Schematic picture of the potential surface of a target function.

In the rest of this chapter, we introduce machineries to quantify theoretical uncertainties in CI-type calculations and to extract much information from the results.

Before discussing technical details about the parameter inference, especially Bayesian inference, in nuclear many-body problems, let us make some important remarks from general viewpoints. Bayesian inference has been widely used in various scientific fields. We must note that Bayesian is not special at all, but it is quite natural. If one has a belief like "XX must be YY", this belief may be only the case with what he/she has already observed or known. The validity of beliefs is strengthen or weakened by objective evidence and then the beliefs are modified by them. This, updating posterior distributions by observing data, is the way of thinking for human beings.

In some modern scientific studies, what people are doing can be rephrased as *seeking some optical values*, *e.g.* the optimal values for parameters in a model, the optimal weight values of neurons in artificial neural networks, the optimal conditions for materials to realize some

special properties, and so on. In all those studies, the best value to maximize (minimize) some target function, say the cost function, is used. We must note that the best value is the best only for the data used and it is much more important to see how much narrow or broad are their target functions in the parameter space around the best value so that one can see the robustness of the *best* value.

The schematic picture, which is nothing to do with physics, is shown in FIG. 2.6. Suppose the z-axis means some target function to be minimized in two-dimensional parameter space (θ_1, θ_2). The best parameter is $(0, 0)$. The structure of the such target function, let us call potential surface, tells us the robustness of the argument and whether or not there are some directions in parameter space which is sensitive or insensitive to the target functions. This tells us whether or not one still needs additional data to constrain the theoretical model.

It is worth showing the structure of the potential surface more than finding a temporary optimal value, which will be modified by adding new data in future. This is especially true for nuclear physics because we cannot write down (at least at this moment) the fundamental interaction in a closed form and models for the nuclear force have free parameters[8].

For that purpose, it is a natural choice to regard the parameters in theory not as points but as probability distributions, and Bayesian machinery is the method of choice. Bayes' theorem is written as

$$P(\boldsymbol{\theta}|D) = \frac{P(D|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(D)} \propto P(D|\boldsymbol{\theta})P(\boldsymbol{\theta}), \tag{2.39}$$

where $\boldsymbol{\theta}$ is multidimensional parameter and D is the data set taken into account for parameter estimations. In most practical cases the marginal likelihood $P(D)$ is difficult to evaluate and regarded as a normalizing constant.

From here we use the following ordinary likelihood function unless otherwise mentioned:

$$P(D|\boldsymbol{\theta}) = \exp\left(-\chi^2(\boldsymbol{\theta})/2\right), \tag{2.40}$$

with the squared errors,

$$\chi^2(\theta) \equiv \sum_{n=1}^{N_D} \left(\frac{\mathcal{O}_n^{\rm exp} - < \mathcal{O}[\boldsymbol{\theta}] >_n}{\Delta\mathcal{O}} \right)^2, \tag{2.41}$$

where N_D is the number of data, $\mathcal{O}_n^{\rm exp}$ is experimental value, $< \mathcal{O}[\boldsymbol{\theta}] >_n$ is corresponding theoretical estimates at $\boldsymbol{\theta}$, and $\Delta\mathcal{O}$ is the adopted error of observables $\{\mathcal{O}_n\}$; which are determined by theoretical errors, since theoretical ones have the largest contributions in many

[8]The parametrization itself is of course not unique in general.

practical cases of nuclear structure physics. We must note that there is no unique choice for the denominator in Eq. (2.41), which can be dependent on n. How to parametrize the errors and how to account for correlations among the errors are still open questions to be studied in many scientific fields. Here we adopted the global error $\Delta\mathcal{O}$ over the entire data as the starting point of those comprehensive studies.

While we use non-informed prior $P(\boldsymbol{\theta}) \sim 1$ throughout this thesis, one can use the informed prior if the prior information is *plausible*. We also mention that if we use Gaussian prior like $P(\boldsymbol{\theta}) \sim \exp\left(-a\boldsymbol{\theta}^T \boldsymbol{M} \boldsymbol{\theta}\right)$, this is nothing but introducing L2-norm regularization term in the χ^2. The prior selection itself can be formulated in terms of Bayesian inference by hierarchical usage of Bayesian machinery. This flexibility is the one of the advantages of Bayesian methods.

2.5 How to evaluate posteriors?

Once the prior and the likelihood are determined, the problem is how to evaluate the posterior, because in general the posterior cannot be obtained in a closed form. We typically have two classes to evaluate the posterior $P(\boldsymbol{\theta}|D)$. One is a stochastic method such as Markov Chain Monte Carlo (MCMC) and the other one is an approximation scheme such as the Laplace approximation (LA) explained below.

If one could achieve infinite number of iterations for MCMC within a realistic time scale, one obtains the samples that are exactly obeying posterior distributions.

In the case of nuclear many-body problems taking account of strong correlations among many nucleons, one can naively expect that the true posterior is highly multimodal; The mappings between parameters and outputs of many-body calculations are expected to be highly nonlinear. This nonlinearity may prevent us from obtaining converged results by MCMC within realistic time scales.

Approximate inferences such as the LA are literally approximation methods for posteriors by writing down a posterior in a simple and closed form such as a Gaussian. Therefore, computational costs are much less than MCMC. However, in general, they never give samples obeying the true posteriors. The advantages and disadvantages for these two classes are complementary.

2.5.1 Stochastic method: Markov Chain Monte Carlo

In this section, we briefly review a few popular Markov Chain Monte Carlo methods from practical perspectives. More comprehensive explanation can be found in e.g. [151].

Random-walk Metropolis-Hastings (RWMH) method

The random-walk Metropolis-Hastings (RWMH) algorithm is highly popular sampling method due to its simplicity and applicability. In RWMH algorithms, the samples obeying a target distribution are generated through updating the state, which is represented by $\boldsymbol{\theta}$ hereafter.

The updating procedure is the following:

- When the current state is $\boldsymbol{\theta}$, a new parameter $\boldsymbol{\theta}'$ is proposed with the transition probability $t(\boldsymbol{\theta} \to \boldsymbol{\theta}')$.
- Calculate the Hastings ratio:

$$r(\boldsymbol{\theta}, \boldsymbol{\theta}') \equiv \frac{t(\boldsymbol{\theta}' \to \boldsymbol{\theta}) P(\boldsymbol{\theta}')}{t(\boldsymbol{\theta} \to \boldsymbol{\theta}') P(\boldsymbol{\theta})}. \tag{2.42}$$

- Accept the proposed $\boldsymbol{\theta}'$ with the probability:

$$r = min\{1, r(\boldsymbol{\theta}, \boldsymbol{\theta}')\} \tag{2.43}$$

The efficiency of sampling crucially depends on the scaling of the proposal distributions. If the variance of the proposal is too small, samples easily get stuck in local minima. On the other hand, if the variance is too large, the sampling can become inefficient. It is one of the most important issues in MCMC to design the appropriate proposal distributions. The highly expertise and many trial and errors are needed to achieve that. Of course there are several techniques to design the proposal distributions more or less automatically such as adaptive MCMC [152].

However, we note that naive usage of adaptive MCMC may get stuck in higher dimension or in the case of multimodal target functions. From the author's experience, if we apply adaptive design of proposal distributions in 17 dimensional space, which is the number of valence shell-model effective interactions of p-shell space to be considered in Chapter 3, several component of covariance matrix of the Gaussian proposal become almost zero to acquire the desired acceptance ratio. This leads to biased samplings.

Hamiltonian Monte Carlo

The Hamiltonian Monte Carlo (originally known as Hybrid Monte Carlo in lattice field theory [153]) algorithm is known as the way to avoid redundancy in random-walk Metropolis-

Hastings algorithm. This has been widely used in various fields and many packages are available, see e.g. Stan [154] and PyMC [155]. We also refer the interested reader to the review by Radford M. Neal [156].

The key idea of HMC is to use Hamiltonian dynamics to propose an update of the state. We introduce momentum $\boldsymbol{p}$ following standard normal distribution,

$$f(\boldsymbol{p}) = \prod_{i=1}^{d} f(p_i), \tag{2.44}$$

$$f(p_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{p_i^2}{2}\right), \tag{2.45}$$

where let d denote the dimension of the parameter $\boldsymbol{\theta}$ and momenta $\boldsymbol{p}$. We consider the marginal probability density of $\boldsymbol{\theta}$ and $\boldsymbol{p}$:

$$f(\boldsymbol{\theta}, \boldsymbol{p}|D) = f(\boldsymbol{\theta}|D) f(\boldsymbol{p}), \tag{2.46}$$

where $\boldsymbol{\theta}$ is the parameter of interest and D is the data as the previous section. To employ standard normal distribution corresponds to considering Hamiltonian dynamics of $m = 1$ particle. We can rewrite the above equation as

$$\begin{aligned} f(\boldsymbol{\theta}, \boldsymbol{p}|D) &= \exp\left(\log f(\boldsymbol{\theta}, \boldsymbol{p}|D)\right), \\ &\propto \exp\left(\log f(\boldsymbol{\theta}|D) - \sum_i p_i^2/2\right). \end{aligned} \tag{2.47}$$

By using $U(\boldsymbol{\theta}) \equiv -\log f(\boldsymbol{\theta}|D)$, we obtain the familiar form $H(\boldsymbol{p}, \boldsymbol{\theta}) \equiv U(\boldsymbol{\theta}) + \sum_i p_i^2/2$. The logarithm of posterior distribution can be regarded as the potential of Hamiltonian $H(\boldsymbol{p}, \boldsymbol{\theta})$. The marginal probability density is given as follows:

$$f(\boldsymbol{\theta}, \boldsymbol{p}|D) = \exp\left(-H(\boldsymbol{p}, \boldsymbol{\theta})\right). \tag{2.48}$$

Next, let us consider the updating procedure of the parameter $\boldsymbol{\theta}$ using Hamiltonian dynamics in the $(\boldsymbol{p}, \boldsymbol{\theta})$ phase-space. By introducing the time τ, Hamilton equation is given as

$$\frac{dp_i(\tau)}{d\tau} = -\frac{\partial H(\boldsymbol{p}, \boldsymbol{\theta})}{\partial q_i} = -\frac{\partial U(\boldsymbol{\theta})}{\partial q_i}, \tag{2.49}$$

$$\frac{d\theta_i(\tau)}{d\tau} = \frac{\partial H(\boldsymbol{p}, \boldsymbol{\theta})}{\partial p_i} = p_i, \tag{2.50}$$

for $i = 1, 2, ..., d$. We have so far considered the case with *independent* momenta. However, one

can employ the correlated momenta by replacing $f(\boldsymbol{p})$ by a multivariate normal distribution

$$f(\boldsymbol{p}) \propto \exp\left(-\frac{1}{2}\boldsymbol{p}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{p}\right). \tag{2.51}$$

Our choice corresponds to its special case, $\boldsymbol{\Sigma} = I_d$.

In practical application of HMC, one has to solve the Hamiltonian dynamics in a numerical way, and this induces numerical errors. This deteriorate the acceptance ratio, i.e. sampling efficiency of HMC algorithm. The method of choice is to employ the so-called *leapfrog* method:

$$\boldsymbol{p}(\tau + 1/2) = \boldsymbol{p}(\tau) - \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}} U(\boldsymbol{\theta}(\tau)), \tag{2.52}$$

$$\boldsymbol{\theta}(\tau + 1) = \boldsymbol{\theta}(\tau) + \epsilon\boldsymbol{p}(\tau + 1/2), \tag{2.53}$$

$$\boldsymbol{p}(\tau + 1) = \boldsymbol{p}(\tau + 1/2) - \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}} U(\boldsymbol{\theta}(\tau + 1/2)), \tag{2.54}$$

where the time step for one iteraction is represented as 1 and ϵ is the infinitesimal constant. After L steps, we propose the updated parameter $\boldsymbol{\theta}' \equiv \boldsymbol{\theta}(\tau + L)$. Then, we calculate the quantity, $H(\boldsymbol{p}(\tau + L), \boldsymbol{\theta}') - H(\boldsymbol{p}(\tau), \boldsymbol{\theta})$. In principle, this must be 0 due to the Hamiltonian conservation, but, in numerical level, there is finite numerical error due to the finite ϵ to describe time evolution in phase space. The Hastings ratio in this case becomes

$$r = \frac{t(\boldsymbol{\theta}, \boldsymbol{p}|\boldsymbol{\theta}', \boldsymbol{p}')f(\boldsymbol{\theta}', \boldsymbol{p}'|D)}{t(\boldsymbol{\theta}', \boldsymbol{p}'|\boldsymbol{\theta}, \boldsymbol{p})f(\boldsymbol{\theta}, \boldsymbol{p}|D)} \tag{2.55}$$

In the Hamiltonian dynamics, the time evolution is reversible, i.e. $t(\boldsymbol{\theta}, \boldsymbol{p}|\boldsymbol{\theta}', \boldsymbol{p}') = t(\boldsymbol{\theta}', \boldsymbol{p}'|\boldsymbol{\theta}, \boldsymbol{p})$. Then, the acceptance rate is given as

$$r(\boldsymbol{\theta}, \boldsymbol{\theta}') = min\{1, \exp\{-(H(\boldsymbol{p}', \boldsymbol{\theta}') - H(\boldsymbol{p}, \boldsymbol{\theta}))\}\}. \tag{2.56}$$

One can see from the above equation, the acceptance ratio of HMC algorithm is always high if one keeps the numerical accuracy of Hamiltonian dynamics using e.g. the leapfrog method.

Here let us move to more detailed formulations which are relevant to CI-type calculations. We use the ordinary likelihood again:

$$f(D|\boldsymbol{\theta}) = \exp\left(-\chi^2/2\right), \tag{2.57}$$

$$\chi^2(\boldsymbol{\theta}) = \sum_i^{N_D} \frac{(E_i^{\text{exp.}} - E_i^{\text{theo.}}(\boldsymbol{\theta}))^2}{(\Delta E)^2} \tag{2.58}$$

where i runs over selected nuclei and states, and ΔE is arbitrary given accuracy, which can be dependent on the label of the states, i. Now we assume prior distribution $f(\boldsymbol{\theta})$ as uniform

distribution and $f(D)$ can be regarded as the normalization factor. Thus, $U(\boldsymbol{\theta})$ in HMC algorithm is now identical with $\chi^2/2$ in problem of interest. We also note that we restrict ourselves to consider energy eigenvalues from now on[9].

In the application of HMC, the main problem is how to evaluate gradients of $U(\boldsymbol{\theta})$ with respect to parameters $\boldsymbol{\theta}$. Besides its importance for HMC, the gradients are also needed in some approximation schemes for the posterior. In the rest part of this subsection, we explain the concrete procedure to obtain the gradient in CI-type calculations.

We start with a general shell-model Hamiltonian up to two-body level:

$$H_{\rm SM} = \sum_i \epsilon_i a_i^\dagger a_i + \frac{1}{4}\sum_{ijkl} \bar{v}_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \equiv S + V, \tag{2.59}$$

where $\bar{v}$ is anti-symmetric two-body interaction. We define the shorthand of the uncoupled two-nucleons states:

$$|ab\rangle \equiv \frac{1}{\sqrt{2}}(|j_a m_a, j_b m_b\rangle - |j_b m_b, j_a m_a\rangle), \tag{2.60}$$

where the factor $1/\sqrt{2}$ is for normalization. We from now on use two-different expressions of TBMEs which are useful as inputs of shell-model codes. One is the JT-coupled form, which is called isospin-formalism, with explicit isospin symmetry and the other one is neutron-proton (np) formalism. Traditionally, phenomenological shell-model effective interactions are given in isospin-formalism. We note that the SPEs are common between neutrons and protons in isospin-formalism and one has to make sure to take account of derivatives with respect to SPEs for both protons and neutrons in considering the summation of derivatives. On the other hand, when one uses derived interactions from chiral potentials, which includes Coulomb force explicitly, one must work in np-formalism.

To obtain the two-body matrix element in np-formalism, let us define the following two-nucleons state coupled to J and M:

$$|ab; JM\rangle = \mathscr{N}_{ab}(J) \sum_{m_a m_b} (j_a j_b m_a m_b | JM) |ab\rangle. \tag{2.61}$$

[9]It is easier to evaluate the derivatives of energy eigenvalues than other observables.

Here the normalization constant is evaluated by considering the norm of the above equation

$$
\begin{aligned}
1 =& < ab; JM|ab; JM\rangle \qquad (2.62)\\
=&[\mathcal{N}_{ab}(J)]^2 \sum_{m_a m_b m'_a m'_b} (\delta_{m_a m'_a}\delta_{m_b m'_b} - \delta_{j_a j_b}\delta_{m_a m'_b}\delta_{m_b m'_a})(j_a j_b m_a m_b|JM)(j_a j_b m'_a m'_b|JM),\\
=&[\mathcal{N}_{ab}(J)]^2 \sum_{m_a m_b} ((j_a j_b m_a m_b|JM)(j_a j_b m_a m_b|JM) - \delta_{j_a j_b}(j_a j_b m_a m_b|JM)(j_a j_b m_b m_a|JM)),\\
=&[\mathcal{N}_{ab}(J)]^2 \sum_{m_a m_b} (j_a j_b m_a m_b|JM)(j_a j_b m_a m_b|JM)(1 + \delta_{j_a j_b}(-1)^J),\\
=&[\mathcal{N}_{ab}(J)]^2(1 + \delta_{j_a j_b}(-1)^J). \qquad (2.63)
\end{aligned}
$$

Then the normalization factor for non-vanishing angular momentum coupled state is given as $1/\sqrt{1+\delta_{j_a j_b}}$. Inversely, the uncoupled two-nucleons state is also expressed as the superposition of coupled two-nucleons states as

$$|ab\rangle = \sum_{JM}(j_a j_b m_a m_b|JM)[\mathcal{N}_{ab}(J)]^{-1}|ab; JM\rangle. \qquad (2.64)$$

We have used the orthogonal relations of Clebsch-Gordan coefficients:

$$\sum_{JM}(j_1 j_2 m_1 m_2|JM)(j_1 j_2 m'_1 m'_2|JM) = \delta_{m_1 m'_1}\delta_{m_2 m'_2}, \qquad (2.65)$$

$$\sum_{m_1 m_2}(j_1 j_2 m_1 m_2|JM)(j_1 j_2 m_1 m_2|J'M') = \delta_{JJ'}\delta_{MM'}. \qquad (2.66)$$

In this thesis, we followed the notations in Ref. [157] by Prof. Arima and Prof. Ichimura[10].

Under these, the two-body part of the shell-model Hamiltonian Eq. (2.59) is represented as

$$
\begin{aligned}
V =& \frac{1}{4}\sum_{abcd}\sum_{JM}\sum_{J'M'}(j_a j_b m_a m_b|JM)(j_c j_d m_c m_d|J'M')\\
&\times [\mathcal{N}_{ab}(J)]^{-1}[\mathcal{N}_{cd}(J')]^{-1}\langle ab; JM|V|cd; J'M'\rangle a^\dagger_a a^\dagger_b a_c a_d, \qquad (2.67)\\
=& \frac{1}{4}\sum_{abcd}\sum_{JM}\sum_{J'M'}[\mathcal{N}_{ab}(J)]^{-1}[\mathcal{N}_{cd}(J')]^{-1}\langle ab; JM|V|cd; J'M'\rangle\\
&\times (-1)^{J'-M'}[a^\dagger_a \times a^\dagger_b]^{(J)}_M [a_c \times a_d]^{(J')}_{-M'}, \qquad (2.68)\\
& \qquad (2.69)
\end{aligned}
$$

[10]This book (in Japanese) has been used for a long time as a formulary in nuclear structure physics group of the university of Tokyo.

where we used Eqs. (2.65) and (2.66) and operators in square brackets are defined as

$$[a_a^\dagger \times a_b^\dagger]_M^{(J)} = \sum_{m_\alpha m_\beta} (j_a j_b m_\alpha m_\beta | JM) a_a^\dagger a_b^\dagger, \tag{2.70}$$

$$[a_c \times a_d]_{M'}^{(J')} = \sum_{m_\gamma m_\delta} (j_c j_d m_\gamma m_\delta | J'M') a_c a_d = [\tilde{a}_c \times \tilde{a}_d]_{-M'}^{(J')}, \tag{2.71}$$

with $\tilde{a}_c \equiv (-1)^{j_c - m_c} a_c$. Using the Wigner-Eckart theorem and the fact that V must be scaler in terms of spherical tensor operators, the two-body matrix elements are rewritten as follows:

$$\langle ab; JM|V|cd; J'M'\rangle = \frac{(J'0M'0|JM)}{\sqrt{2J+1}} \langle ab; J||V||cd; J'\rangle = \frac{\delta_{J'J}\delta_{M'M}}{\sqrt{2J+1}} \langle ab; J||V||cd; J'\rangle, \tag{2.72}$$

where $\langle ||\cdot|| \rangle$ is the reduced matrix element. For rank-k spherical tensor operators, the following relation holds:

$$\frac{(-1)^k}{\sqrt{2k+1}} \sum_q (-1)^q U_q^{(k)} V_{-q}^{(k)} = [U^{(k)} \times V^{(k)}]_0^{(0)}. \tag{2.73}$$

Then, Eq. (2.68) becomes

$$V = \frac{1}{4} \sum_{abcd} \sum_J [\mathcal{N}_{ab}(J)]^{-1} [\mathcal{N}_{cd}(J)]^{-1} \langle ab; J||V||cd; J\rangle \left[[a_a^\dagger \times a_b^\dagger]^{(J)} \times [\tilde{a}_c \times \tilde{a}_d]^{(J)} \right]_0^{(0)}. \tag{2.74}$$

Dividing the each term of V into the coefficient $\langle ab; J||V||cd; J\rangle$ and the operator part, the expectation value $< V >$ is represented in a form $< V >= \sum_\alpha \langle ab; J||V||cd; J\rangle < [\cdot]_0^{(0)} >$. Now let α denote (a, b, c, d, J), and we define the TBMEs in np-formalism as $V_\alpha^{np} \equiv \langle ab; J||V||cd; J\rangle$. We note that in np-formalism indices (a, b, c, d) include the labels of proton and neutron.

When the two-body interaction is isoscalar too, one can obtain the corresponding expressions of V in isospin-formalism by extending the operators:

$$V = \frac{1}{4} \sum_{abcdJT} [\mathcal{N}_{ab}(JT]^{-1} [\mathcal{N}_{cd}(JT)]^{-1} \langle ab; JT||V||cd; JT\rangle \left[[a_a^\dagger \times a_b^\dagger]^{(J,T)} \times [\tilde{a}_c \times \tilde{a}_d]^{(J,T)} \right]_{0,0}^{(0,0)}. \tag{2.75}$$

We refer to the reduced matrix element $\langle ab; JT||V||cd; JT\rangle \equiv V_\alpha^{iso}$ as TBMEs in isospin-formalism. In this case, the greek index α denotes (a, b, c, d, J, T).

The gradients of the potential $U(\boldsymbol{\theta})$ with respect to TBMEs in *np*-formalism are given as

$$\begin{aligned}\frac{\partial U}{\partial V_\alpha^{np}} &= \frac{\partial}{\partial V_\alpha^{np}}(-\log f(\boldsymbol{\theta}|D)),\\ &= -\frac{\partial}{\partial V_\alpha^{np}}(\log f(D|\boldsymbol{\theta}) + \log f(\boldsymbol{\theta})),\\ &= -\frac{\partial}{\partial V_\alpha^{np}}(\log f(D|\boldsymbol{\theta})) = \frac{\partial}{\partial V_\alpha^{np}}\frac{\chi^2}{2},\\ &= -\sum_i^{N_D}\left[\frac{(E_i^{\text{exp.}} - (\sum_\beta \langle S_\beta\rangle + \sum_\beta \langle V_\beta^{np}\rangle_i)}{(\Delta E)^2}\langle V_\alpha^{np}\rangle_i\right]. \end{aligned} \tag{2.76}$$

In the forth line, we used Hellmann-Feynman theorem. It is convenient to work in *np*-formalism in terms of codes and derivatives of the quantity. However, it is much more convenient to work in isospin-formalism when it comes to update the parameters. This is simply because the number of parameters in isospin-formalism is less than that of *np*-formalism.

Relations of TBMEs between proton-neutron formalism isospin-formalism is given by

$$V^{np}(p_a p_b; p_c p_d; J) = V^{np}(n_a n_b; n_c n_d; J) = V^{iso}_{J,T=1}(ab; cd), \tag{2.77}$$

$$\begin{aligned} V^{np}(p_a n_b; p_c n_d; J) = \frac{1}{2}\Big\{&\sqrt{(1+(-1)^J\delta_{ab})(1+(-1)^J\delta_{cd})}V^{iso}_{J,T=1}(ab;cd)\\ &+ \sqrt{(1-(-1)^J\delta_{ab})(1-(-1)^J\delta_{cd})}V^{iso}_{J,T=0}(ab;cd)\Big\}, \end{aligned} \tag{2.78}$$

$$\equiv c_0 V^{iso}_{J,T=0}(ab;cd) + c_1 V^{iso}_{J,T=1}(ab;cd), \tag{2.79}$$

where p and n represent proton and neutron.

Note that the above relations are true only when TBMEs in neutron-proton formalism have good isospin symmetry. We can use the above equations to construct *np*-formalism interaction having isospin symmetry if only you have TBMEs in isospin formalism.

By considering infinitesimal deviations of the function $f(V)$, we obtain

$$\begin{aligned}&\left\{\frac{\partial f}{\partial V^{iso}_{\alpha,T=0}} - c_0\frac{\partial f}{\partial V^{np}(\pi\nu)}\right\}dV^{iso}_{J,T=0}(ab;cd)\\ &\quad+\left\{\frac{\partial f}{\partial V^{iso}_{\alpha,T=1}} - \frac{\partial f}{\partial V^{np}(\pi\pi)} - \frac{\partial f}{\partial V^{np}(\nu\nu)} - c_1\frac{\partial f}{\partial V^{np}(\pi\nu)}\right\}dV^{iso}_{J,T=1}(ab;cd) = 0, \end{aligned} \tag{2.80}$$

where we assumed that infinitesimal variations in neutron-proton formalism are done simultaneously $dV^{np}(\nu\nu) = dV^{np}(\pi\nu)$ and we used $\pi\pi, \nu\nu$, and $\pi\nu$ to represent proton-proton, neutron-neutron, and proton-neutron interactions, respectively. Then, we obtain the rela-

tions:

$$\frac{\partial f}{\partial V^{iso}_{\alpha,T=0}} = c_0 \frac{\partial f}{\partial V^{np}(\pi\nu)}, \tag{2.81}$$

$$\frac{\partial f}{\partial V^{iso}_{\alpha,T=1}} = \frac{\partial f}{\partial V^{np}(\pi\pi)} + \frac{\partial f}{\partial V^{np}(\nu\nu)} + c_1 \frac{\partial f}{\partial V^{np}(\pi\nu)}. \tag{2.82}$$

We have shown so far the way to calculate gradients in shell-model calculations. This can be also used in approximate method with Laplace approximation for posterior.

2.5.2 Approximate inference

In the previous subsection, we introduced widely used MCMC methods to obtain samples from the posterior distribution. As already mentioned at the beginning of this section, the MCMC is computationally demanding.

We again start with the non-informed prior $P(\boldsymbol{\theta}) \propto 1$ and the ordinary likelihood function,

$$P(D|\boldsymbol{\theta}) = \exp\left(-\chi^2(\boldsymbol{\theta})/2\right), \tag{2.83}$$

with the squared errors,

$$\chi^2(\theta) \equiv \sum_{n=1}^{N_D} \left(\frac{E_n^{\rm exp} - < E[\boldsymbol{\theta}] >_n}{\Delta E} \right)^2, \tag{2.84}$$

where N_D is the number of data, $E_n^{\rm exp}$ is experimental value, $< E[\boldsymbol{\theta}] >_n$ is corresponding theoretical estimates with the parameter $\boldsymbol{\theta}$, and ΔE is the adopted error of the energy eigenvalues. This ΔE, which could be dependent on the label n, is determined by the theoretical error, since theoretical ones have the largest contributions in many practical cases. It should be noted again that E can be extended to any type of observables and we restrict ourselves to consider only energy eigenvalues for simplicity.

In the Laplace approximation (LA), the posterior is approximated by the multivariate Gaussian distribution around Maximum A posteriori (MAP):

$$\boldsymbol{\theta}_{\rm MAP} \equiv \arg\max_{\boldsymbol{\theta}} P(D|\boldsymbol{\theta}) P(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}} \chi(\boldsymbol{\theta}), \tag{2.85}$$

$$\begin{aligned} P(\boldsymbol{\theta}|D) &\simeq \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_{\rm MAP}, \boldsymbol{A}^{-1}), \\ &= \sqrt{\frac{|\boldsymbol{A}|}{(2\pi)^k}} \exp\left(-\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}_{\rm MAP})^T \boldsymbol{A} (\boldsymbol{\theta} - \boldsymbol{\theta}_{\rm MAP}) \right), \end{aligned} \tag{2.86}$$

with the Hessian matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = -\nabla\nabla \log P(D|\boldsymbol{\theta})P(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\mathrm{MAP}}}. \tag{2.87}$$

Here we used Eqs. (2.39) and (2.83) and $P(\boldsymbol{\theta}) \propto 1$. Each element of the Hessian matrix $\boldsymbol{A}$ is written as

$$A_{ij} = \sum_{n=1}^{N_D} \frac{1}{(\Delta E)^2} \frac{\partial\langle E_n[\boldsymbol{\theta}]\rangle}{\partial\theta_i} \frac{\partial\langle E_n[\boldsymbol{\theta}]\rangle}{\partial\theta_j} + \sum_{n=1}^{N_D} \frac{(\langle E_n[\boldsymbol{\theta}]\rangle - E_n^{\mathrm{exp}})}{(\Delta E)^2} \frac{\partial\langle E_n[\boldsymbol{\theta}]\rangle}{\partial\theta_i\partial\theta_j}. \tag{2.88}$$

The second derivative term above is evaluated by using finite differences of first derivative terms [158]:

$$\frac{\partial\langle E_n[\boldsymbol{\theta}]\rangle}{\partial\theta_i\partial\theta_j} = \frac{1}{2\epsilon}\left\{\frac{\partial\langle E_n[\theta_j+\epsilon]\rangle}{\partial\theta_i} - \frac{\partial\langle E_n[\theta_j-\epsilon]\rangle}{\partial\theta_i}\right\} + \mathcal{O}(\epsilon^2), \tag{2.89}$$

where $E_n[\theta_j \pm \epsilon]$ denotes the E_n value evaluated by the parameters whose jth components are slightly shifted by ϵ from $\boldsymbol{\theta}_{\mathrm{MAP}}$. In the practical applications in this thesis, we varied this ϵ from 1.e-5 to 1.e-4 and confirmed that the results are independent of the choice. Under this Eq. (2.89), the calculation cost of the second derivative term is $2d$ times larger than that of the first term of Eq. (2.88).

Chapter 3

Uncertainty quantification in valence shell-model calculations

In this chapter, we discuss a systematic way to quantify theoretical uncertainties in valence shell-model calculations, and how these information would help us to understand *exotic* properties in nuclear many-body systems.

The valence CI calculation, which is commonly called shell-model calculation in nuclear physics community, has been providing successful and systematic descriptions of a wide variety of properties of light- to medium-mass nuclei. This indicates that the shell-model calculations well approximates the wave functions of nuclei. See the reviews of shell-model calculations especially from phenomenological viewpoints, [159–161]. As already discussed in the previous chapter, also from microscopic viewpoints, the shell model plays a key role with recent developments in nuclear potentials by chiral EFT and many-body methods to derive effective interactions for a physically motivated model space. We refer the interested readers to the references cited in the previous chapter and the very recent review [162] for more details. This kind of study combining *ab initio* methods and shell-model calculations can act as a foothold for better understandings of the nuclear potential or pinning down the LECs, which still have relatively large uncertainties.

Under this circumstance, it is an urgent task to assess the validity of valence CI itself through evaluating their uncertainties which stems from the input effective interactions[1]. This is because by performing detailed analyses of uncertainties from the inputs, one can obtain deeper understandings on such as origins of the discrepancies between theory and experiments, or which observables could be and should be reproduced by the theory within a given model space.

[1] Uncertainties from many-body problem solvers are discussed in Chapter 4.

3.1 Inference of shell-model effective interactions

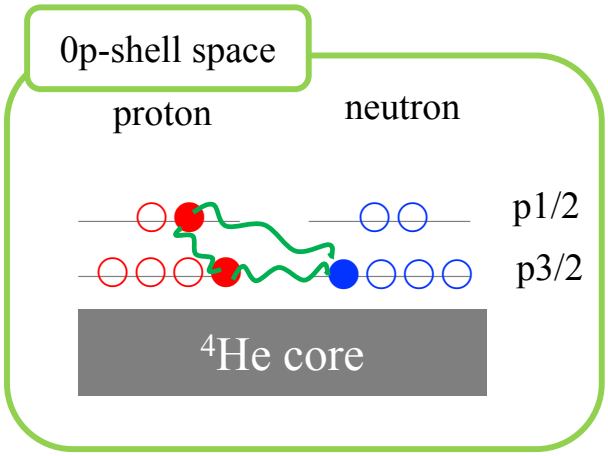

FIG. 3.1: Schematic figure of $0p$-shell space on top of the ^{4}He core.

In what follows, we introduce a way to quantify uncertainties in valence CI results so that one can make more insightful comparisons than what is schematically shown in FIG. 1.1.

We take the $0p$-shell space on top of the ^{4}He core as the model space. As shown in FIG. 3.1, the $0p$ shell consists of two orbitals $0p_{1/2}$ and $0p_{3/2}$ and these are abbreviated as $p1/2$ and $p3/2$ in this chapter. The orbitals above the $0p$ shell such as the $1s0d$ shell and the $1p0f$ shell are not included in the model space, i.e. the excitations from $0p$ shell to those shells are not allowed in the calculations. Under the isospin symmetry, $0p$-shell interactions consist of two single particle energies and fifteen two-body matrix elements, i.e. 17 parameters in total.

In this $0p$-shell region, there is an well-known phenomenological interaction by Cohen and Kurath (CK) [110]. Since they did not introduced mass-dependence of TBMEs, which turned out to improve the fit after CK's work, and there is a few updated experimental data in this region, we also re-examine the best fit of $0p$-shell interaction below. In addition to this, we do inference of valence effective interactions on the $0p$-shell space, then properly propagate the uncertainties to the observables.

All calculations within the $0p$-shell space can be done by an exact diagonalization method (using the Lanczos method) with an ordinary laptop. This means that there is no uncertainty from the many-body calculation itself. Thus, we can extract a theoretical uncertainty coming merely from the input parameters. In this sense, the valence shell-model calculation is appropriate test ground and starting points for uncertainty estimates in nuclear many-body problems.

In order to do inference of the parameters, we need to introduce the measure of *goodness* of the effective interactions. Following literatures, we use the squared errors shown in Eq. (2.84).

To compare experimental binding energies with shell-model results, the Coulomb corrections to energy values are added to $E_n^{\rm th}[\boldsymbol{\theta}]$ in accordance with Ref. [110]. We again show the definition of the chi-square deviations in Eq. (2.41):

$$\chi^2(\theta) \equiv \sum_{n=1}^{N_D} \left(\frac{\mathcal{O}_n^{\rm exp} - < \mathcal{O}[\boldsymbol{\theta}] >_n}{\Delta\mathcal{O}} \right)^2 . \tag{3.1}$$

For $\{E_n^{\rm exp}\}$, we use a fixed 33 energy values of ground and excited states of the $0p$-shell nuclei[2] throughout this thesis. This data set is essentially the same one used by CK with the exception of a few updated data. The entire dataset is summarized in the tables shown in Appendix A.

We should mention that our choice of the measure of the goodness is not unique. An additional possible term which could be added to the squared error is the so-called regularization term such as L1-norm or L2-norm. An alternative choice is to consider some informed prior. For example, one can evaluate the mean and variance of effective interactions derived from different input nuclear potentials, and use them as e.g. a Gaussian prior. Of course there is no obvious reason to take those interactions on an equal footing.

Now the problem is to evaluate the posterior distribution of the $0p$-shell interaction over the seventeen-dimensional parameter space. In general the posterior cannot be evaluated analytically. We employ the Laplace approximation, which was introduced in Sec. 2.5.2, to evaluate the posterior. Our task is to evaluate the Hessian matrix in Eq. (2.88).

If we write down a shell-model Hamiltonian in form of

$$H = \sum_i S_i a_i^\dagger a_i + \sum_{ijkl} V_{ijkl} a_i^\dagger a_j^\dagger a_k a_l, \tag{3.2}$$

where indices (i, j, k, l) run all single particle states in the model space and S and V are respectively SPE and TBME. The energy eigenvalues can be represented as

$$E \equiv \langle H \rangle = \sum_i S_i \langle a_i^\dagger a_i \rangle + \sum_{ijkl} V_{ijkl} \langle a_i^\dagger a_j^\dagger a_k a_l \rangle. \tag{3.3}$$

Thanks to the Hellmann–Feynman theorem, this factorization indicates that there is no additional computation to evaluate first derivative terms in the Hessian matrix, Eq. (2.88), as far as $\langle a_i^\dagger a_i \rangle$ and $\langle a_i^\dagger a_j^\dagger a_k a_l \rangle$ are computed. The first derivative terms of the Hessian matrix can be evaluated using the expectation values of the operators.

The concrete procedures to quantify the uncertainties in shell-model calculations by the

[2]The notation "XX-shell nuclei" simply means that the naive filling configuration is given in the XX-shell space. That does not necessarily means that additional excitations beyond the XX-shell are irrelevant.

LA are the followings:

a) to search the optimal interaction $\boldsymbol{\theta}_{\rm MAP}$ in Eq. (2.85) with respect to the given data set

b) to calculate the Hessian matrix in Eq. (2.88)

c) to generate sufficient number of effective interactions and to diagonalize them

For a), we prefixed the optimal interaction for the dataset by Adam [163], which is famous as an optimization method in the machine learning community. We introduced the mass-dependence of the form $(A/6)^p$ on two-body matrix elements (TBMEs) and we optimized this p to minimize the total root mean square errors by changing the p by 0.1. This scaling of TBMEs is originally introduced in Ref. [111] for the universal *sd*-shell interaction and accounts for the mass dependence of the radial wave functions more or less. This significantly improves the quality of the effective interaction. The root mean square errors for the given data set are summarized in Tab. 3.1 in comparison with the results for three sets of the interaction proposed by CK [110]. Three CK interactions, denoted CKpot, CKtb1, and CKtb2, correspond, respectively, to (8-16)POT, (8-16)2BME, and (6-16)2BME in the original work. Here, the numbers in the parentheses denote the mass range of the nuclei used in the fitting procedure. In what follows, we use the total root mean square error for the optimal interaction as the ΔE in Eq. (2.84), $\Delta E = 0.34$ MeV.

Table 3.1: The root mean square (RMS) errors of energies for the 33 data in fit with the our optimized interaction and Cohen-Kurath interactions. All errors are in units of MeV.

	$\boldsymbol{\theta}_{\rm MAP}$	CKpot	CKtb1	CKtb2
total RMS	0.34	0.57	0.47	0.54

We summarize the $\boldsymbol{\theta}_{\rm MAP}$ and its (projected) uncertainty in Tab. 3.2. In addition to the mean values, the standard errors evaluated from the covariance matrix are shown. Here (full) and (1st) are the results with the Hessian matrices evaluated by Eq. (2.88) and by only the first term in Eq. (2.88), respectively. The TBMEs are abbreviated as v($abcd$; JT) with total spin J, total isospin T, and the orbits (a to d), which are $1 = p_{1/2}$ or $3 = p_{3/2}$. The 1σ deviations of parameters are given by the square root of diagonal components of $\boldsymbol{A}^{-1}$. We can see from the difference in the standard deviations between (full) and (1st) that the contribution of second derivative terms are as a whole small compared to the first derivative terms in Eq. (2.88). If it is a good approximation to omit the second term, the calculation cost to evaluate Hessian matrices is reduced by a factor of $1/(2d+1)$.

Table 3.2: Projected posterior distributions for the p-shell effective interactions. The mean values correspond to the optimal value and std values denote associated standard deviations. The values are in the unit of MeV.

SPE/TBME	mean	std (full)	std (1st)
SPE p1	2.49	0.97	0.84
SPE p3	3.89	0.37	0.35
v(1 1 1 1: 0 1)	−1.66	0.69	0.72
v(1 1 1 1: 1 0)	−4.84	0.61	0.58
v(1 1 1 3: 1 0)	−3.34	0.83	0.80
v(1 1 3 3: 0 1)	−5.35	0.66	0.81
v(1 1 3 3: 1 0)	3.68	1.06	1.07
v(1 3 1 3: 1 0)	−6.48	0.93	0.86
v(1 3 1 3: 1 1)	0.75	0.54	0.56
v(1 3 1 3: 2 0)	−6.44	0.69	0.77
v(1 3 1 3: 2 1)	−1.82	0.28	0.30
v(1 3 3 3: 1 0)	−6.94	1.24	1.22
v(1 3 3 3: 2 1)	1.96	0.27	0.27
v(3 3 3 3: 0 1)	−4.15	0.71	0.84
v(3 3 3 3: 1 0)	−3.29	1.14	1.30
v(3 3 3 3: 2 1)	−2.13	0.48	0.50
v(3 3 3 3: 3 0)	−8.73	0.39	0.38

Some TBMEs, whose total angular momentum J and total isospin T are $(J,T) = (1,0)$, show relatively large uncertainties. If we take Gamow-Teller transition strengths or electromagnetic observables into fit, which are sensitive to those parameters, we expect that the uncertainties would become smaller. However, this brings additional issues on the quenching factor or the effective charges. In this thesis, we do not enter into the detail of fitting procedure including those quantities for the sake of simplicity.

3.2 Credible intervals of valence shell model for the 0p-shell nuclei

For the procedure c) in the last section, we performed shell-model calculations with 50,000 LA samples. This number is large enough to suppress the error coming from stochastic choices of samples. The typical size of Monte Carlo error in the mean values of energy eigenvalues is less than 0.1%. All the calculations below are done with KSHELL [12, 13][3].

[3]The public version is available on the page of Prof. N. Shimizu https://sites.google.com/a/cns.s.u-tokyo.ac.jp/kshell/home.

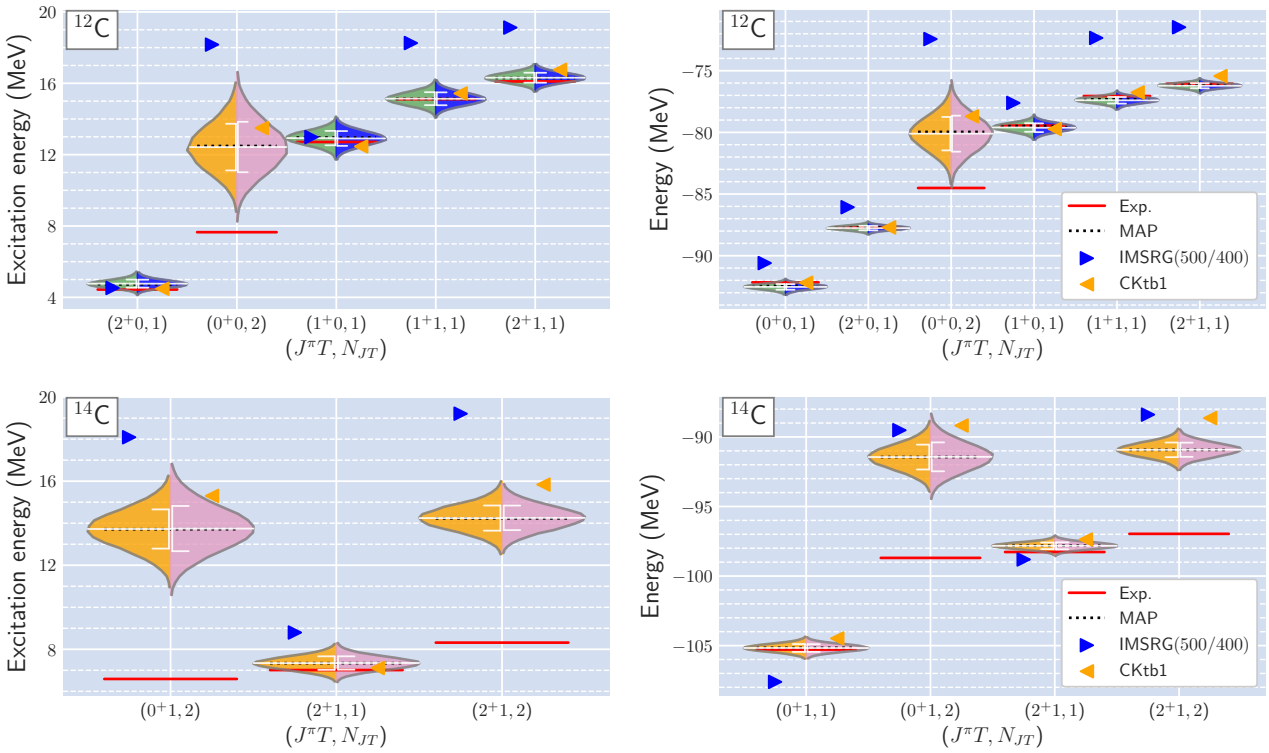

FIG. 3.2: Excitation energies (left column) and energies (right column) of ^{12}C and ^{14}C. Each state is classified by its $(J^\pi T, N_{JT})$, where $(J^\pi T, N_{JT})$ stands for the N-th lowest states with the total angular momentum J^π and the total isospin T. The sold lines colored in red show experimental data and the corresponding theoretical results with the LA samples distribute as violin plots. In each violin plot, mean value and $1\bar{\sigma}$ error are shown by horizontal line and error bar in white, respectively, where $\bar{\sigma}$ means one standard deviation of the results. Results by the our $\boldsymbol{\theta}_{\rm MAP}$, CK interaction and the VS-IMSRG are also shown. While the left half of the violin plot shows the results with LA samples with full evaluation of the Hessian, the right half show the counterpart with the Hessian approximated by only the first term of Eq. (2.88). The bluish and reddish colors, respectively, mean that the state is in and not in fitting procedure.

Now we are ready to evaluate credible intervals of valence shell-model calculations. In FIG. 3.2, we show some selected results of energy spectra. Theoretical results by the LA samples are shown by the so-called *violin plots* in comparison with (i) experimental data on ENSDF databese [164], (ii) the results by one of the CK interactions, and (iii) those by the derived interactions using VS-IMSRG [138], if available. The height and width of violins show, respectively, $3\bar{\sigma}$ and appearance frequencies of the quantities with respect to all the 50,000 LA samples. Here we use the notation $\bar{\sigma}$ for standard deviations of the results with all the LA samples in order to distinguish from the ordinary statistical term σ for Gaussian distributions. The distributions for most of the states are actually Gaussian-like, but we must mention that the distribution of the calculated energy eigenvalues do not necessarily distribute as a Gaussian.

The mean values and $1\bar{\sigma}$ credible intervals are shown by horizontal solid lines and error bars colored in white. Experimental data and theoretical results are classified with the angular momentum J, the parity π, and the isospin T. We note that the total isospin is not determined in some data, in which cases most plausible values are taken from the corresponding theoretical results. The CKtb1 was determined from the closest data to that for our $\boldsymbol{\theta}_{\rm MAP}$, and gives the minimum total root mean square errors among the three CK interactions (see Tab. 3.1). One can see the significant variances in the height of the credible intervals. This enables us to quantify the *relative* reliabilities of the theoretical estimates. We would like to emphasize here that the overall scale of the credible intervals can be controlled by the factor in denominator of the Eq. (2.84). If we adopted unrealistically large ΔE, we can make all the violins cover the corresponding experimental data. In that sense, the *relative* ratio of the height is much more informative than its absolute size. In the followings, we omit *relative* and *relatively* just before nouns or adjectives to avoid the redundancy, but it should be remembered occasionally. Those difference in size of uncertainty for each state would imply much information such as complicated correlations among parameters and many-body configurations, and sensitivity of the parameters to the states.

Excitation spectra like the left panels in FIG. 3.2 are convenient for comparisons with experimental studies with gamma-ray measurements, e.g. to determine which states are more likely relevant to the newly measured gamma rays. One can also work with energy eigenvalues, as shown in the right panels in FIG. 3.2. These are convenient for purposes to see systematic deviations (e.g. with respect to the mass numbers) of shell-model results from experimental data.

From Fig. 3.2, we can make another important remark. Obviously, the second 0^+ state of ^{12}C (Hoyle state) and the second 0^+ and 2^+ states of ^{14}C show large deviations from corresponding experimental data. For the three states, it has been suggested that the 0p shell is insufficient to describe these state and it is now common understanding today, see e.g. [56, 165–173].

We propose that these large discrepancies between the credible intervals and experimental data can be interpreted as an indicator of exotic structures (α clustering, intruder configurations, core excitations, etc.). Since both configurations in the given model space and outside the model space contribute to the true wave function simultaneously, it is not giving a distinct criteria for exotic structures. However, we can deduce which states are more likely to have exotic structures listed above by looking at the relative size of the deviations. This, discussed in the Sec. 3.2.1, is an important outcome of the evaluation of uncertainties in shell-model calculations.

Furthermore, we can assess the validity of the LA by looking into the shapes of violin plots.

If the LA failed to capture the global structure of χ^2, i.e. in the case that there are many local minima or plateaus in the χ^2-potential over the parameter space, it is expected that the mean values of observables calculated with the LA samples will deviate from the evaluation with the MAP estimate and that the shape of violin plots will become asymmetric. However, the deviation is not significant and the shapes of violin are almost symmetric for most of the states: Regarding all the 72 states, the typical discrepancy in energy values between a result by the $\boldsymbol{\theta}_{\rm MAP}$ and a mean value of the results with the LA samples is about 0.1%.

3.2.1 Detailed analyses of the results

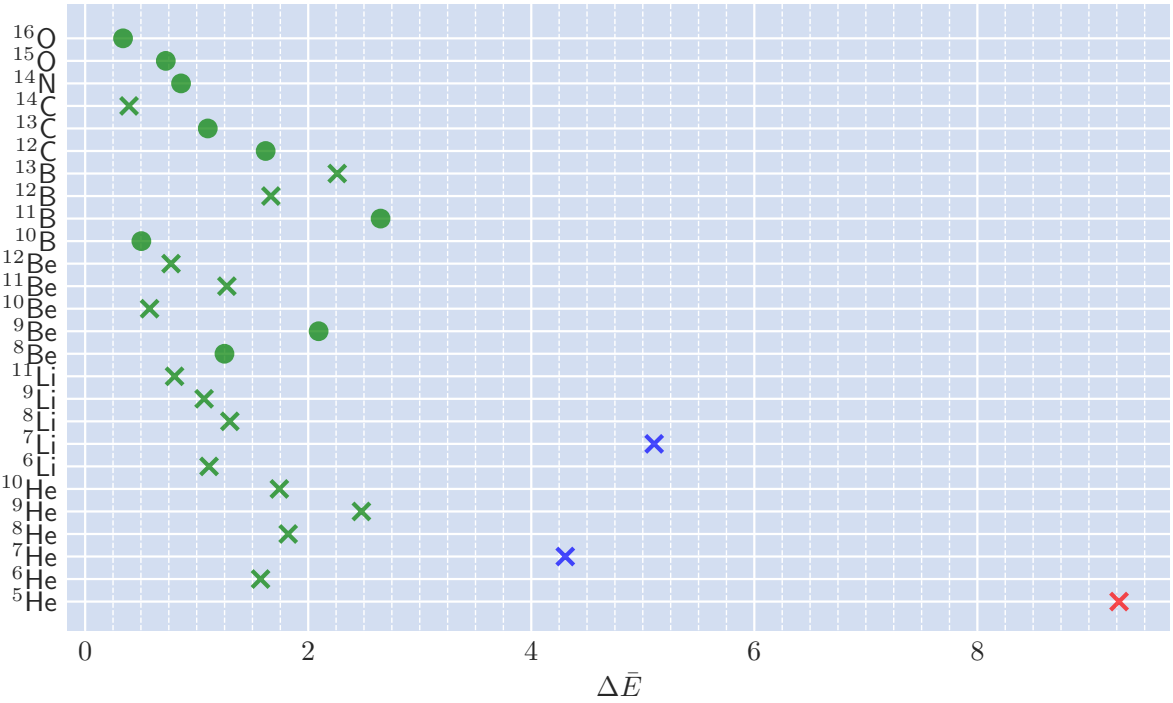

FIG. 3.3: The plot showing the $\Delta\bar{E}$ for the ground states. The states in fit (not in fit) are shown by filled circle (cross) symbols. All the states are classified into the three clusters by the so-called k-means++ method [174].

Hereafter, we explore a sort of generalization ability[4] of the valence CI calculation, (i.e. to confirm that the parameter distributions are not overfitted to the 33 states) by considering 72 low-lying states for 0p-shell nuclei. All these states and corresponding results using $\boldsymbol{\theta}_{\rm MAP}$ and LA samples are summarized in Tables A.1–A.2 in Appendix A. One can find more detailed data including CK and VS-IMSRG results on GitHub [4]. The 33 data used in the parameter

[4]The terminology "generalization ability" is an important concept in Machine learning and in any scientific fields using *parameters*. We expect this will appear in dictionaries for the general public in near future.

inference are specified by the check-mark symbols ✓ in the tables. We note that the g.s. of ^{11}Be is thought to have unnatural parity, which cannot be described in the 0p shell, so the excitation energies are considered from the lowest natural parity state.

For that purpose, let us define the following quantity:

$$\Delta\bar{E}^{(n)}(D) \equiv \frac{|\bar{E}^{(n)}(D)_{\text{theo.}} - E^{(n)}_{\text{exp.}}|}{\bar{\sigma}^{(n)}(D)}, \tag{3.4}$$

where the state is labeled by (n). Here $E^{(n)}_{\text{exp.}}$, $\bar{E}^{(n)}_{\text{theo.}}$, and $\bar{\sigma}^{(n)}(D)$ are experimental energy value, the mean of theoretical estimates using the LA samples, and one standard deviation, respectively. We write dependence on the number of data D explicitly to emphasize the mean value and standard error is dependent on which experimental data is used for the parameter inference. This $\Delta\bar{E}$ is the deviation between experimental data and the mean of the LA results normalized by $1\bar{\sigma}$ error, and this can be defined for both energy eigenvalues and excitation energies.

In FIG. 3.3 and FIG. 3.4 we summarized $\{\Delta\bar{E}\}$ for the all 72 states with the data set $D = 33$. FIG. 3.3 is showing the ground states and FIG. 3.4 is the counterpart for the excited states. The values on x and y axis in FIG. 3.4 are $\Delta\bar{E}$ for energy eigenvalues and excitation energies, respectively. The states in fit (not in fit) are shown by filled circle (cross) symbols. In both plots, for intuitive understanding, all the states are classified into the three clusters by the k-means++ method [174], and colored in green, blue, and red. This is just for rough classification; the number of clusters is not essential.

Regarding the ground states, say, the cluster colored in green, including 20 out of 26 states, can be said to be described (relatively) well by the shell model. For the rest, especially a few odd-mass nuclei, deviations are large. It has been already discussed in e.g. [175] that the continuum effect and the ^{4}He+^{3}He component has large contribution to the ground state energy of ^{7}Li. In such a way, combining information of uncertainties of valence CI calculations with others, we can conclude that the origin of discrepancy is not likely due to the quality of the effective interactions.

Now let us move to the excited states. The three exotic states mentioned above are located far from the origin in FIG. 3.4. From this figure and spectra in Appendix A, one can see that ^{9}Be and ^{11}B respectively show systematic underbinding and overbinding, whereas the excitation spectra are in relatively good agreement with experimental data. In the nuclei located at $x > y$, it is expected that shell-model properly account for the correlation among the ground state and excited states. On the other hand, in the nuclei located at $x < y$ such as ^{8}Be and ^{10}Be, the shell-model apparently miss the correlation among the ground state and the excited states, even though the absolute value of ground state is reproduced relatively well. To

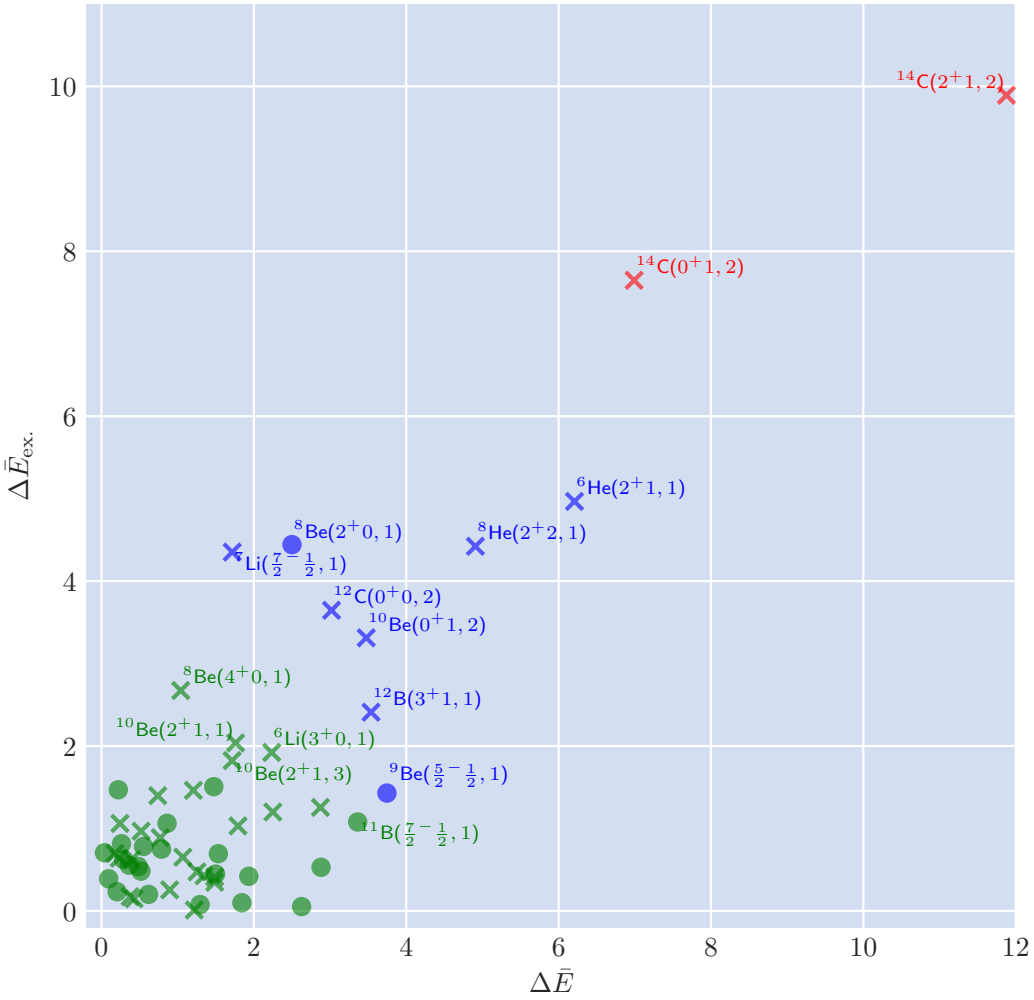

FIG. 3.4: The 2D plot showing $\Delta\bar{E}$ (x-axis) and $\Delta\bar{E}$ for the excitation energy (y-axis). See the main text for more details.

see whether or not this is general tendency for nuclei having α-cluster component, it is needed to perform more comprehensive studies with shell model in other regions and cluster-model or AMD calculations with uncertainty quantification.

Even if there is no preliminary knowledge of which states have exotic structures, one can quantify how the states are to be (or not to be) taken into account through fitting procedures. For that purpose, the distance from the origin in FIGs. 3.3-3.4 or a method like the k-means clustering would provide us with a criterion. In this way, one can identify certain states which should not be included in the fitting procedure, and then update the optimal interaction to more *utilitarian* one. We note that there is a risk of redundancy of data and we do not claim that more and more data should be included, because the number of data to evaluate parameter distributions must be suppressed in future applications for heavier systems. If

available, however, the iteration of inclusion/exclusion of data based on such a majority rule could be a general strategy of data selection to this kind of phenomenological methods.

3.3 Scalability

In this section, we briefly discuss the scalability of our analysis to heavier systems. We have demonstrated the distribution of energy eigenvalues obtained by 50,000 LA samples for $0p$-shell nuclei. In the mass region beyond the $0p$ shell, however, computing 10 low-lying states for each nuclei with 50,000 different interactions is generally not realistic to achieve.

We consider the additional approximation to reduce computational cost significantly. Let us define the following quantity:

$$E^{(n)}_{\mathrm{LFCA}} \equiv \sum_i S^{(n)}_i \langle a^\dagger_i a_i \rangle_{\mathrm{MAP}} + \sum_{ijkl} V^{(n)}_{ijkl} \langle a^\dagger_i a^\dagger_j a_k a_l \rangle_{\mathrm{MAP}}, \tag{3.5}$$

where LFCA is the abbreviation of **L**aplace plus **F**ixed **C**onfiguration **A**pproximation, and the n is the label for LA samples. Instead of diagonalizing all the samples, expectation values of operators are fixed as ones by the MAP interaction and propagate only the uncertainty in SPEs and TBMEs. If the computational cost for shell-model calculation is dominant and the approximation works well, the whole computational time is reduced by a factor of 1/(Number of samples to be diagonalized) compared to the original computation. The validity of this approximation is determined by how much the many-body configurations are sensitive to the parameters.

Very recently, this approximation has been applied in UQ work under the Laplace approximation for sd-shell nuclei [176][5].

3.4 Perspectives and remarks

If one intends to change the data set in the fitting procedure, one can modify the parameter distributions relatively easily by regarding the previously obtained posterior as the prior and by doing the similar inference for the data added. Furthermore, if one makes pseudo data for unknown states, one can estimate their impacts on parameters in the same manner. This flexibility is a benefit of introducing Bayesian inference.

[5] In addition to this approximation, they introduce the principal component analysis to find a linear combinations which are sensitive to the quantity. This gives us more efficient parametrization of the effective interactions.

It should be noted that our comparison between theory and experimental data is based on the the order of the states with given J^π and T counting from the lowest one; one should keep in mind the possibility to make *wrong* comparisons.

This quantification also plays an important role. We have found there are some interesting excited states in ^{14}N, the third 1^+ state at 6.204 MeV and the first 3^+ at 6.446 MeV listed in the ENSDF database [164]. This nuclei has been extensively studied by many theories such as shell-model calculations in the *p* shell [110,177] (and this thesis) and the *psd* shell [165,178,179], the no-core shell model [180] and the antisymmetrized molecular dynamics (AMD) method [181]. Among these, there is no corresponding states reported.

So far we have assumed that it is feasible to find a global minimum under the given data set in the parameter space. However, as can be easily imagined, this is not the case in heavier region even if one moved from (SPEs, TBMEs) space to the space spanned by a fewer number of parameters such as linear combinations of SPEs and TBMEs. For that reason, it is strongly desired to work with the bottom-up approaches and to figure out some way to achieve uncertainty quantifications on the derived effective interactions.

If we could achieve proper uncertainty quantifications in any theoretical models, it must be very helpful to assign spin parities of the newly measured states and to reexamine tentative assignment of the spin parities. That leads to complemental developments of theories and experiments.

Chapter 4

Uncertainty quantification of extrapolation problem in CI-type calculations

In the previous chapter, we have discussed uncertainties coming from input interactions when the eigenvalue problems can be exactly solved under a given interaction and a fixed model space. In this chapter, we consider intrinsic uncertainties in many-body methods by taking full configuration interaction (FCI) method as an example. Some parts of this chapter are also discussed in [182].

4.1 Extrapolation problems in configuration interaction methods

In FIG. 4.1, we show schematic figures of problems to be addressed. Suppose that we have four calculated data y as a function of $1/L$, where L is the system size of the calculations. Usually, the value at $L = \infty$ corresponds to the exact value, and it is often the case that the current limitation of numerical simulations is very far from $L = \infty$. As shown in the right panel, one usually interpolate and/or extrapolate the results to estimate the value which have not been calculated by assuming some parametric function and then optimizing its coefficients to minimize χ^2 deviation between the function and given data points (or an alternative with additional term such as L2 norm). We show three different examples in the right panel. Important remark here is that this type of parametric approach[1] has always a risk of overfitting

[1] More precisely, the point estimation of the parametric model.

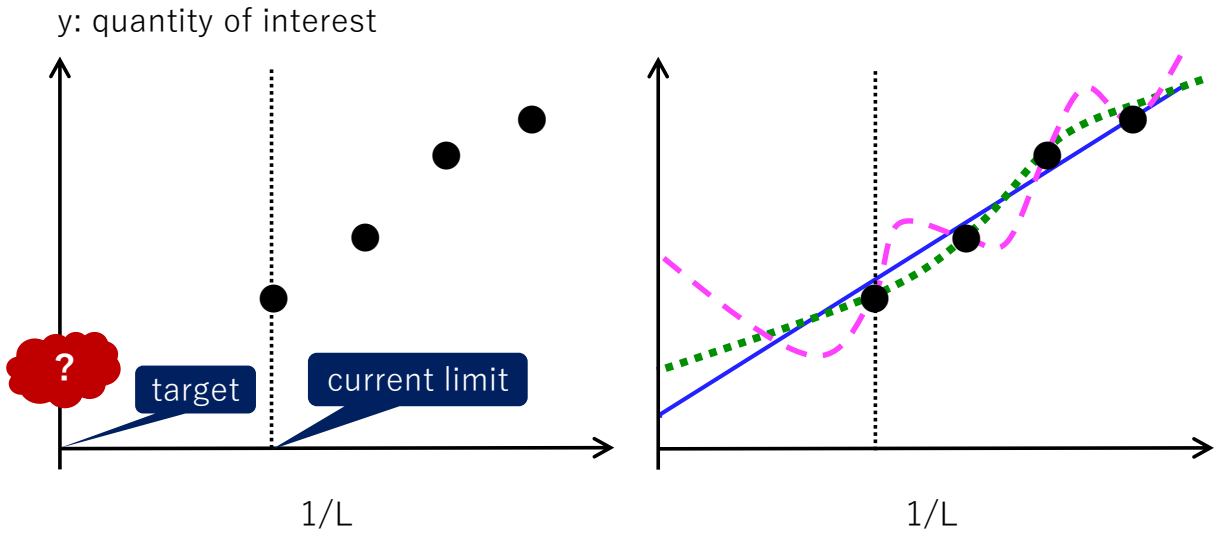

FIG. 4.1: Schematic figure of difficulties in many-body calculations (left) and of conventional approach to do interpolation and extrapolation (right). See the text for more details.

to the given data, i.e. lack of predictive power, unless the underlying mechanisms are exactly expressed by the adopted parametric function, which is not often the case.

We now introduce examples of extrapolation problems in nuclear many-body theories, and will discuss the novel interpolation/extrapolation technique that may alleviate overfitting problems.

The full configuration interaction (FCI) method, which is also known as no-core full configuration (NCFC)/no-core shell model (NCSM) [183, 184][2]

We summarized in FIG. 4.2 published FCI results of ground state energy of ^{6}Li using JISP16/NNLO$_{\rm opt}$ interaction with $\hbar\Omega = 17.5$ MeV [185] and N3LO interaction with $\hbar\Omega = 16.0$ MeV [186] as a function of $N_{\rm max}$. These will be analyzed later.

Despite enormous efforts for developing efficient algorithms and advances in computing power, the currently available $N_{\rm max}$ for the upper $0p$-shell nuclei is around 10 (see e.g. [187]). Since this is still far from exact calculations ($N_{\rm max} = \infty$), one usually extrapolates the results with different $N_{\rm max}$ to $N_{\rm max} = \infty$ to estimate the exact one. In previous studies, several extrapolation methods were proposed and the dependence on them was analyzed [187–193]. The most intuitive way to achieve the extrapolation is the one based on an exponential function [187],

$$E(N_{\rm max}) = E_\infty + a\exp\left(-bN_{\rm max}\right), \tag{4.1}$$

where (E_∞, a, b) are the free parameters.

[2]Note that in some precious studies the authors call the ones using effective interaction by Lee-Suzuki transformation as NCSM and distinguish it from NCFC. In this thesis, we consider only NCFC in this definition.

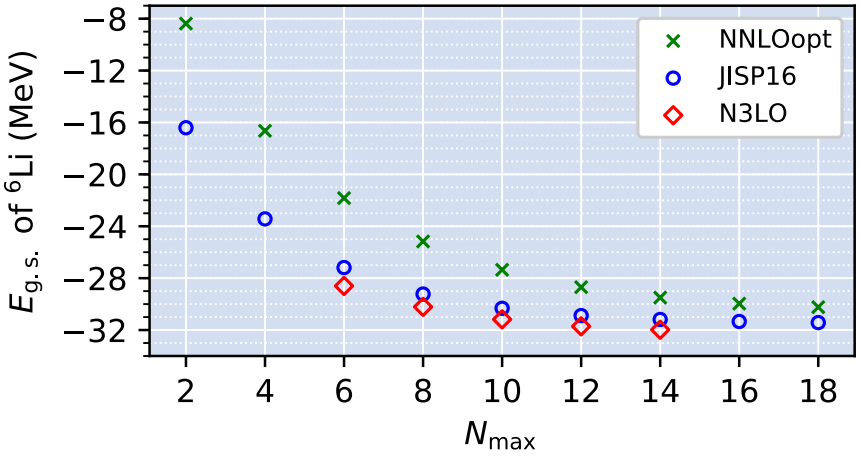

FIG. 4.2: The FCI results of g.s. energy of ^{6}Li using JISP16/NNLO$_{\text{opt}}$ [185] and N3LO [186].

In addition to FCI, such extrapolation techniques are also required in the valence CI method using additional truncations. Representative examples of the truncation are importance truncation scheme [26–29] and Monte Carlo shell model (MCSM) [15, 16] discussed in Sec. 2.2. In those methods, the rapid growth of the number of many-body basis for a valence space is suppressed by selecting a small subset of the many-body basis states which is physically more relevant. For example, the MCSM enables us to perform shell-model calculations for Sm isotopes in which the largest M-scheme dimension is beyond 10^{31} [25]. Note that the current limitation is around 10^{11}. These truncation schemes have been successfully applied to valence CI and also FCI calculations in previous works [30, 31, 194–197].

Any of the extrapolation methods in those studies give intuitively reasonable extrapolated results. However, there is a risk of overfitting, i.e., lack of predictive power for the true exact values. This overfitting is because the χ^2 minimization of a parametric model and point estimation of the parameter leads to too deterministic predictions due to limited expression power of the function. This is a problematic situation if one intends to discuss quantitative issues like a level ordering of states with small energy differences, the positions of proton and neutron drip lines, impacts of three-body forces, etc.

We also note that extrapolation techniques using artificial neural-network (ANN) are proposed very recently [198, 199]. In general, one usually requires large data sets to train the networks, while it is still tough to achieve enormous number of FCI calculations while varying their inputs, $\hbar\Omega$ and N_{max}. In future applications of FCI and also valence CI with importance truncations to heavier systems, it is strongly desired to develop an extrapolation technique which is applicable even to sparse data sets.

In order to overcome the difficulties above, in this chapter, we propose a novel non-

parametric extrapolation method for CI calculations using constrained Gaussian Processes that can give extrapolated results with uncertainty quantification and is applicable even to small data sets. We demonstrate the validity of our model by taking the ground state energy obtained by FCI calculations as an example. The code written in Julia is also available on GitHub [5]. Some related technical issues are discussed in Appendix B.

4.2 Gaussian processes

Gaussian Process (GP) is a popular statistical method as a non-parametric regression model [200]. It is also becoming popular in physics due to its flexibility. We refer the interested readers to recent publications in APS journals [97, 201–206].

The GP regression can be interpreted as a method to describe a distribution over a function space and doing inference of probability for each function. This is just what we need, because this enables us to consider an ensemble of many possible functions for the extrapolation, infer probability of each function, and then quantify uncertainties in the extrapolated value in a statistical manner.

Interestingly, GPs are mathematically equivalent or related to many other models such as ANN, support vector machines, spline models, and so on. We refer the interested readers to e.g. [200,207,208]. The excellent textbook on GPs by C. E. Rasmussen and C. K. I. Williams [200] is available for free in the authors' page. We would like to mention a very recent book written in Japanese introducing GPs as a stochastic generative model of the functions [209]. This would help to understand the usage of GPs for the purpose in this thesis.

Here we introduce some notations. As in statistics literature, $P(a|b)$ denotes the probability distribution of a under the condition b, and we use $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to express the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. In what follows, we consider two variable sets, *data* and *prediction*. The terminology *data*[3] is used to express a set of $X = \{x_i | i = 1, ..., D\}$ and $Y = \{y_i | i = 1, ..., D\}$. Here we assumed that we have D input points. The *prediction* represents positions $X^* = \{x_i^* | i = 1, ..., P\}$ and values $Y^* = \{y_i^* | i = 1, ..., P\}$ for P points where the target values are not known. Specifically, X denotes currently computable $N_{\max}$, and X^* is a set of $N_{\max}$ where it is hard or almost impossible to carry out FCI calculations.

In Gaussian processes, it is assumed that the two target values at the two arbitrary points in the vicinity must be similar, and so-called kernel functions express the similarities. Then the data values Y and prediction values Y^* are assumed to be generated from the multivariate

[3]This is distinguished from *experimental data* in this chapter.

Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ whose covariance matrix $\boldsymbol{\Sigma}$ is given as

$$\boldsymbol{\Sigma} = \begin{bmatrix} K_{XX} & K_{XX^*} \\ K_{XX^*}^T & K_{X^*X^*} \end{bmatrix}. \tag{4.2}$$

Here K_{XX}, K_{XX^*}, and $K_{X^*X^*}$ are respectively $D \times D$, $D \times P$, and $P \times P$ matrices, and these elements are evaluated with a kernel function. One of the most popular choices for the kernel function is the Matérn kernel [210, 211], which is defined for, e.g., two data points x_i and x_j as follows:

$$k_{\mathrm{M}}(x_i, x_j; \nu) = \tau \frac{2^{1-\nu}}{\Gamma(\nu)} \xi^\nu K_\nu(\xi), \tag{4.3}$$

$$\xi \equiv \frac{\sqrt{2\nu}|x_i - x_j|}{\ell}, \tag{4.4}$$

where Γ is the gamma function and K_ν is the modified Bessel function of the second kind. Here the global strength τ[4] and correlation length ℓ are the hyperparameters, and let $\boldsymbol{\theta}$ denote the vector of hyperparameters. We will revisit the issue of hyperparameters later. The other matrix elements, $k_{\mathrm{M}}(x_i, x_j^*; \nu)$ for K_{XX^*} and $k_{\mathrm{M}}(x_i^*, x_j^*; \nu)$ for $K_{X^*X^*}$, are also defined in a similar way.

We follow the typical choice in literature, Matérn kernel with $\nu = 5/2$:

$$k_{\mathrm{M}}(r \equiv |x_i - x_j|; \nu = 5/2) = \tau \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right), \tag{4.5}$$

which makes sample functions second differentiable, and we replace $r \equiv |x_i - x_j|$ in Eq. (4.5) by $r \equiv |\ln x_i - \ln x_j|$ so as to make results independent on the scale of the x-axis and to capture the non-stationary nature of FCI results [212], i.e. results rapidly converge to certain values as functions of $N_{\max}$. The radial basis function (RBF):

$$k_{\mathrm{RBF}}(r \equiv |x_i - x_j|) = \tau \exp\left(-\frac{r^2}{2\ell}\right), \tag{4.6}$$

which corresponds to Matérn kernel with $\nu = \infty$ [200], is also popular choice for the kernel. However, its smoothness of the sample functions is often regarded as too high and, in practice, this too smooth nature sometimes breaks down the positive semi-definiteness of covariant matrices in numerical calculations when the number of sampling points is increased. These technical issues are summarized in Appendix B and the author's GitHub page [213].

Once the kernel function and its hyperparameters are fixed, one can define the joint covari-

[4]The global strength is not introduced in some textbooks and papers. However, we prefer to have τ.

ance matrix in Eq. (4.2) for data/prediction as a function of $\boldsymbol{\theta}$. Then, the joint distribution of data $\boldsymbol{y}$ and predictions $\boldsymbol{y}^*$ under given hyperparameters is given as

$$P(\boldsymbol{y}, \boldsymbol{y}^*|\boldsymbol{\theta}) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu}^* \end{bmatrix}, \Sigma(\boldsymbol{\theta})\right). \tag{4.7}$$

Regarding the mean vectors, we use two different mean functions in this thesis:

- (case a) zero mean: $\boldsymbol{\mu} = \mathbf{0}_D, \boldsymbol{\mu}^* = \mathbf{0}_P$
- (case b) B3 fit: mean of data and prediction are both determined by B3 fit [187], i.e. minimizing χ^2 deviation between the largest three $N_{\max}$ data and the exponential function in the form of $E_\infty + c_0 \exp(-c_1 N_{\max})$ with three free parameters (E_∞, c_0, c_1). In this choice, it can be said that preliminary knowledge on the behavior of the quantity is included in terms of the mean function of GPs.

We will analyze both cases hereafter.

By definition of the conditional probabilities, the left-hand side of Eq. (4.7) can be rewritten as

$$P(\boldsymbol{y}, \boldsymbol{y}^*|\boldsymbol{\theta}) = P(\boldsymbol{y}^*|\boldsymbol{y}, \boldsymbol{\theta}) P(\boldsymbol{y}|\boldsymbol{\theta}). \tag{4.8}$$

Under given $\boldsymbol{\theta}$, one can write down $P(\boldsymbol{y}^*|\boldsymbol{y}, \boldsymbol{\theta})$ and $P(\boldsymbol{y}|\boldsymbol{\theta})$ in a closed form:

$$P(\boldsymbol{y}^*|\boldsymbol{y}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}, \Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}), \tag{4.9}$$

$$\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}(\boldsymbol{\theta}) = \boldsymbol{\mu}^* + K_{XX^*}^T K_{XX}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}), \tag{4.10}$$

$$\Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}(\boldsymbol{\theta}) = K_{X^*X^*} - K_{XX^*}^T K_{XX}^{-1} K_{XX^*}, \tag{4.11}$$

$$P(\boldsymbol{y}|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}, K_{XX}). \tag{4.12}$$

It is common practice to use the so-called maximum a posteriori (MAP), that is the one to maximize the hyperparameter posterior $P(\boldsymbol{\theta}|\boldsymbol{y})$. However, we do not use a single value for the hyperparameters. We do inference of their probability distributions in a Bayesian manner so as to integrate out the hyperparameter dependence, in which case the posterior distribution of $\boldsymbol{y}^*$ for unobserved input $\boldsymbol{x}^*$ can be written as

$$P(\boldsymbol{y}^*|\boldsymbol{y}) \propto \int P(\boldsymbol{y}^*|\boldsymbol{y}, \boldsymbol{\theta}) P(\boldsymbol{y}|\boldsymbol{\theta}) P(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{4.13}$$

4.2.1 Constrained Gaussian Processes (CGP)

In many practical situations, the target function is known to have shape constraints (e.g., monotonicity or convexity) or inequality constraints. That is also the case with problems of interest, i.e. energy eigenvalues in FCI are monotonic and (almost) convex with respect to $N_{\max}$. In general, the accuracy of a statistical model like GP is improved by including such physics information. To this end, we extend Eq. (4.13) to

$$P(\boldsymbol{y}^*|\boldsymbol{y},\alpha,\beta,...) \propto \int P(\boldsymbol{y}^*|\boldsymbol{y},\boldsymbol{\theta})P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})P(\alpha,\beta,\ldots|\boldsymbol{y}^*,\boldsymbol{y})d\boldsymbol{\theta}, \tag{4.14}$$

where $P(\alpha,\beta,...|\boldsymbol{y}^*,\boldsymbol{y})$ is the probability that the conditions $\alpha,\beta,...$ are satisfied under the given $\boldsymbol{y}^*$ and $\boldsymbol{y}$. This follows from the fact that $\boldsymbol{\theta}$ are independent of the conditions, i.e. $P(\alpha,\beta,...|\boldsymbol{y}^*,\boldsymbol{y},\boldsymbol{\theta}) = P(\alpha,\beta,...|\boldsymbol{y}^*,\boldsymbol{y})$ These constraints are introduced independently for each problem of interest.

In general, the integration in Eq. (4.14) cannot be evaluated analytically. Therefore, some approximation or sampling scheme is required. We evaluate the integration in Eq. (4.14) by weighted N_p samples as follows:

$$P(\boldsymbol{y}^*|\boldsymbol{y},\alpha,\beta,...) \simeq \sum_{i=1}^{N_p} w^{(i)} P(\boldsymbol{y}^{*(i)}|\boldsymbol{y},\boldsymbol{\theta}^{(i)}), \tag{4.15}$$

$$w^{(i)} \equiv \frac{P(\boldsymbol{y}|\boldsymbol{\theta}^{(i)})P(\boldsymbol{\theta}^{(i)})P(\alpha,\beta,...|\boldsymbol{y}^{*(i)},\boldsymbol{y})}{\sum_{i=1}^{N_p} P(\boldsymbol{y}|\boldsymbol{\theta}^{(i)})P(\boldsymbol{\theta}^{(i)})P(\boldsymbol{y}^{*(i)}|\boldsymbol{y},\boldsymbol{\theta}^{(i)})P(\alpha,\beta,...|\boldsymbol{y}^{*(i)},\boldsymbol{y})}. \tag{4.16}$$

We employ the particle filtering method [214] (also known as Sequential Monte Carlo) as a sampling scheme to evaluate the summation in Eq. (4.15). In our particle filtering algorithm, *states* $\{\boldsymbol{\theta}^{(i)}, \boldsymbol{y}^{*(i)}\}$ are assigned to *particles* labeled by $i = 1, 2, ..., N_p$ and those particles are evolved independently according to the Metropolis-Hastings method. As can be seen from Eq. (4.16), the plausibility of hyperparameters and predictions, namely $w^{(i)}$, is determined by the balance among the likelihood, the prior, and the fidelity to the constraints.

Particle filter

The pseudo code for our particle filter method is shown in Algorithm 1. At certain points of runs of the code, we do the so-called resampling of the particles according to the fidelity of the constraints, i.e. $P(\alpha,\beta,...|\boldsymbol{y}^{*(i)},\boldsymbol{y})$. This prevents the samplings from being biased by the localization at physically irrelevant region and significantly improve the efficiency. This procedure is explained in Sec. 4.3.

Algorithm 1 Particle filter for GP hyperparameters

input: N (sample size), T (number of step), $P(\boldsymbol{\theta})$ (prior for hyperparameters), $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{x}^*$.

Generating initial hyperparameters:

$\boldsymbol{\theta}^{(i)} (i = 1, ..., N) \sim P(\boldsymbol{\theta})$

Generating N predictions from $P(\boldsymbol{y}^{*(i)}|\boldsymbol{y}, \boldsymbol{\theta}^{(i)})$.

for step $= 1$ to T **do**

  for $i = 1$ to N **do**

    $\boldsymbol{\theta}^{(i)}_{\rm cand} \sim T_1(\boldsymbol{\theta}^{(i)})$ (T_1 : proposal for candidate $\boldsymbol{\theta}^{(i)}$)

    $\alpha^{(i)} \equiv \min\{1, P(\boldsymbol{y}^{*(i)}, \boldsymbol{\theta}^{(i)}_{\rm cand}|\boldsymbol{y})/P(\boldsymbol{y}^{*(i)}, \boldsymbol{\theta}^{(i)}|\boldsymbol{y})\}$

    if $U(0,1)$ (uniform distribution) $< \alpha^{(i)}$ **then**

      $\boldsymbol{\theta}^{(i)} := \boldsymbol{\theta}^{(i)}_{\rm cand}$

    end if

    $\boldsymbol{y}^{*(i)}_{\rm cand} \sim T_2(\boldsymbol{y}^{*(i)})$ (T_2 : proposal for candidate $\boldsymbol{y}^{*(i)}$)

    $\beta^{(i)} \equiv \min\{1, P(\boldsymbol{y}^{*(i)}_{\rm cand}, \boldsymbol{\theta}^{(i)}|\boldsymbol{y})/P(\boldsymbol{y}^{*(i)}, \boldsymbol{\theta}^{(i)}|\boldsymbol{y})\}$

    if $U(0,1) < \beta^{(i)}$ **then**

      $\boldsymbol{y}^{*(i)} := \boldsymbol{y}^{*(i)}_{\rm cand}$

    end if

  end for

  if step == e.g. 100 **then**

    Resampling of N particles according to the weights

  end if

end for

4.3 Application of CGP to extrapolation problems in full CI method

In what follows, we apply the constrained GP model to extrapolation problems in FCI calculations. We summarize again the published FCI results of the ground state energy of ^{6}Li. These are obtained by JISP16/NNLO$_{\rm opt}$ interaction with $\hbar\Omega = 17.5$ MeV [185] and N3LO interaction with $\hbar\Omega = 16.0$ MeV which is softened by similarity renormalization group (SRG) method with a flow parameter $\lambda = 2.02$ fm^{-1} [186]. The results are shown in FIG. 4.2 as a function of $N_{\rm max}$.

Let $\{(x_1, y_1), (x_2, y_2), ..., (x_D, y_D)|x_1 < x_2 < \cdots < x_D\}$ denote the data, i.e. $(x_1, y_1) = (6, -28.602)$, ..., $(x_D, y_D) = (14, -31.977)$ in the case of N3LO results. Unlike least squares fitting of parametric models in which one should remove outliers from data, there is no reason to reduce data in the GP model and we use all $N_{\rm max}$ results as data unless otherwise mentioned.

The extrapolation problem addressed below is to estimate the ground state energies at $N_{\rm max}$ larger than x_D and we express them as $\{(x^*_1, y^*_1), (x^*_2, y^*_2), ..., (x^*_P, y^*_P)|x^*_1 < x^*_2 < \cdots < x^*_P\}$; Here

$x_1^* = x_D + 2$ and P is large even integer. For the sake of simplicity, we restrict ourselves to consider the ground state of ^{6}Li with natural parity, i.e. we only consider even $N_{\max}$. In practice, we truncate at certain finite P value where predictions are converged with respect to $N_{\max}$. A detailed discussion about this P will be given later.

As a minimal constraints to FCI calculations to capture true asymptotic behavior of FCI results, we impose the following two constraints. We assign the labels α and β to the conditions as in Eq. (4.14).

The condition α is variational property, i.e. the monotonicity of energy eigenvalues with respect to $N_{\max}$:

$$P(\alpha|\boldsymbol{y}^*, \boldsymbol{y}) = \Phi(y_D - y_1^*) \times \Phi(y_1^* - y_2^*) \times \cdots \times \Phi(y_{P-1}^* - y_P^*). \tag{4.17}$$

Here we introduced the Probit function:

$$\Phi(z, \kappa) \equiv \int_{-\infty}^{z\kappa} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt, \tag{4.18}$$

where κ is the parameter which controls the strictness of constraints.

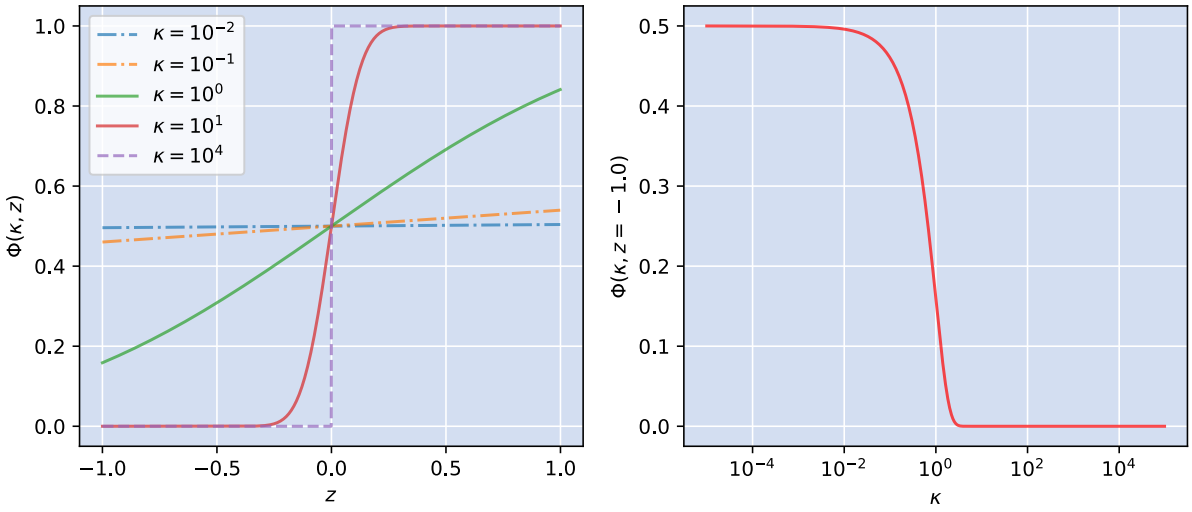

FIG. 4.3: The behavior of the Probit function as a function of z (left) and κ (right).

As shown in the left panel of FIG. 4.3, this Probit function approaches the step function at $z = 0$ when $\kappa \to \infty$. In the right panel, we plot this Probit function as a function of κ, When we suppose that the z is given in unit of MeV and fixed as -1.0, the $\Phi(\kappa, z = -1.0)$ has non-zero value only if $\kappa <\sim 38$. This κ is set as about 10^6 in our code, which is large enough to impose the variational property with satisfactory accuracy less than 0.01 keV level. More

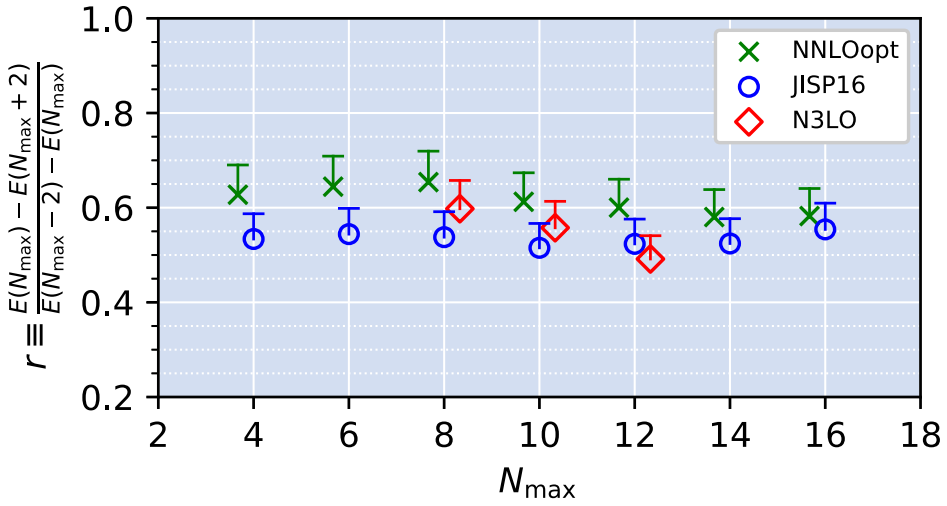

FIG. 4.4: The ratios of energy gains associated with 10% fluctuation to larger value (error bar). The symbols for FCI results are the same as FIG. 4.2. For the visibility of the figure, $\mathrm{NNLO_{opt}}$ and N3LO results are slightly shifted to the left and right, respectively. See the text for more details.

precisely, in our code, the κ is gradually increased to $\sim 10^6$ so as to avoid possible localizations at the early steps of the Monte Carlo sampling. We confirmed that the form of this κ as a function of the Monte Carlo step does not affect the extrapolation results other than the sampling efficiency.

Second condition β is the one about the convergence pattern. Let r_j denotes a ratio of energy gain at nearest neighbor three points labeled by i, j, and k:

$$r_j \equiv \frac{y_j^{(*)} - y_k^{(*)}}{y_i^{(*)} - y_j^{(*)}} \tag{4.19}$$

$$x_i^{(*)} + 2 = x_j^{(*)} = x_k^{(*)} - 2, \tag{4.20}$$

where $(x^{(*)}, y^{(*)})$ denotes data and/or prediction. This $\{r_j\}$ can be used as a measure of convergence pattern of the FCI results. We plot the $\{r_j\}$ in FIG. 4.4 for the three published FCI results. If calculated results of g.s. energy exactly obey exponential function, this means r_j is a constant. However, this is generally not the case as shown in FIG. 4.4.

We impose the constraint on the $\{r_j\}$ as follows:

$$P(\beta|\boldsymbol{y}^*, \boldsymbol{y}) = \Phi\left(R_E - \left|\frac{y_1^* - y_2^*}{y_D - y_1^*}\right|\right) \times \Phi\left(R_E - \left|\frac{y_2^* - y_3^*}{y_1^* - y_2^*}\right|\right) \times \cdots \times \Phi\left(R_E - \left|\frac{y_{P-1}^* - y_P^*}{y_{P-2}^* - y_{P-1}^*}\right|\right), \tag{4.21}$$

where R_E is upper threshold determined as follows:

$$R_E = r_{\text{mean}} + r_{\text{std}}, \tag{4.22}$$

$$r_{\text{mean}} \equiv \frac{y_{D-1} - y_D}{y_{D-2} - y_{D-1}} \geq 0, \tag{4.23}$$

$$r_{\text{std}} \equiv \sigma_r r_{\text{mean}}, \tag{4.24}$$

We use $\sigma_r = 0.1$ throughout this thesis. As can be expected from FIG. 4.4 and Eqs. (4.21)–(4.24), this condition is rather soft constraint on the convergence pattern. When σ_r is large enough, results agree with ones with only the condition α. We refer to the GP extrapolation model using constraints as the constrained Gaussian Process (cGP) model. More specifically, we call cGPs defined with zero-mean vectors and with B3 fit as cGP-a and cGP-b, respectively.

It should be noted that the choice of β can be also regarded as the hyperparameter and one can introduce $P(\beta)$ and consider its marginalization. This is left as a future study.

4.4 Results and Discussions

4.4.1 Extrapolation of g.s. energies

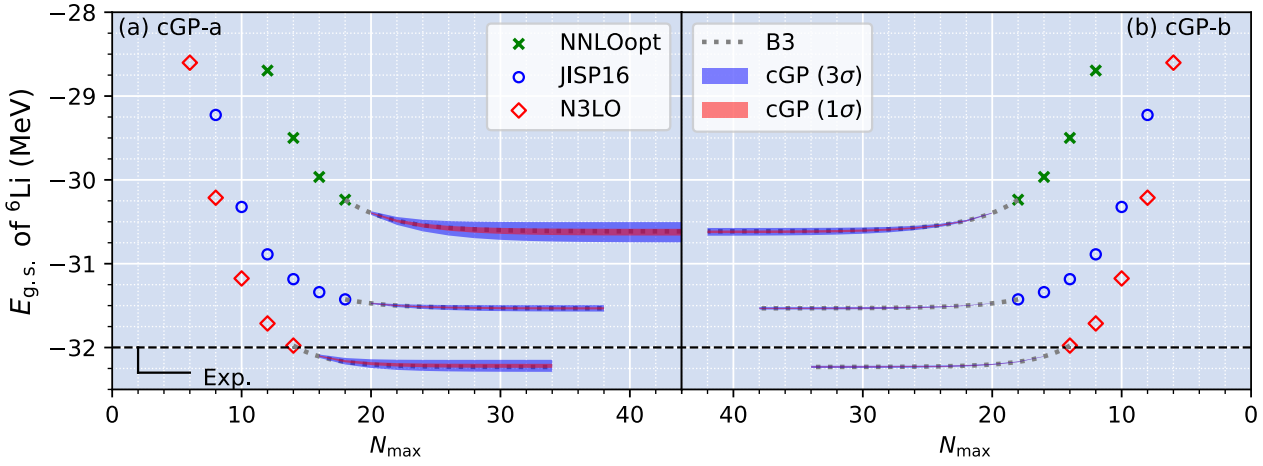

FIG. 4.5: The extrapolated values of ^{6}Li g.s. energy using the two constrained GP models. The 1σ and 3σ confidence intervals of cGP results are shown by bands colored in red and blue. The other symbols are the same as in FIG. 4.2.

We now present the results of cGP predictions for the ground-state energies of ^{6}Li in FIG. 4.5. Extrapolated results of cGP-a and cGP-b are shown by red and blue bands. These are drown by assuming the posterior at each point to be a Gaussian distribution. The extrapolated

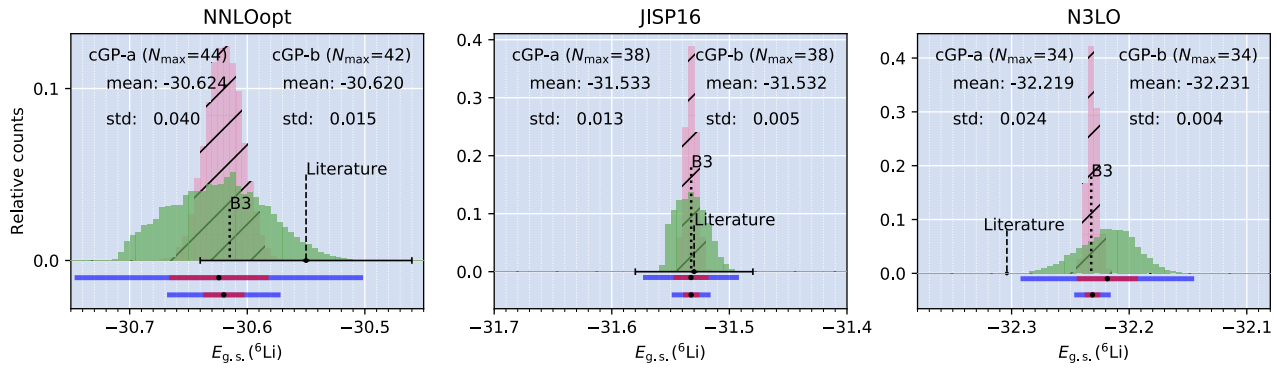

FIG. 4.6: Distribution of predictions for the three interactions at the point where results are converged with respect to $N_{\max}$. The credible intervals shown below the histograms are obtained by assuming a Gaussian distribution.

values are cut at a certain $N_{\max}$ where the mean value is converged. The minimal values of $N_{\max}$ which give convergence are 44, 38 and 34 for $\mathrm{NNLO_{opt}}$, JISP16, and N3LO, respectively. These numbers are consistent with our intuition that harder (repulsive) interaction requires larger $N_{\max}$ to obtain the converged result, while we should remind that these results are for different $\hbar\Omega$. We note that the convergence threshold is set as 0.2 keV and that the possible deviation of the energy values from $E(N_{\max} = \infty)$ are suppressed less than 1 keV because of the constraint β.

In FIG. 4.6, the predictions at which results converged are shown by histograms in comparison with other extrapolated values. The values for *Literature* are from Ref. [185, 186]. Here we note that literature value for N3LO [186] might be obtained by an exponential fit using all five data, though that is not explicitly stated in the references. It must also be noted that it is a highly non-trivial task to fairly compare the results with different extrapolation techniques, because some data is truncated in parametric models and some extrapolation methods use data with multiple $\hbar\Omega$ as in A5 extrapolation of Ref. [185]. Our predictions include parametric B3 fit as special cases within 1σ deviation.

We have confirmed that the particle filtering algorithm gives converged results within a few keV accuracy in case of 20,000 particles after 2,000 times Metropolis-Hastings updates and that independent runs reproduce the same results within Monte Carlo errors.

Our predictions of the exact value of ^{6}Li g.s. energies with uncertainty quantification are summarized in FIG. 4.6.

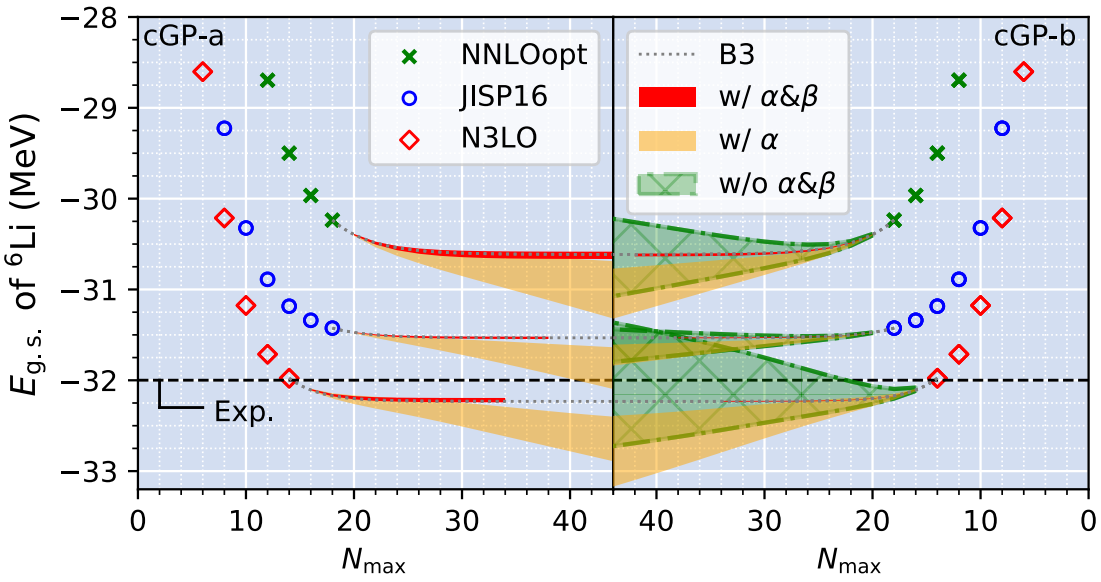

FIG. 4.7: The plot showing the impact of the constraints α (monotonicity) and β (convexity-like). The bands in red, orange and green, show 1σ errors of predictions with α and β, only α, and without constraints, respectively. The red bands are cut at a certain $N_{\max}$ which gives converged results and these are identical with what are shown by red bands in FIG. 4.5.

4.4.2 Impact of the constraints

Here we discuss how extrapolation results are influenced by two constraints imposed. In FIG. 4.7, the impact of the constraints is shown. All symbols are the same with FIG. 4.2, and the cGP-a and cGP-b results are summarized, respectively, in the left and right region. The bands colored with red, orange, and green, correspond to 1σ error of GP predictions with α and β, with the only α, and without constraints, respectively.

As shown in textbooks such as [200] and can be easily confirmed numerically, predictions of unconstrained zero-mean GPs at points far from the data domain converge to zero with a fixed standard error. For this reason, we omit the case of cGP-a without α and β which is obviously not appropriate for the current purpose. It can be seen from FIG. 4.7 that green and orange ones predict to get additional bindings even after $N_{\max} = 40$. This is unlikely under the assumption that we are interested in low-lying states and that the wave function is dominated by relatively lower $N_{\max}$ configurations[5]. For that reason, it is expected that predictions with both constraints α and β are more reliable and suitable than the others.

[5]Of course we can think of exceptions such as ground state of ^{11}Li.

4.4.3 Data dependence

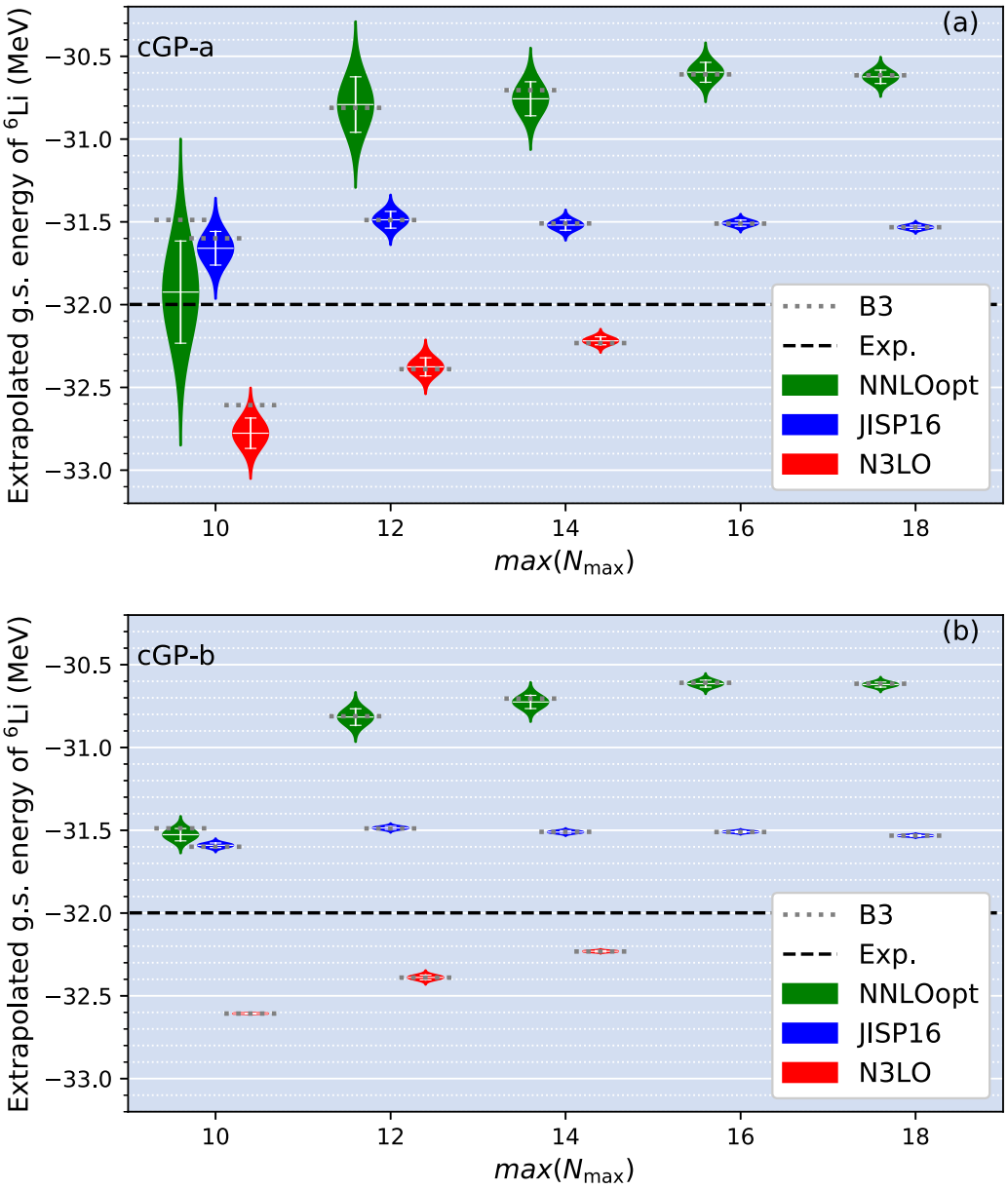

FIG. 4.8: Extrapolated ground state energies as a function of maximum $N_{\max}$ used as data.

We have used all $N_{\max}$ results as data so far. Here we explore the dependence of extrapolated values on the used data to test the potential predictive power of our extrapolation method. In FIG. 4.8, extrapolated values for both cGP-a and cGP-b are shown as a function of the maximum $N_{\max}$ used as data, i.e. $x_D = \max(N_{\max})$.

The height of the violin plot shows 3σ error, and the error bar in white corresponds to 1σ. We also note that the width of violins is scaled to the same, and some violin plots are slightly

shifted to left or right for visibility.

As a whole, the size of credible intervals for cGP-a are larger than that for cGP-b, and the credible intervals of cGP predictions become smaller as higher $N_{\max}$ data added with only one exception. The exception is the one for N3LO interaction with $\max(N_{\max} = 10)$ for cGP-b. This exception can be understood from Eqs. (4.8)–(4.12). In this case, the exponential function exactly fit the given three data and then $\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}$ in Eq. (4.10) is identical with $\boldsymbol{\mu}^*$. Any fluctuation of the joint mean value $\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}$ is not allowed, and this significantly reduces probability weights for functions other than the B3 fit.

In the rest of this subsection, let us regard the mean values at $\max(N_{\max})= 18$ for $\mathrm{NNLO_{opt}}$ and JISP16 and at $\max(N_{\max})= 14$ for N3LO as the tentative exact values.

For $\mathrm{NNLO_{opt}}$ and JISP16, an important remark is that the tentative exact values are covered by cGP-a predictions with relatively lower $\max(N_{\max})$, whereas the cGP-b prediction seems to underestimate the uncertainties in extrapolation. This is because that the cGP-b takes account of fluctuation of the functional form only around the exponential function, while cGP-a includes a broad class of the functions for the extrapolation.

For N3LO, the extrapolated values for both cGP-a and cGP-b are much sensitive to the $\max(N_{\max})$ than the results with $\mathrm{NNLO_{opt}}$ and JISP16. Any predictions by cGP-a, cGP-b, and B3 fit with lower $N_{\max}$ have almost no overlap with the tentative exact value.

The result shown in FIG. 4.8 indicates that the extrapolation for results with N3LO interaction would be a more non-trivial problem than others. This can be partly understood from the behavior of the ratio of energy gains r. Especially in the N3LO case, the r is unstable with respect to $N_{\max}$, as seen from FIG. 4.4. One of the reasons which causes this non-flat behavior of r is that the dominant part of the exact wave function is still not exhausted by the model space with lower $N_{\max}$, i.e. two more excitations still give significant effects. This is a fundamental difficulty in extrapolations. Another possible reason is the SRG evolution of the input nuclear potential. To reveal that, one needs to achieve more comprehensive studies while varying the SRG flow parameter. If one could figure out an additional constraint on the behavior of extrapolated values as a function of $\max(N_{\max})$, this kind of problem seen in N3LO results of ^{6}Li could be alleviated.

4.5 Short summary of this chapter

We introduced an extrapolation method for CI-type calculations using the constrained Gaussian Processes. This method has the following advantages that are required for future

generations of *ab initio* studies to make more quantitative discussions on observables of interest and on quality of the nuclear interaction adopted.

Firstly, our extrapolation method has applicability to sparse data set (e.g., three or five data), which is strongly needed for future FCI calculations.

Secondly, one can naturally incorporate domain knowledge into the model. It is often the case especially in physics that one knows in advance behavior of the target quantity at a certain level, which is ranging from empirical laws to physical principles. One can expect that imposing such information improves the accuracy of the predictions by the model such as GPs. This flexibility might be useful to alleviate difficulties in extrapolation mentioned at the end of the previous section.

Thirdly, uncertainty quantification under the given model and the assumption is available. The overfitting associated with chi-square minimization (e.g., in a single parametric model) can be alleviated by marginalizing over the parameters instead of making point estimation of the values [158]. The dependence of the extrapolated results on the extrapolation function is partly integrated out by means of Gaussian processes, and marginalization of the hyperparameters in Gaussian processes. For those reasons, it is expected that the overfitting problem in terms of the selection of the extrapolation function is partly alleviated by our method, though it should be also mentioned that any type of model are not completely immune to overfitting.

In this thesis, we discussed only the ground state energies obtained by FCI calculations. When it comes to the extrapolation problem of other quantities or in other systems, the main problem is to find the minimal conditions to capture the asymptotic behavior of the quantities of interest as a function of $N_{\max}$. It is a possible future direction along this line to extend the cGP model to a higher dimension. In case of FCI calculations, for example, one can impose the following additional constraint on GP by extending the formulation to $(N_{\max}, \hbar\Omega)$ space: extrapolated values with different $\hbar\Omega$ should converge more or less to the same value. The extension of the formulation to a multi-dimensional space is rather straightforward, while it is expected that one needs more technical analyses in numerical studies such as positive semi-definiteness of the covariance matrices. Our model can also be applied to valence CI techniques using an importance-truncation in which extrapolation function is much more non-trivial than the FCI calculations. It would also be interesting to apply this kind of constrained GP to the finite-size scaling analyses in other systems.

Chapter 5

Summary and future perspectives

Brief summary of recent developments in nuclear structure theory

In recent decades, we have been witnessing very rapid developments in nuclear structure physics community, mainly from two aspects. One is the fundamental interaction owing to the chiral effective field theory, and the other one is many-body problem solvers, especially *ab initio* ones. The scope of those *ab initio* methods has been rapidly extended to medium-mass nuclei, including doubly open-shell nuclei in the last decade. We are now able to compare various *ab initio* methods to check the validity of many-body methods themselves and consistency to the others by using the same chiral potentials. Even the nuclei with $A \sim 20$, such as oxygen isotopes, can be used as a testing ground now.

In addition to this, we are now able to derive shell-model effective interactions starting from free-space nuclear potentials such as chiral forces. This enables us to achieve systematic studies of various properties of nuclei in a consistent manner. Our scope has been enlarged to the systems with $10 \leq A \leq 100$. Until, say, 10 years ago, the many-body perturbation theory (in a broad sense) was the only choice to derive shell-model Hamiltonians. The situation has changed by the extensive studies on the post Hartree-Fock methods such as coupled-cluster and in-medium similarity renormalization group.

All these are remarkable and encouraging. We are getting a path from the fundamental interaction to nuclear many-body problems in a wide range of the nuclear chart. At the same time, however, it is still ambiguous what is the model-dependent or model-independent view of nuclear phenomena caused by a specific model or approximations. In common with most areas of science, many modern theoretical models have been developed by iterative cycles of solving "forward modeling problems" such as methods listed above and "inverse modeling problems", i.e. to make feedbacks from experimental observations, comparisons with other models, etc. They have provided deep insight into phenomena for sure, but it may not work for the nuclei

that are not accessible to experiments. We need a better idea to consider how much we can know about properties of nuclei using the various theories.

The motivation of this work

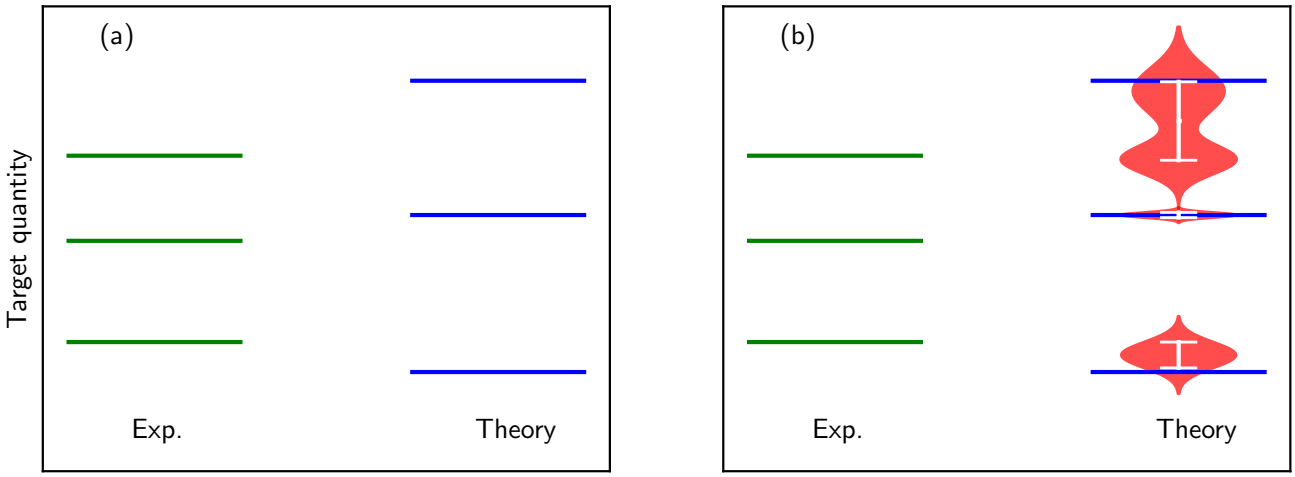

FIG. 5.1: Schematic figure of impacts of UQ studies. See the main text for more details.

Under these circumstances, uncertainty quantification for theoretical calculations, in particular *ab initio* ones, is one of the most urgent and pervasive issues in the low-energy nuclear physics community today. The aim of this work is to contribute in that direction.

In FIG. 5.1, we showed a schematic figure to show the possible impacts of this study in contrast to FIG. 1.1. In the left and right panels, we plotted lines as in FIG. 1.1 in Chapter 1. When one looked at green and blue lines in the left panel (a), one might have the impression that the second-lowest one shows good agreement with data and the agreement in the third one is the worst. However, if we could obtain the uncertainties coming from input parameters and the many-body method itself, and drew the figure like the right panel (b), our impression is completely different from that from the panel (a). While the middle one by the theory has no overlap at all with the experimental counterpart, the first- and third-lowest value are covered by the credible intervals. Those values are more likely to be described by the theory. Moreover, in the third-lowest one, we got a bimodal distribution of the quantity. The lower peak seems to be much more consistent with the data. This may indicate that there still remain some parameters that are not determined well. Uncertainty quantifications would tell us in such a way more insightful information than the conventional comparisons between theories and experiments, though it is not often to observe an extreme case like FIG. 5.1.

Here we emphasize that the benefits of the uncertainty quantification are not limited to putting error bars in predictions. It enables us to visualize non-trivial relation between input

and output of many-body calculations and leads to deeper understandings of the many-body method itself.

Conclusions on the analyses in this thesis

In Chapter 3, we have introduced a way to quantify the theoretical uncertainty from input parameters in valence shell-model calculations by taking the 0*p*-shell as an example. Our work (including [177]) is the first study of systematic uncertainty quantification of the valence shell model. We have proposed that one can quantify a sort of capability of the shell model under the given model space and the assumptions for the probability density of the model parameters. The characteristic information is found in e.g. FIG. 3.2 (spectra of ^{12}C and ^{14}C) and FIGs. 3.3-3.4 (the normalized deviations). We proposed that one can use these plots and the associated quantities to see which states are more likely to have *exotic* structures, i.e. states being out of the model space. It is expected that such information is of great importance for deeper understandings of those states by combining them with other theoretical models.

In Chapter 4, we have proposed a new method to quantify the intrinsic uncertainty of the full configuration interaction method. The uncertainty stems from the ambiguity in the extrapolation method to estimate the exact value. The method has some advantages over the conventional parametric ones. Firstly, this gives us uncertainty estimates in a systematic way. One can think of ensembles of many possible extrapolation functions by means of Gaussian processes. This may alleviate the overfitting, which is caused by the selection of a specific class of function and point estimation of the parameters. In addition to this, we integrate out the dependence on hyperparameter by Monte Carlo sampling; this enables us to consider wider class of functions. Secondly, this method tells us where to stop the FCI calculations; it is not obviously inadvisable to carry out FCI calculations while increasing $N_{\max}$ forever. The point to stop can be deduced from an analysis like a plot in FIG. 4.8, i.e. from the behavior of the extrapolation function as the function of max ($N_{\max}$). In addition to these, its applicability to sparse data set and the flexibility to incorporate domain knowledge into the model are notable points of the method. We should note that the credible intervals by the cGP model show the uncertainty under the given model and assumption; we may underestimate the *true* extrapolation uncertainty as seen in e.g. the panel (b) of FIG. 4.8. It is desired to achieve more comprehensive studies, which are partly described in the next subsection.

If we could evaluate the extrapolation uncertainty in any nucleus calculated by the CI-methods, this significantly improves the quality of comparisons with the state-of-the-art experimental observations. This facilitates the complemental studies among theories and experiments.

Future perspectives

Our method for UQ in a valence CI calculation has already been applied to other systems and other quantities very recently [176]. We hope that more comprehensive studies of UQ also in other models will appear in the near future.

One can consider a future extension of our analyses in this thesis by means of hierarchical usage of Bayesian machinery; one can think of the posterior model distribution:

$$P(m|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|m)P(m)}{\sum_{m'\in\mathcal{M}} P(m', \boldsymbol{y})}, \tag{5.1}$$

where m denotes each model adopted. In this case, the terminology *model* is used for the choice of the likelihood, the choice of the Kernel for GPs, the choice of extrapolation model (GPs, ANN, etc.), and so on. From Eq. 5.1, one can marginalize the model dependence using the so-called Bayesian model averaging:

$$P(\boldsymbol{y}^*|\boldsymbol{y}) = \sum_{m\in\mathcal{M}} P(\boldsymbol{y}^*|\boldsymbol{y}, m)P(m|\boldsymbol{y}). \tag{5.2}$$

Recently, it has been shown that the eigenvector continuation (EC), which is introduced in Ref. [215], can be used as an efficient emulator of *ab inito* methods. This EC has been applied for uncertainty quantification and sensitivity analyses on input parameters such as low-energy constants in the chiral EFT potentials [216, 217]. It is expected that the EC facilitates comprehensive studies of uncertainty[1] propagation in *ab initio* methods. It is also of great importance to consider the possibilities to apply the EC to the CI-type calculations. This is left as one of the possible future studies.

We also mention the acceleration of communications in the nuclear physics community. Recent studies in our community are actively opening the raw data, the calculated data, the code used. For example, the code for the VS-IMSRG is opened by the leading author's GitHub page [218], and the code for eigenvector continuation with no-core shell model [216] is also available from the arXiv page. We hope this trend will accelerate.

[1]The uncertainty here is one in nuclear potentials.

Acknowledgements

First of all, I would like to express my gratitude to my family for their supports, patience, and kindness.

My sincere gratitude also goes to my two supervisors: Professor Takaharu Otsuka and Professor Kenji Fukushima. Professor Otsuka has given me a lot of advices based on his deep knowledge. I am always impressed with his insight into nuclear structure physics. Professor Fukushima kindly accepted me as a member of his group after Prof. Otsuka's retirement. I am always impressed by his wide-ranging knowledge of physics, enthusiastic discussions in his group, and their high research productivity.

I would like to thank those who are (were) in nuclear theory group in UTokyo and/or the theory group of Center for Nuclear Study too: I have learned many things from discussing with them. Dr. Yutaka Utsuno taught me a sort of guideline to write a paper when I wrote the paper [219]. This experience still helps me a lot when writing a paper. My sincere thanks also go to Prof. Noritaka Shimizu who gave me a lot of advice regarding his excellent KSHEL code and computational machinery for UQ studies. I really enjoyed discussing with him and learned many things. I thank Dr. Takashi Abe, Dr. Javier Menéndez, and Dr. Takayuki Miyagi too. They taught me various important points in establishing a research style.

I would like to thank people I met at Institute for Nuclear Physics (IKP) in Technische Universität Darmstadt (TUD). I thank Prof. Achim Schwenk, who kindly accepted my visit twice. Regarding uncertainty quantification of the valence shell-model, it is partly motivated by the question by Dr. Joel Lynn in my seminar at TUD. That was something like..."*How much those monopole components of TBMEs should be converged?*". Prof. Robert Roth must be acknowledged too. When I had the second opportunity to give a seminar in TUD, he gave me a comment like..."*You have shown "data-driven" uncertainties. How can you know "intrinsic uncertainties" in your calculation?*" I hope our study about uncertainties from the extrapolations in CI-methods partly answers this question.

During my Ph.D. program including the master course, I have received many supports. I was supported by the leading program in UTokyo, Materials Education program for the future leaders in Research, Industry, and Technology (MERIT). I was also supported by JSPS as a JSPS doctoral fellow, and some works were supported by JSPS KAKENHI (Grants No. 17J06775). The CNS-RIKEN joint project for large-scale nuclear structure calculations is also acknowledged. Recent studies are partly inspired during three months stay in TU Darmstadt with the aid of JSPS Overseas Challenge Program for Young Researchers (No. 201880031). The Liberal and General Education Center at Utsunomiya University hired me as a project research associate. I really appreciate all the supports from these organizations and budgets.

Appendix A

Summary of credible intervals of valence CI calculations

In this Appendix A, we summarize the complete results of Chapter 3.

In Tables A.1 and A.2, the all states discussed in Chapter 3 are summarized. Here we note that the results shown as "LA samples" are the results with the full evaluation of the Hessian matrix using Eq. (2.88). All distribution of the results are shown by the violin plot in FIG. A.26–FIG. A.7. The excitation energies and energy eigenvalues are shown in the left and right panel respectively. We note that the excitation energies are calculated from the state which has the lowest experimental energy among the natural parity states in the ENSDF database. For that reason, the excitation energy of $(J^\pi T, N) = (1^+0, 1)$ state of ^{10}B can be negative. That can be seen from the distribution of the results of the LA samples. It is known that the ordering of $(J^\pi T, N) = (1^+0, 1)$ and $(J^\pi T, N) = (3^+0, 1)$ states is numerically very subtle and sensitive to the three-body force. For more discussions, see e.g. [65, 138].

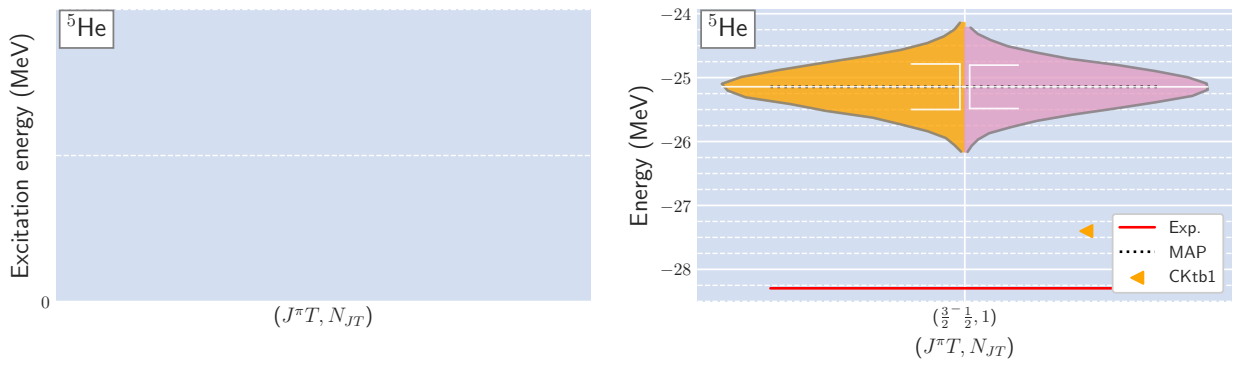

FIG. A.1: Energy spectra for ^{5}He: Excitation enegies (left) and energy eigenvalues (right).

Table A.1: Summary of energy and excitation energies (Ex.) for experimental data and theoretical estimations with the MAP and the LA samples. The total isospin T values with ? symbol are tentative asignments and the states with check-mark symbol are taken into the fit. The associated standard deviations ($1\bar{\sigma}$) are also shown in parentheses for the LA samples.

			Exp.		Theory			
					MAP		LAsamples	
nuclei	$(J^\pi T, N_{JT})$	fit	Energy	Ex.	Energy	Ex.	Energy ($1\bar{\sigma}$)	Ex.($1\bar{\sigma}$)
^{5}He	$(\frac{3}{2}^-\frac{1}{2}, 1)$		-28.30	-	-27.40	-	-25.14(0.36)	-
^{6}He	$(0^+1, 1)$		-29.27	-	-31.34	-	-30.37(0.77)	-
	$(2^+1, 1)$		-27.47	1.80	-27.82	3.52	-25.29(0.37)	5.08(0.66)
^{7}He	$(\frac{3}{2}^-\frac{3}{2}, 1)$		-28.86	-	-28.84	-	-26.97(0.46)	-
^{8}He	$(0^+2, 1)$		-31.40	-	-31.68	-	-30.30(0.59)	-
	$(2^+2, 1)$		-28.30	3.10	-27.09	4.60	-25.28(0.60)	5.03(0.44)
^{9}He	$(\frac{1}{2}^-\frac{5}{2}, 1)$		-29.04	-	-27.06	-	-26.40(1.04)	-
^{10}He	$(0^+3, 1)$		-29.95	-	-26.01	-	-27.77(1.21)	-
^{6}Li	$(1^+0, 1)$		-31.99	-	-32.86	-	-32.66(0.59)	-
	$(3^+0, 1)$		-29.80	2.19	-30.72	2.14	-28.25(0.71)	4.40(1.15)
	$(0^+1, 1)$		-28.42	3.57	-30.35	2.51	-29.37(0.77)	3.28(0.67)
^{7}Li	$(\frac{3}{2}^-\frac{1}{2}, 1)$		-39.24	-	-40.98	-	-40.50(0.26)	-
	$(\frac{7}{2}^-\frac{1}{2}, 1)$		-34.60	4.64	-36.20	4.79	-33.95(0.40)	6.55(0.44)
	$(\frac{5}{2}^-\frac{1}{2}, 1)$		-32.55	6.69	-33.58	7.41	-33.52(0.71)	6.99(0.72)
^{8}Li	$(2^+1, 1)$		-41.28	-	-41.80	-	-41.66(0.30)	-
	$(1^+1, 1)$		-40.30	0.98	-40.72	1.08	-40.58(0.34)	1.08(0.39)
	$(3^+1, 1)$		-39.02	2.25	-40.11	1.69	-39.50(0.32)	2.16(0.27)
	$(1^+1, 2)$		-38.07	3.21	-39.03	2.76	-37.66(0.69)	4.00(0.81)
	$(1^+1, 3)$		-35.88	5.40	-36.84	4.96	-36.15(0.62)	5.51(0.75)
	$(4^+1, 1)$		-34.75	6.53	-35.59	6.21	-34.65(0.45)	7.02(0.46)
^{9}Li	$(\frac{3}{2}^-\frac{3}{2}, 1)$		-45.34	-	-46.12	-	-45.72(0.35)	-
	$(\frac{1}{2}^-\frac{3}{2}, 1)$		-42.65	2.69	-42.78	3.34	-42.91(0.68)	2.80(0.64)
^{11}Li	$(\frac{3}{2}^-\frac{5}{2}, 1)$		-45.71	-	-44.40	-	-46.30(0.72)	-
^{8}Be	$(0^+0, 1)$	✓	-56.50	-	-55.96	-	-56.86(0.30)	-
	$(2^+0, 1)$	✓	-53.47	3.03	-52.55	3.41	-52.91(0.22)	3.95(0.21)
	$(2^+1, 1)$	✓	-39.87	16.63	-40.16	15.80	-40.03(0.30)	16.83(0.43)
	$(4^+0, 1)$		-45.15	11.35	-45.45	11.28	-44.59(0.52)	13.04(0.63)
	$(1^+1, 1)$	✓	-38.86	17.64	-39.08	16.88	-38.95(0.34)	17.91(0.34)
^{9}Be	$(\frac{3}{2}^-\frac{1}{2}, 1)$	✓	-58.16	-	-57.76	-	-57.77(0.19)	-
	$(\frac{5}{2}^-\frac{1}{2}, 1)$	✓	-55.73	2.43	-55.11	2.64	-55.00(0.18)	2.76(0.23)
	$(\frac{7}{2}^-\frac{1}{2}, 1)$		-51.77	6.39	-51.57	6.19	-50.97(0.30)	6.80(0.33)
^{10}Be	$(0^+1, 1)$		-64.98	-	-65.13	-	-65.15(0.29)	-
	$(2^+1, 1)$		-61.61	3.37	-61.43	3.70	-61.24(0.22)	3.90(0.26)
	$(2^+1, 2)$		-59.02	5.96	-59.71	5.43	-59.34(0.27)	5.80(0.33)
	$(0^+1, 2)$		-58.80	6.18	-52.79	12.35	-55.45(1.00)	9.70(1.06)
	$(2^+1, 3)$		-57.43	7.54	-55.89	9.24	-56.78(0.42)	8.37(0.45)
	$(2^+1, 4)$		-55.42	9.56	-54.78	10.35	-54.93(0.41)	10.22(0.45)
^{11}Be	$(\frac{1}{2}^-\frac{3}{2}, 1)$		-65.16	-	-64.32	-	-64.65(0.39)	-
	$(\frac{3}{2}^-\frac{3}{2}, 1)$		-62.80	2.35	-62.07	2.23	-62.55(0.35)	2.08(0.30)

Table A.2: Counterpart of Tab. A.1 for nuclei with relatively larger mass numbers.

			Exp.		Theory			
					MAP		LAsamples	
nuclei	$(J^\pi T, N_{JT})$	fit	Energy	Ex.	Energy	Ex.	Energy ($1\bar{\sigma}$)	Ex.($1\bar{\sigma}$)
^{12}Be	$(0^+2,1)$		-68.65	-	-66.91	-	-68.18(0.61)	-
	$(2^+2,1)$		-66.55	2.10	-63.79	3.12	-65.34(0.51)	2.84(0.61)
^{10}B	$(3^+0,1)$	✓	-64.75	-	-64.76	-	-64.62(0.25)	-
	$(1^+0,1)$	✓	-64.03	0.72	-63.99	0.77	-64.44(0.27)	0.19(0.35)
	$(0^+1,1)$	✓	-63.01	1.74	-63.16	1.60	-63.17(0.29)	1.45(0.37)
	$(1^+0,2)$	✓	-62.60	2.15	-62.35	2.41	-62.62(0.34)	2.00(0.39)
	$(2^+0,1)$	✓	-61.17	3.58	-61.38	3.38	-61.42(0.29)	3.21(0.35)
	$(2^+1,1)$	✓	-59.59	5.16	-59.46	5.30	-59.27(0.22)	5.35(0.28)
	$(4^+0,1)$	✓	-58.73	6.02	-58.81	5.95	-58.81(0.29)	5.81(0.34)
	$(2^+1,2)$	✓	-57.28	7.47	-57.73	7.03	-57.37(0.27)	7.25(0.39)
^{11}B	$(\frac{3}{2}^-\frac{1}{2},1)$	✓	-76.20	-	-76.35	-	-76.70(0.18)	-
	$(\frac{1}{2}^-\frac{1}{2},1)$	✓	-74.08	2.12	-74.57	1.78	-74.72(0.22)	1.98(0.28)
	$(\frac{5}{2}^-\frac{1}{2},1)$	✓	-71.76	4.44	-71.89	4.47	-72.24(0.18)	4.45(0.25)
	$(\frac{3}{2}^-\frac{1}{2},2)$	✓	-71.19	5.02	-71.34	5.01	-71.71(0.32)	4.99(0.35)
	$(\frac{7}{2}^-\frac{1}{2},1)$	✓	-69.46	6.74	-70.42	5.93	-70.24(0.23)	6.45(0.27)
^{12}B	$(1^+1,1)$		-79.58	-	-79.38	-	-80.01(0.26)	-
	$(2^+1,1)$		-78.62	0.95	-78.06	1.33	-78.84(0.21)	1.18(0.34)
	$(0^+1,1)$		-76.85	2.72	-74.71	4.67	-76.68(0.80)	3.33(0.88)
	$(2^+1,2)$		-75.82	3.76	-74.59	4.79	-75.96(0.34)	4.05(0.47)
	$(1^+1,2)$		-74.58	5.00	-74.73	4.65	-75.02(0.37)	5.00(0.36)
	$(3^+1,1)$		-73.96	5.61	-74.25	5.13	-75.23(0.37)	4.78(0.35)
	$(1^+1,3)$		-72.97	6.60	-72.80	6.58	-72.46(0.69)	7.56(0.68)
^{13}B	$(\frac{3}{2}^-\frac{3}{2},1)$		-84.45	-	-84.02	-	-85.28(0.36)	-
^{12}C	$(0^+0,1)$	✓	-92.16	-	-92.19	-	-92.53(0.23)	-
	$(2^+0,1)$	✓	-87.72	4.44	-87.71	4.48	-87.76(0.18)	4.77(0.23)
	$(0^+0,2)$		-84.51	7.65	-78.70	13.49	-80.11(1.35)	12.42(1.31)
	$(1^+0,1)$	✓	-79.45	12.71	-79.73	12.46	-79.60(0.35)	12.93(0.41)
	$(1^+1,1)$	✓	-77.05	15.11	-76.76	15.43	-77.39(0.26)	15.14(0.36)
	$(2^+1,1)$	✓	-76.06	16.11	-75.43	16.76	-76.21(0.21)	16.31(0.28)
^{13}C	$(\frac{1}{2}^-\frac{1}{2},1)$	✓	-97.11	-	-96.37	-	-96.88(0.21)	-
	$(\frac{3}{2}^-\frac{1}{2},1)$	✓	-93.42	3.68	-92.72	3.65	-93.08(0.18)	3.80(0.27)
	$(\frac{5}{2}^-\frac{1}{2},1)$	✓	-89.56	7.55	-89.12	7.24	-89.54(0.23)	7.33(0.30)
	$(\frac{1}{2}^-\frac{1}{2},2)$		-88.22	8.89	-87.76	8.61	-88.31(0.46)	8.56(0.51)
^{14}C	$(0^+1,1)$		-105.28	-	-104.48	-	-105.17(0.28)	-
	$(0^+1,2)$		-98.69	6.59	-89.18	15.30	-91.45(0.89)	13.72(0.93)
	$(2^+1,1)$		-98.27	7.01	-97.37	7.11	-97.83(0.23)	7.34(0.32)
	$(2^+1,2)$		-96.97	8.32	-88.64	15.84	-90.93(0.53)	14.24(0.60)
^{14}N	$(1^+0,1)$	✓	-104.66	-	-104.31	-	-104.40(0.26)	-
	$(0^+1,1)$	✓	-102.35	2.31	-101.48	2.82	-102.17(0.28)	2.23(0.40)
	$(1^+0,2)$	✓	-100.71	3.95	-100.30	4.01	-100.29(0.29)	4.12(0.38)
^{15}O	$(\frac{1}{2}^-\frac{1}{2},1)$	✓	-111.95	-	-112.20	-	-112.09(0.19)	-
	$(\frac{3}{2}^-\frac{1}{2},1)$	✓	-105.78	6.18	-105.17	7.04	-105.84(0.27)	6.26(0.34)
^{16}O	$(0^+0,1)$	✓	-127.62	-	-128.34	-	-127.71(0.26)	-

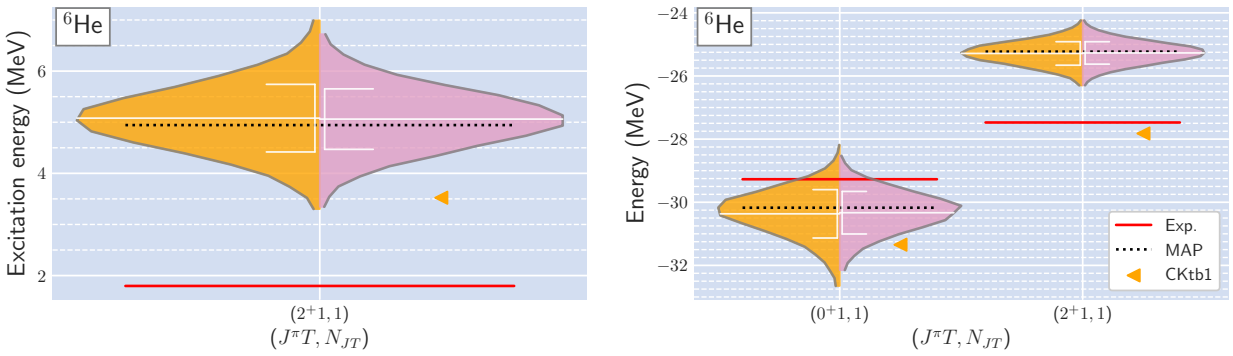

FIG. A.2: Energy spectra for ^{6}He: Excitation enegies (left) and energy eigenvalues (right).

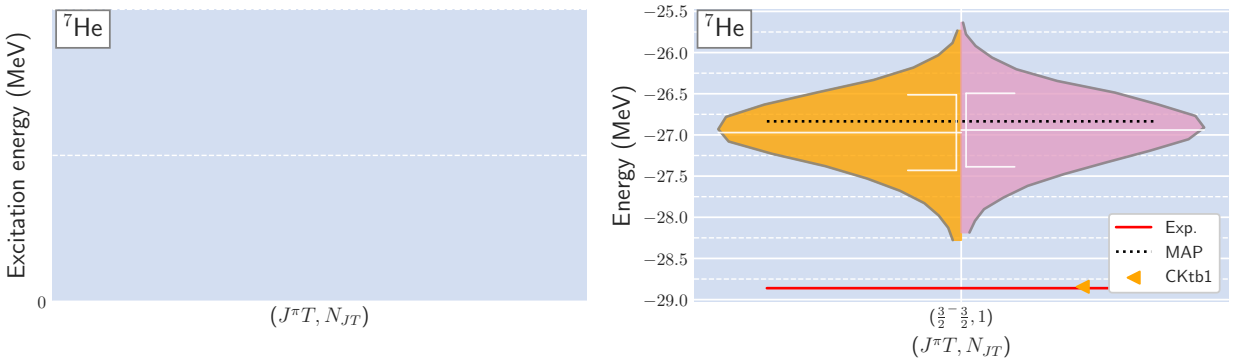

FIG. A.3: Energy spectra for ^{7}He: Excitation enegies (left) and energy eigenvalues (right).

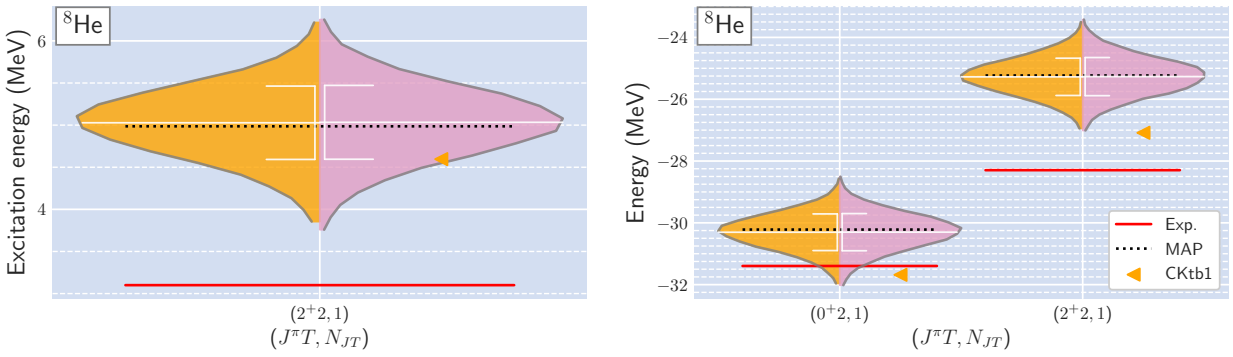

FIG. A.4: Energy spectra for ^{8}He: Excitation enegies (left) and energy eigenvalues (right).

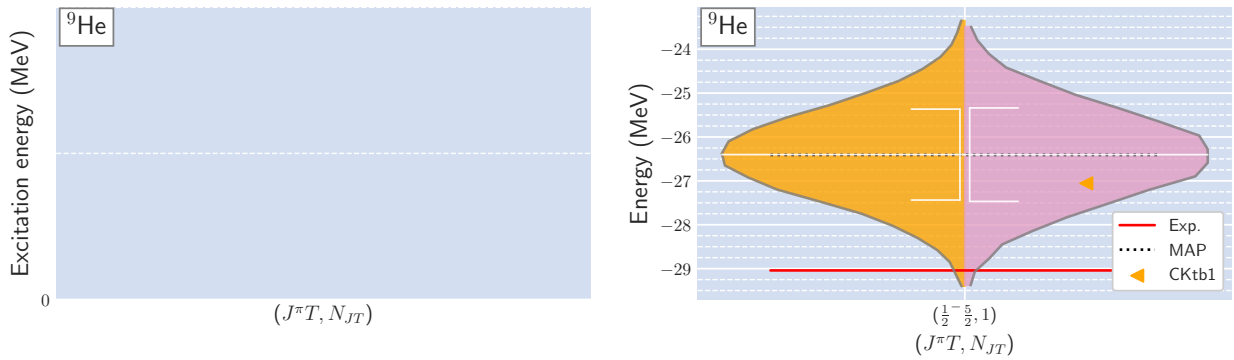

FIG. A.5: Energy spectra for ^{9}He: Excitation enegies (left) and energy eigenvalues (right).

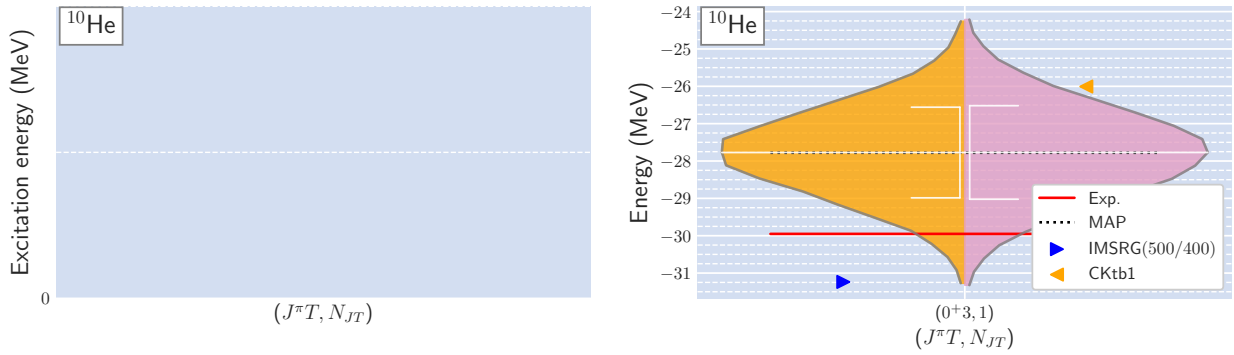

FIG. A.6: Energy spectra for ^{10}He: Excitation enegies (left) and energy eigenvalues (right).

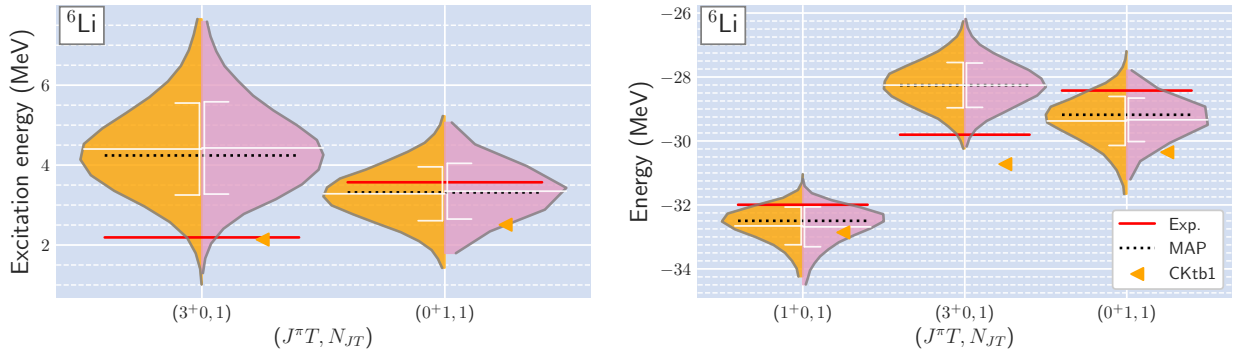

FIG. A.7: Energy spectra for ^{6}Li: Excitation enegies (left) and energy eigenvalues (right).

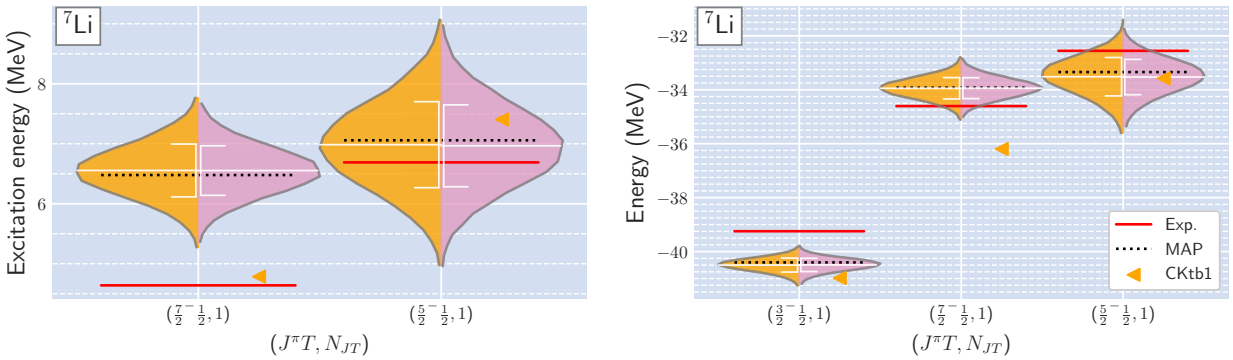

FIG. A.8: Energy spectra for ^{7}Li: Excitation enegies (left) and energy eigenvalues (right).

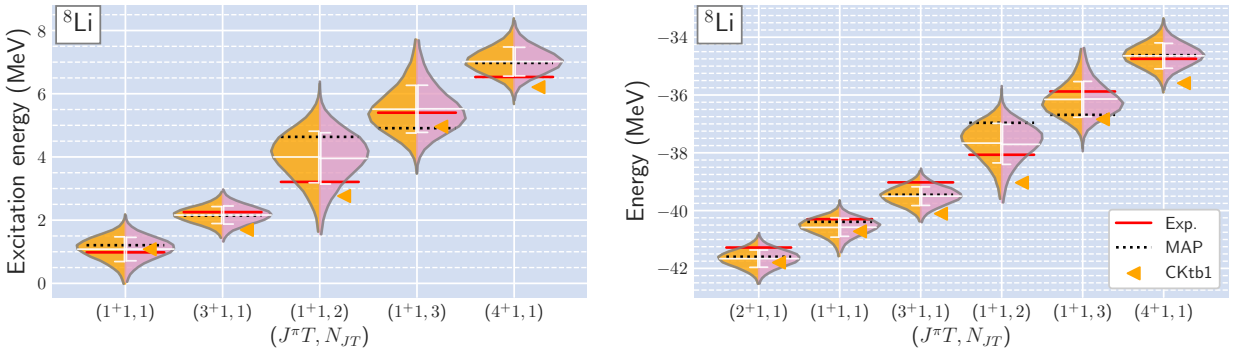

FIG. A.9: Energy spectra for ^{8}Li: Excitation enegies (left) and energy eigenvalues (right).

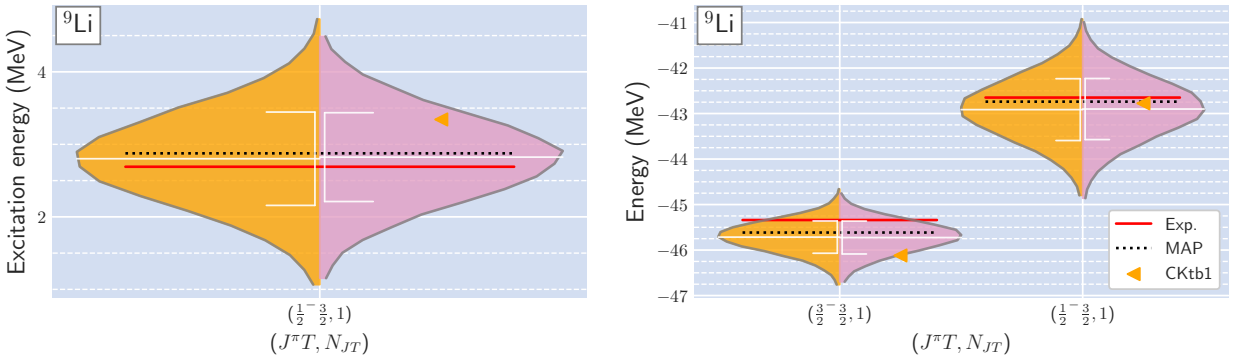

FIG. A.10: Energy spectra for ^{9}Li: Excitation enegies (left) and energy eigenvalues (right).

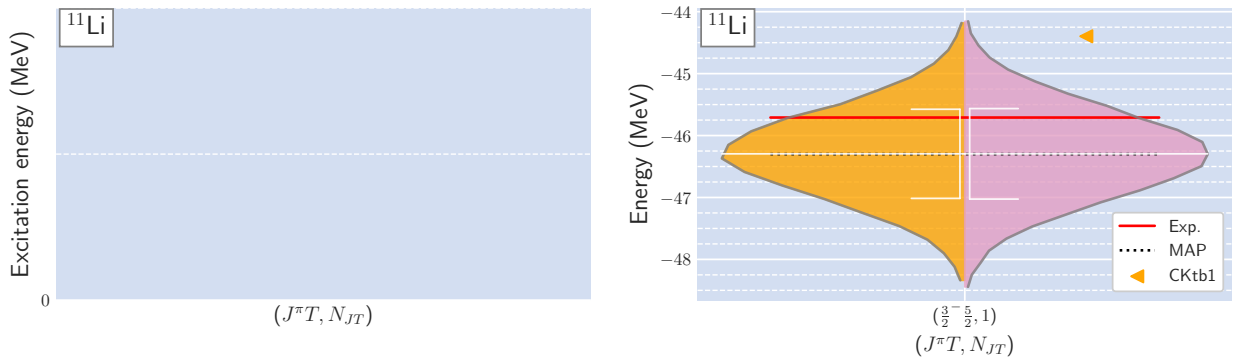

FIG. A.11: Energy spectra for ^{11}Li: Excitation enegies (left) and energy eigenvalues (right).

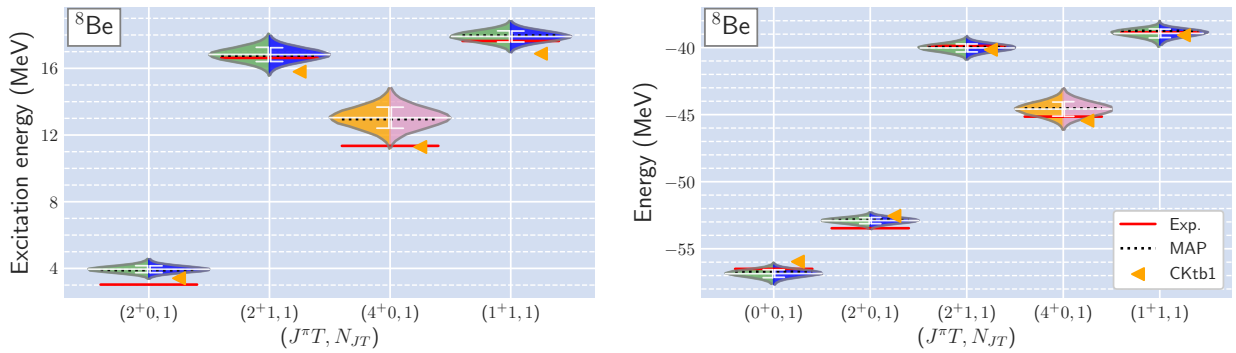

FIG. A.12: Energy spectra for ^{8}Be: Excitation enegies (left) and energy eigenvalues (right).

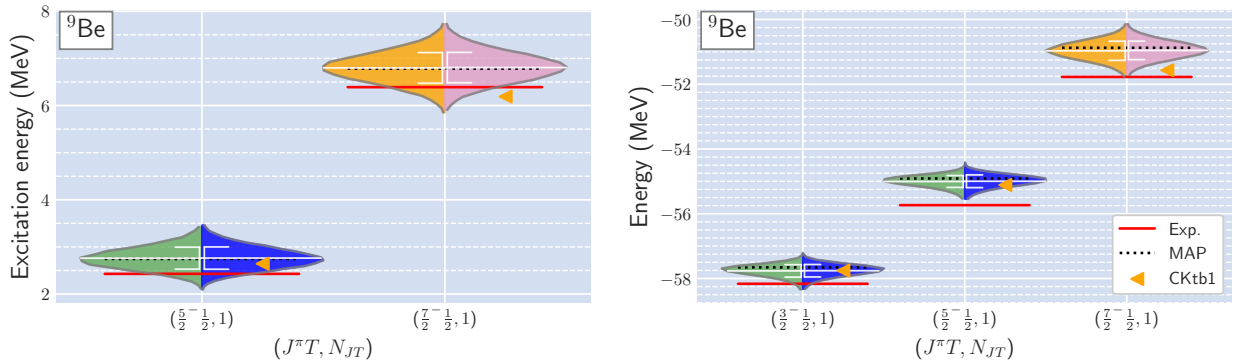

FIG. A.13: Energy spectra for ^{9}Be: Excitation enegies (left) and energy eigenvalues (right).

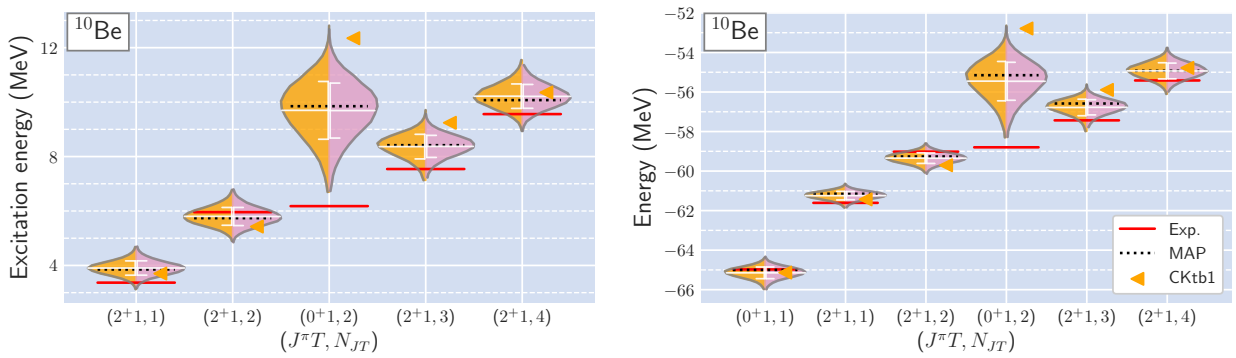

FIG. A.14: Energy spectra for ^{10}Be: Excitation enegies (left) and energy eigenvalues (right).

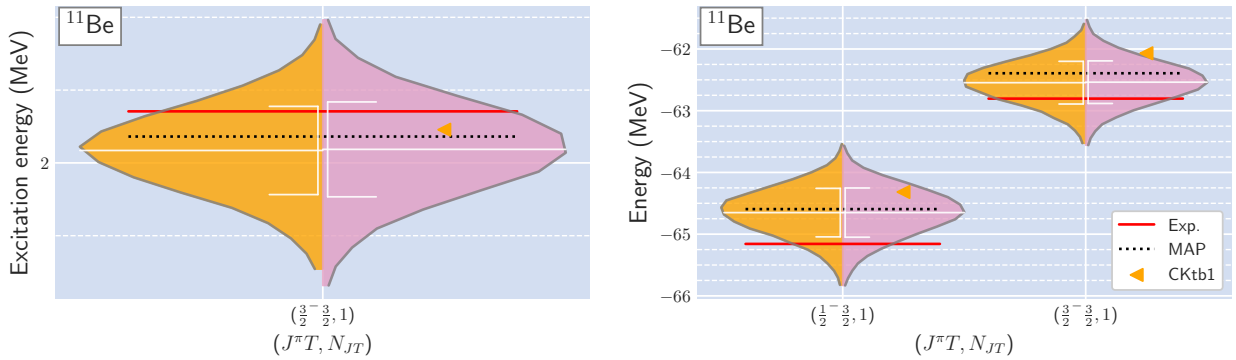

FIG. A.15: Energy spectra for ^{11}Be: Excitation enegies (left) and energy eigenvalues (right).

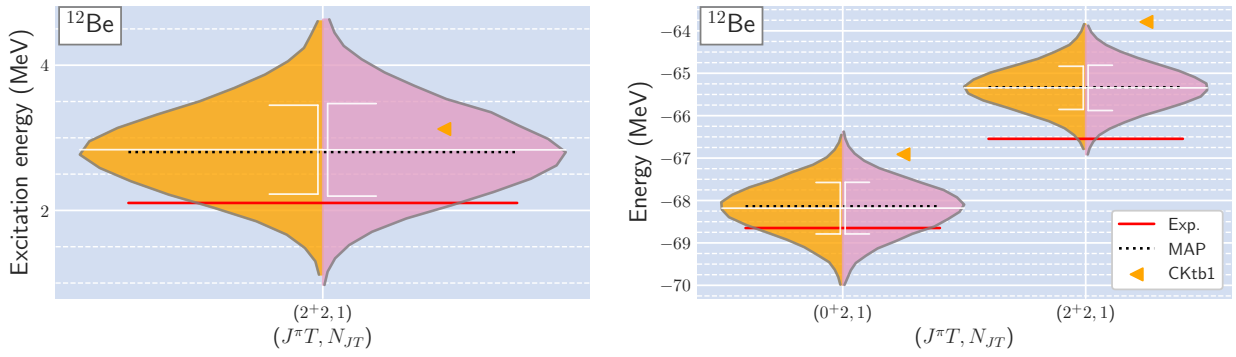

FIG. A.16: Energy spectra for ^{12}Be: Excitation enegies (left) and energy eigenvalues (right).

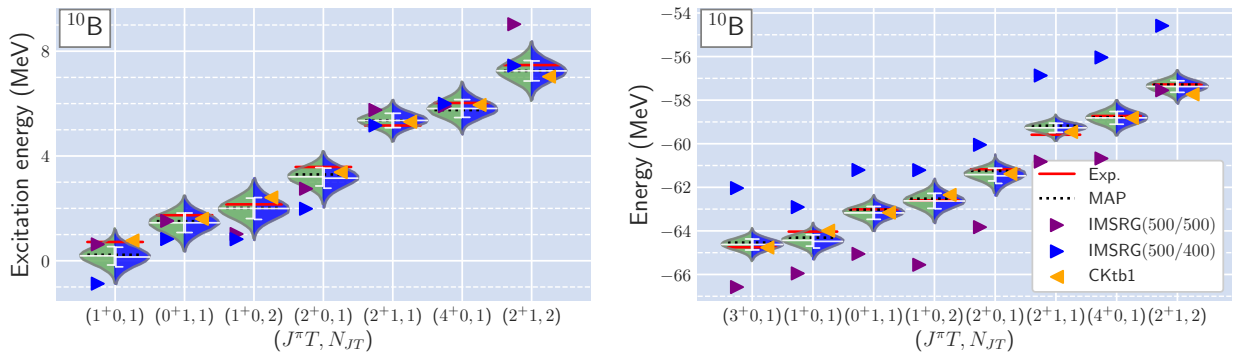

FIG. A.17: Energy spectra for ^{10}B: Excitation enegies (left) and energy eigenvalues (right).

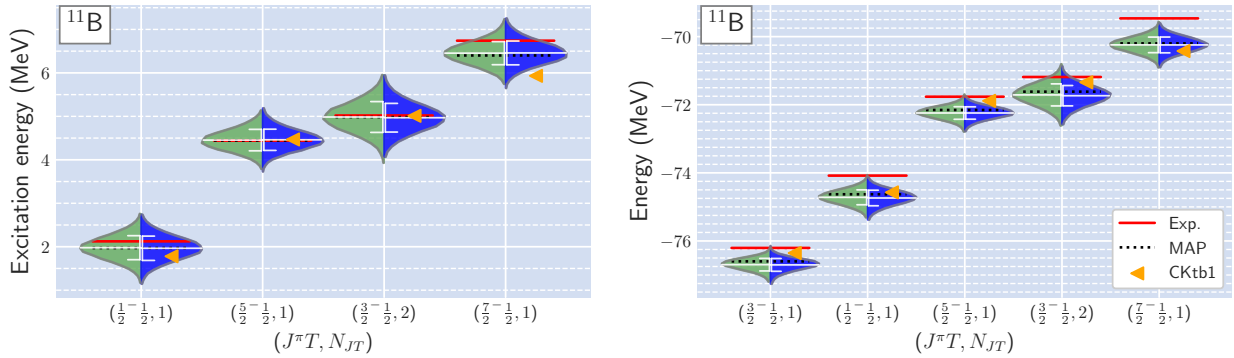

FIG. A.18: Energy spectra for ^{11}B: Excitation enegies (left) and energy eigenvalues (right).

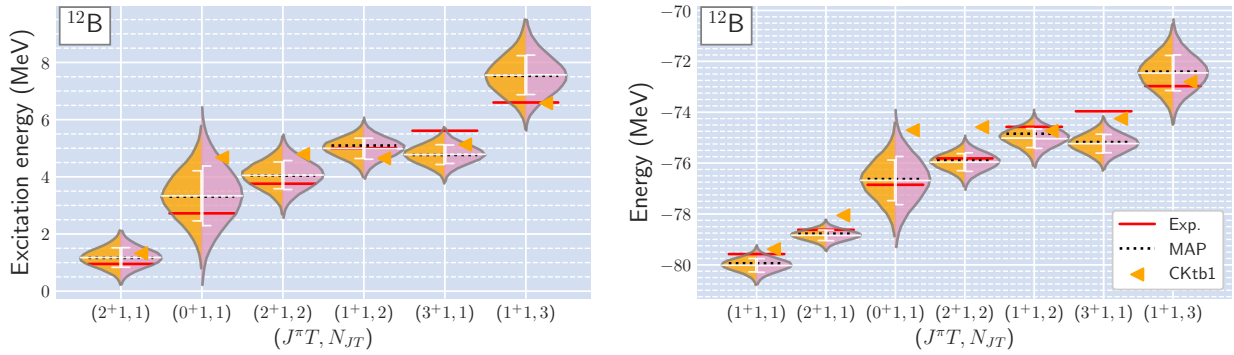

FIG. A.19: Energy spectra for ^{12}B: Excitation enegies (left) and energy eigenvalues (right).

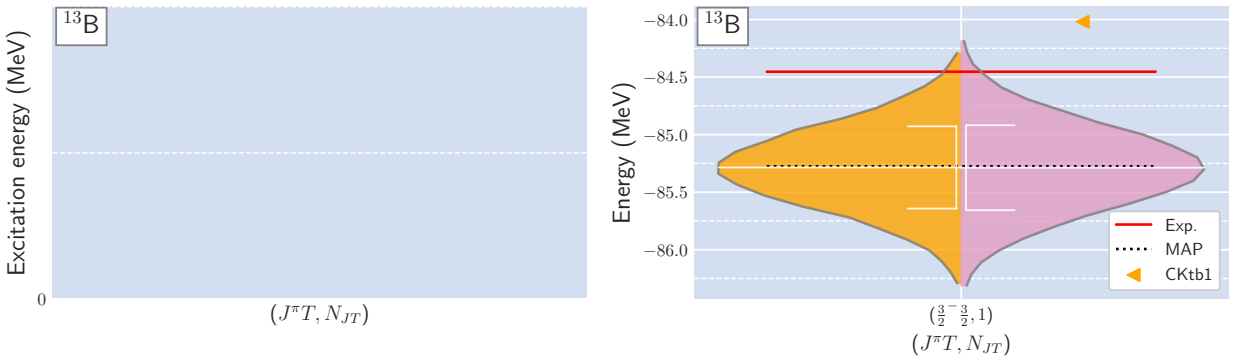

FIG. A.20: Energy spectra for ^{13}B: Excitation enegies (left) and energy eigenvalues (right).

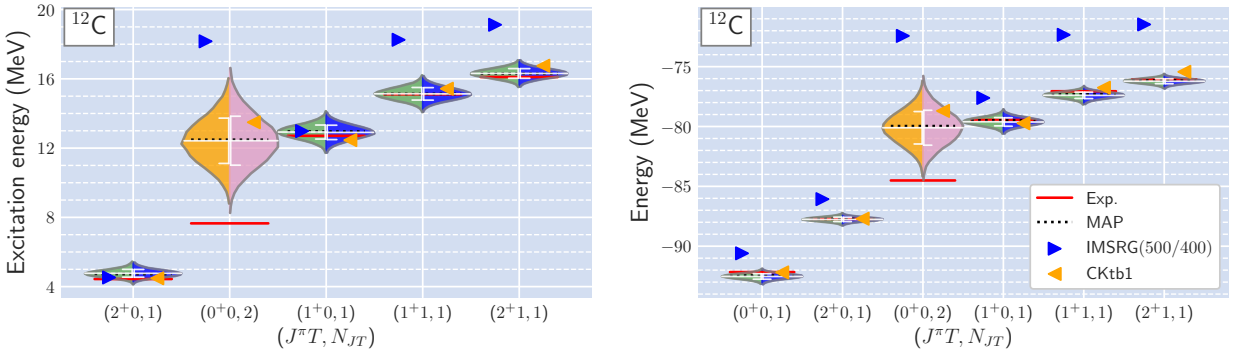

FIG. A.21: Energy spectra for ^{12}C: Excitation enegies (left) and energy eigenvalues (right).

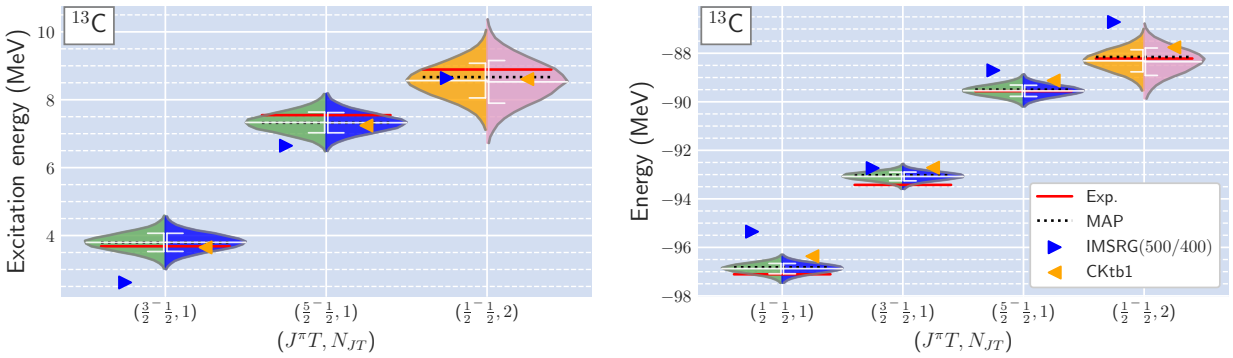

FIG. A.22: Energy spectra for ^{13}C: Excitation enegies (left) and energy eigenvalues (right).

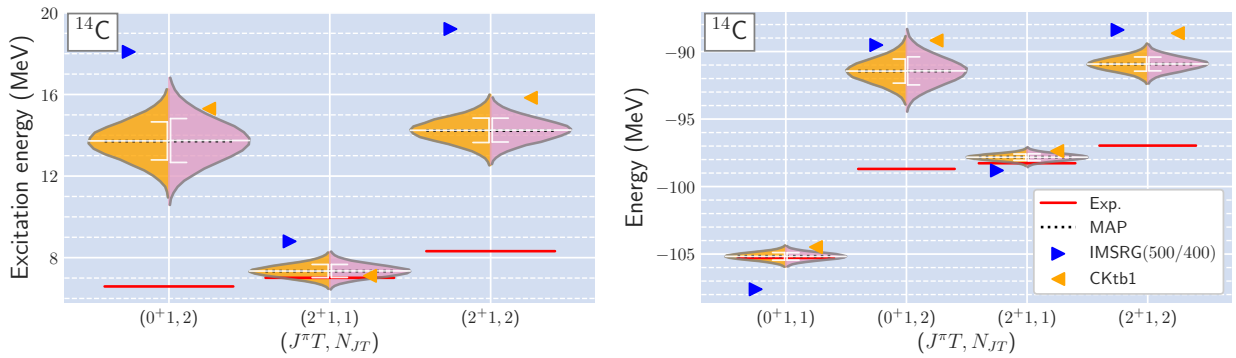

FIG. A.23: Energy spectra for ^{14}C: Excitation enegies (left) and energy eigenvalues (right).

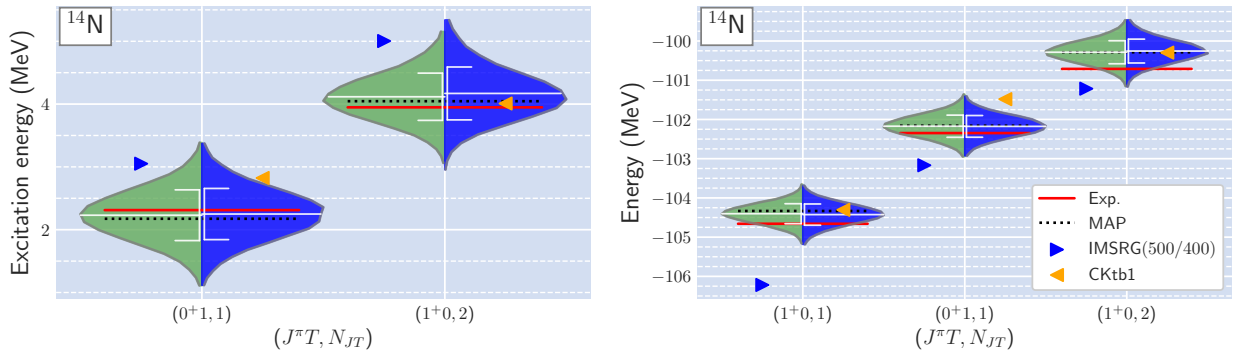

FIG. A.24: Energy spectra for ^{14}N: Excitation enegies (left) and energy eigenvalues (right).

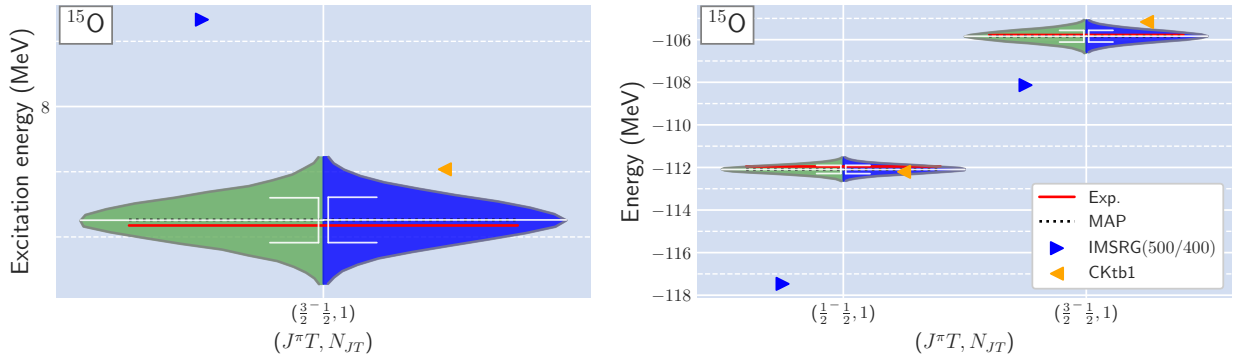

FIG. A.25: Energy spectra for ^{15}O: Excitation enegies (left) and energy eigenvalues (right).

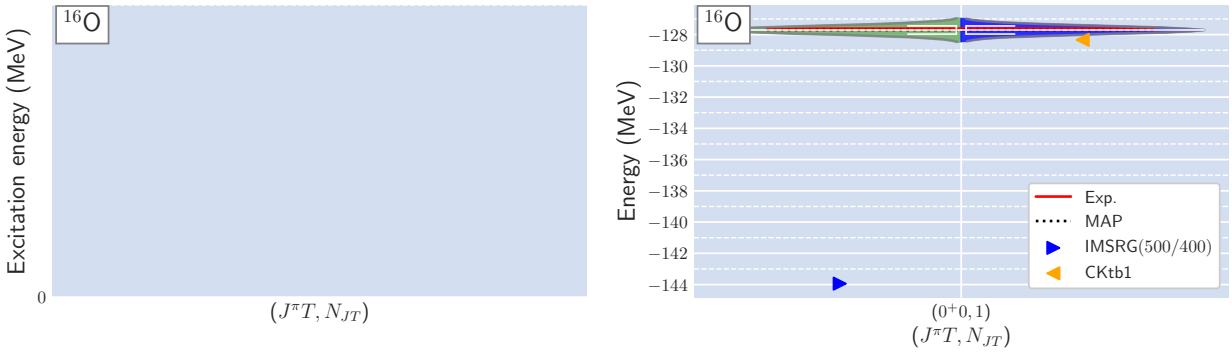

FIG. A.26: Energy spectra for ^{16}O: Excitation enegies (left) and energy eigenvalues (right).

Appendix B

The code and related issues on GPs

In this appendix, we introduce some information of the code used in Chapter 4 and technical issues on implementation of Gaussian Processes (GPs).

The code for constrained Gaussian processes (cGP) [5] is written in Julia language [220]. The marketing catchphrase of the Julia is something like "Efficient as Python and Fast as C/Fortran". The number of users of Julia is increasing rapidly especially in communities such as machine learning, data science, physics, etc. One can achieve fast manipulation of e.g. matrix-matrix product without *magic spells* for optimization and use libraries from Python, R, C-family, Fortran, Java and many handmade packages on GitHub.

We opened the code used in Chapter 4 on GitHub and that uses only the official packages. We would be happy if this code and the technical information below are helpful to someone using GPs.

B.1 Technical issue on Gaussian processes

We, practitioners of Gaussian processes, usually need to calculate conditional mean vectors and covariance matrices for predictions $\{\boldsymbol{y}^*\}$ under the given observations $\{\boldsymbol{y}\}$:

$$P(\boldsymbol{y}^*|\boldsymbol{y},\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}, \Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}), \tag{B.1}$$

$$\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}(\boldsymbol{\theta}) = \boldsymbol{\mu}^* + K_{XX^*}^T K_{XX}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \tag{B.2}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{y}^*|\boldsymbol{y}}(\boldsymbol{\theta}) = K_{X^*X^*} - K_{XX^*}^T K_{XX}^{-1} K_{XX^*}, \tag{B.3}$$

where the $\boldsymbol{\theta}$ is hyperparameter for the kernel function, and X and X^* are the set of observation points and prediction points respectively. The positive semi-definiteness of the covariance

matrices is required to achieve the Cholesky decomposition to evaluate K_{XX}^{-1} and to make samples from the multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{y}^*|\boldsymbol{y}}, \Sigma_{\boldsymbol{y}^*|\boldsymbol{y}})$.

In some cases (e.g., points are located too close to each other), the covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{y}^*|\boldsymbol{y}}(\boldsymbol{\theta})$ and/or K_{XX} become non-PSD numerically, although it must be PSD mathematically. The typical prescription to this is to add infinitesimal diagonal matrix to the K_{XX}. Let us call this *the ϵ prescription* in the followings. For example, in the package GaussianProcesses.jl for Julia [221], the [make_posdef!] function perform this.

The primary motivation of this survey is based on our experience: RBF kernel tends to have singular covariance (i.e. not positive semi-definite), when we write the GP code by ourselves. Then, we thought "Why do many people use RBF kernel in their papers without any problem?"

Here are toy two data sets to be considered:

- set A:
  $X = (6.0, 8.0)$, $Y = (-1.0, 1.0)$, $X^* = (10.0, 12.0)$
- set B:
  $X = (6.0, 8.0, 10.0, 12.0, 14.0)$, $Y = (-28.602, -30.213, -31.176, -31.713, -31.977)$,
  $X^* = (16.0, 18.0, 20.0)$

B.1.1 Who may give the *correct* answer?: set A

Before discussing the PSD property of covariance matrices, we would like to make it clear that the problem below is not due to the our implementation of GPs. The set A is one for that purpose. In this case, Eq. (B.2) and Eq. (B.3) can be analytically calculated under the given hyperparameters.

In the following, we use the three codes listed below:

A [python_gp_own.py]: own code in Python,

B [julia_gp_own.jl]: own code in Julia,

C [julia_gp_library.jl]: using the GaussianProcesses.jl package for Julia.

All are available on GitHub [213].

For the simplicity of discussion, we restrict ourselves to use the RBF kernel:

$$k_{\mathrm{RBF}}(x_i, x_j) = \tau \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right), \tag{B.4}$$

Table B.1: The summary of outputs of the three codes.

	A (own Python)	B (own Julia)	C (Gaussian Processes.jl)
$\boldsymbol{\mu}_{\boldsymbol{y}^*\|\boldsymbol{y}}$	[0.15613, 0.00038795]	[0.15613, 0.00038795]	[0.15613, 0.00038795]
$\mathrm{Diag}(\Sigma_{\boldsymbol{y}^*\|\boldsymbol{y}})$	[0.981355, 0.99999989]	[0.981355, 0.99999989]	[0.981355, 0.99999989]

and we fixed the hyperparameters $\tau = 1.0$ and $\sigma = 1.0$. Since the data values have zero mean, the results are not unexpectedly influenced by a normalization function in the libraries.

The results are summarized in Table B.1. The codes A, B and C gives the same results.

In this case of set A, we can calculate the mean vector Eq. (B.2) and Eq. (B.3) analytically:

$$\mu(x^* = 10) = 1 - \frac{1 - e^{-8} + e^{-12}}{e^4 - 1} \simeq 0.9813550426957166, \tag{B.5}$$

$$\mu(x^* = 12) = 1 - \frac{e^{-12} - 2e^{-24} + e^{-32}}{e^4 - 1} \simeq 0.9999998853666244. \tag{B.6}$$

Now the answer to the title of this subsection is the all codes listed above.

In the GaussianProcesses.jl the default size of observation noise is set as -2.0 in logarithmic scale. We fixed this observation noise as almost zero in the code C. We are afraid that people sometimes use an unrealistically large observation noise in their codes without noticing it.

B.1.2 How to deal with non-PSD covariance in the codes?: set B

Now we explore the PSD property of covariance matrices by using set B. Hereafter, we only use the codes written in Julia.

We use the following three Kernels. One is the RBF Kernel already used in the previous subsection.

$$k_{\mathrm{RBF}}(x_i, x_j) = \tau \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right). \tag{B.7}$$

The other two are Matérn kernels with $\nu = 5/2$

$$k_{\mathrm{M}}(r \equiv |x_i - x_j|; \nu = 5/2) = \tau \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right), \tag{B.8}$$

and with $\nu = 3/2$

$$k_{\mathrm{M}}(r \equiv |x_i - x_j|; \nu = 3/2) = \tau \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right). \tag{B.9}$$

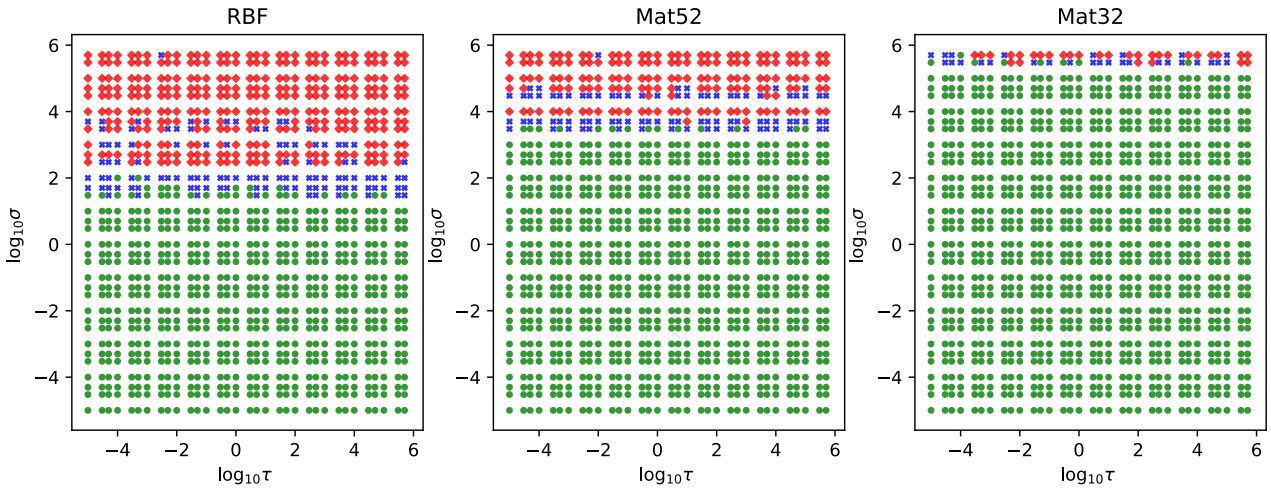

FIG. B.1: The summary of PSD properties for the three Kernels. The covariance matrices having PSD are shown by the filled circle. The diamond symbols mean that both K_{XX} and $\Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}$ are non-PSD, and the cross symbols are assigned if only the $\Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}$ is non-PSD.

We calculated the conditional mean vectors and covariance matrices while varying the hyperparameters τ and σ using our own code in Julia. In Fig. B.1, the positive semi-definiteness of the covariance matrices are summarized. The diamond symbols (red) correspond to the case that both K_{XX} and $\Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}$ are non-PSD, and the cross symbols (blue) are assigned if only the $\Sigma_{\boldsymbol{y}^*|\boldsymbol{y}}$ is non-PSD. This is the reason why the RBF relatively easily gives non-singular covariances in our applications. In terms of the length scale, the RBF gives non-PSD matrices easier than Matérn kernel with $\nu = 5/2$ by two orders of magnitude.

In the GaussianProcesses.jl package, the ϵ prescription is implemented as follows:

```
function make_posdef!(m::AbstractMatrix, chol_factors::AbstractMatrix)
    n = size(m, 1)
    size(m, 2) == n || throw(ArgumentError("Covariance matrix must be square"))
    for $_$ in 1:10
        try
            copyto!(chol_factors, m)
            chol = cholesky!(Symmetric(chol_factors, :U))
            return m, chol
        catch err
            if typeof(err)!=LinearAlgebra.PosDefException
                throw(err)
            end
            ϵ = 1e-6 * tr(m) / n
            @inbounds for i in 1:n
                m[i, i] += ϵ
            end
```

```
        end
    end
    copyto!(chol_factors, m)
    chol = cholesky!(Symmetric(chol_factors, :U))
    return m, chol
end
```

One can see that if the code fails to achieve the Cholesky decomposition of a matrix K, this function do

$$K := K + \epsilon I_n, \tag{B.10}$$

$$\epsilon = 1.\mathrm{e}-6 \times \mathrm{Tr}(K)/n. \tag{B.11}$$

As shown later, this procedure sometimes induces significant changes in resultant mean vectors.

Now we show the effects of the ϵ prescription on the predictions. In FIG. B.2 and FIG. B.3, the following quantity is shown as the so-called heatmap:

$$\Delta \equiv \log_{10}\left(max(|\boldsymbol{\mu}^{\mathrm{own}}_{\boldsymbol{y}^*|\boldsymbol{y}} - \boldsymbol{\mu}^{\mathrm{lib.}}_{\boldsymbol{y}^*|\boldsymbol{y}}|)\right), \tag{B.12}$$

where the superscripts own and lib. mean the conditional mean vector calculated by our own code and one using the GaussianProcesses.jl library, respectively. FIG. B.3 shows the Δ in the case of Kernels defined by the logarithm distance, i.e. $|x_i - x_j|$ is replaced by $|\ln x_i - \ln x_j|$.

The hatched region at $\log_{10}\sigma < 1$ means that the results with the two codes are exactly the same. At the points shown by the diamond symbol in FIG. B.1, our code without the ϵ prescription cannot not give the mean vectors. Those correspond to the white region $\log_{10}\sigma > 1.5$ appeared in FIG. B.2 and FIG. B.3. When we go to the parameters at which the posterior covariance become non-PSD, i.e. points shown by the cross symbol in FIG. B.1, the Δ become large because of the *immune system* in the library. In some cases, the deviations in mean vectors reach $10^3 - 10^4$, which is obviously non-negligible. As immune systems in vivo, the ϵ prescription sometimes reacts too much.

One should pay much attention to how the PSD is treated in codes and whether or not one really can neglect the impact of the prescriptions on the predictions, especially when one would like to integrate out the hyperparameter dependence.

Here we make a remark from physics aspect. In the current case, we assume the y and y^* values to be calculated energies of nuclei by *ab initio* full configuration interaction method. The typical convergence tolerance of the Lanczos method in shell-model codes on the market is 1.e-5 or better[1].

[1] In the BIGSTICK [8] and the KSHELL [12,13], tol = 1.e-5 as default.

When one would like to incorporate this "tol" information into the kernel as the observation noise[2] like $K+\sigma^2_{\text{obs.}}I$, the ϵ in the package is obviously too large for the current purpose and this must be on the order of 1.e−10. As a result, the possible error coming from the ϵ prescription in the package can be non-negligible as shown in FIG. B.2 and FIG. B.3.

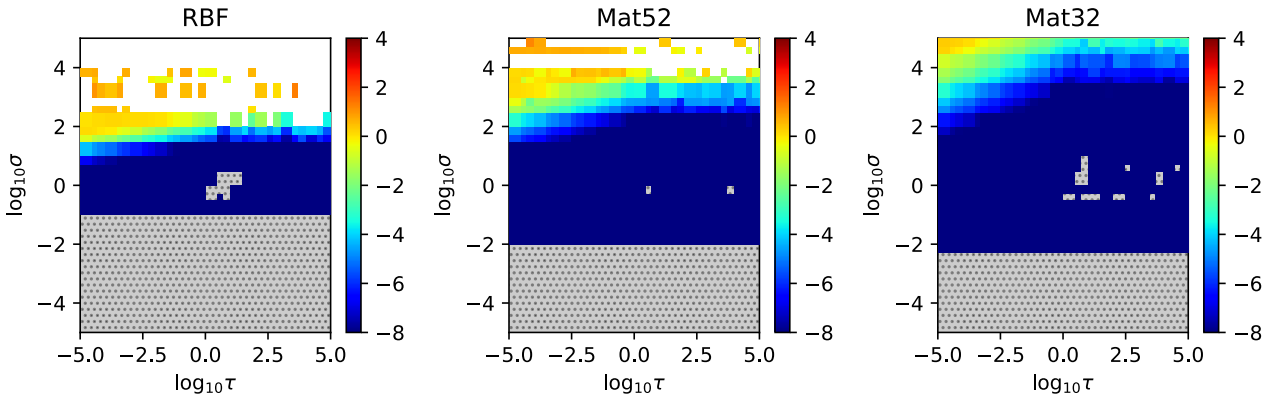

FIG. B.2: The plot showing the impact of the ϵ prescription to avoid non-PSD. The colormap shows the Δ in Eq. (B.12). The hatched region (gray) show the points at which the two codes gives the exactly the same result. The region where the self-made code does not give answer is colored in white.

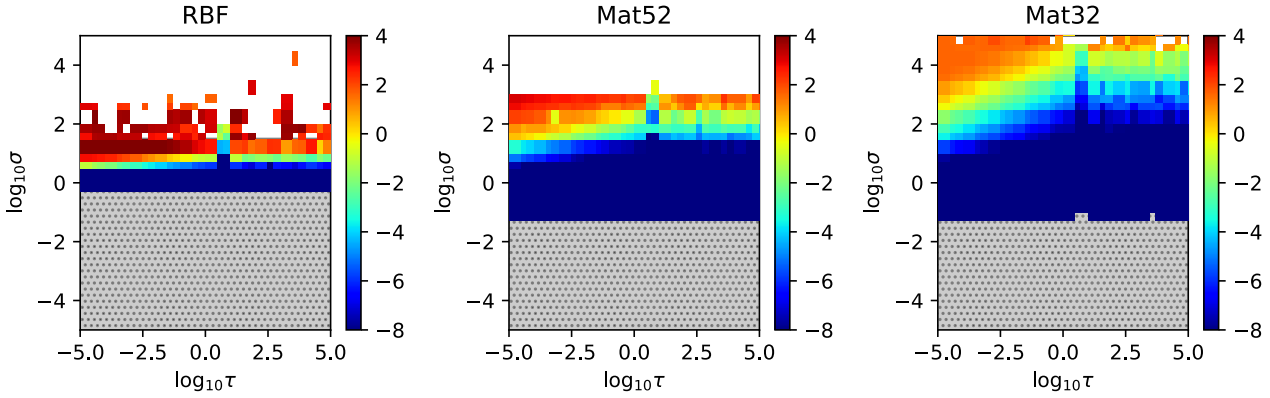

FIG. B.3: The counterpart of FIG. B.2 for the logarithm distance.

B.1.3 Summary

We conclude that the Matérn kernel is safer to use than RBF in terms of positive semi-definiteness in many practical situations, and, the ϵ prescription with **any** Kernel function is **not always** acceptable. One should be careful to the prescription done in the codes, especially when using some packages.

[2]In this context, the term *observation* means doing calculation instead of experimental observation.

Bibliography

[1] T. Editors, Phys. Rev. A **83**, 040001 (2011).

[2] J. Dobaczewski, W. Nazarewicz, and P.-G. Reinhard, Journal of Physics G: Nuclear and Particle Physics **41**, 074001 (2014).

[3] J. Piekarewicz, W.-C. Chen, and F. J. Fattoyev, Journal of Physics G: Nuclear and Particle Physics **42**, 034018 (2015).

[4] S. Yoshida, GitHub repository:ShellModelUncertaintyQuantification `https://github.com/SotaYoshida/ShellModelUncertaintyQuantification` .

[5] S. Yoshida, GitHub repository:Bayesian Constrained Gaussian Proocess model for Extrapolations in CI methods `https://github.com/SotaYoshida/BCGP_ECI`.

[6] C. Lanczos, J. Res. Natl. Bur. Stand. B **45**, 255 (1950).

[7] E. Caurier and F. Nowacki, Acta Physica Polonica B **30**, 705 (1999).

[8] C. W. Johnson, W. E. Ormand, K. S. McElvain, and H. Shan, Bigstick: A flexible configuration-interaction shell-model code (2018), arXiv:1801.08432 .

[9] J. P. Vary, The many-fermion dynamics shell-model code, unpublished (1992).

[10] J. P. Vary and D. C. Zheng, The many-fermion dynamics shell-model code, unpublished (1994).

[11] P. Sternberg, E. G. Ng, C. Yang, P. Maris, J. P. Vary, M. Sosonkina, and H. V. Le, in Proceedings of the 2008 ACM/IEEE Conference on Supercomputing, SC '08 (IEEE Press, Piscataway, NJ, USA, 2008) pp. 15:1–15:12.

[12] N. Shimizu, Nuclear shell-model code for massive parallel computation, "kshell" (2013), arXiv:1310.5431 .

[13] N. Shimizu, T. Mizusaki, Y. Utsuno, and Y. Tsunoda, Computer Physics Communications **244**, 372 (2019).

[14] B. Brown and W. Rae, Nuclear Data Sheets **120**, 115 (2014).

[15] T. Otsuka, M. Honma, T. Mizusaki, N. Shimizu, and Y. Utsuno, Progress in Particle and Nuclear Physics **47**, 319 (2001).

[16] T. Otsuka, Nuclear Physics A **693**, 383 (2001), radioactive Nuclear Beams.

[17] S. Pittel and N. Sandulescu, Phys. Rev. C **73**, 014301 (2006).

[18] O. Legeza, L. Veis, A. Poves, and J. Dukelsky, Phys. Rev. C **92**, 051303 (2015).

[19] G. Puddu, Journal of Physics G: Nuclear and Particle Physics **32**, 321 (2006).

[20] G. Puddu, The European Physical Journal A **34**, 413 (2007).

[21] Y. Tsunoda, T. Otsuka, N. Shimizu, M. Honma, and Y. Utsuno, Phys. Rev. C **89**, 031301 (2014).

[22] N. Shimizu, Y. Utsuno, T. Mizusaki, M. Honma, Y. Tsunoda, and T. Otsuka, Phys. Rev. C **85**, 054301 (2012).

[23] N. Shimizu, T. Abe, Y. Tsunoda, Y. Utsuno, T. Yoshida, T. Mizusaki, M. Honma, and T. Otsuka, Progress of Theoretical and Experimental Physics 2012, 10.1093/ptep/pts012 (2012).

[24] N. Shimizu, T. Abe, M. Honma, T. Otsuka, T. Togashi, Y. Tsunoda, Y. Utsuno, and T. Yoshida, Physica Scripta **92**, 063001 (2017).

[25] T. Otsuka, Y. Tsunoda, T. Abe, N. Shimizu, and P. Van Duppen, Phys. Rev. Lett. **123**, 222502 (2019).

[26] R. Roth and P. Navrátil, Phys. Rev. Lett. **99**, 092501 (2007).

[27] D. J. Dean, G. Hagen, M. Hjorth-Jensen, T. Papenbrock, and A. Schwenk, Phys. Rev. Lett. **101**, 119201 (2008).

[28] R. Roth and P. Navrátil, Phys. Rev. Lett. **101**, 119202 (2008).

[29] R. Roth, Phys. Rev. C **79**, 064324 (2009).

[30] R. Roth, J. Langhammer, A. Calci, S. Binder, and P. Navrátil, Phys. Rev. Lett. **107**, 072501 (2011).

[31] C. Stumpf, J. Braun, and R. Roth, Phys. Rev. C **93**, 021301 (2016).

[32] N. Shimizu, Y. Utsuno, T. Mizusaki, T. Otsuka, T. Abe, and M. Honma, Phys. Rev. C **82**, 061305 (2010).

[33] S. Weinberg, Physics Letters B **251**, 288 (1990).

[34] S. Weinberg, Nuclear Physics B **363**, 3 (1991).

[35] R. Machleidt and D. Entem, Phys. Rept. **503**, 1 (2011).

[36] A. Nogga, R. G. E. Timmermans, and U. van Kolck, Phys. Rev. C **72**, 054006 (2005).

[37] E. Epelbaum and J. Gegelia, The European Physical Journal A **41**, 341 (2009).

[38] M. C. Birse, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **369**, 2662 (2011).

[39] B. Long and C.-J. Yang, Phys. Rev. C **85**, 034002 (2012).

[40] E. Epelbaum and J. Gegelia, Physics Letters B **716**, 338 (2012).

[41] E. Epelbaum and U.-G. Meißner, Few-Body Systems **54**, 2175 (2013).

[42] D. R. Entem and R. Machleidt, Phys. Rev. C **66**, 014002 (2002).

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

[44] E. Marji, A. Canul, Q. MacPherson, R. Winzer, C. Zeoli, D. R. Entem, and R. Machleidt, Phys. Rev. C **88**, 054002 (2013).

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

[46] D. R. Entem, N. Kaiser, R. Machleidt, and Y. Nosyk, Phys. Rev. C **92**, 064001 (2015).

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

[48] E. Epelbaum, A. Nogga, W. Glöckle, H. Kamada, U.-G. Meißner, and H. Witała, Phys. Rev. C **66**, 064001 (2002).

[49] E. Epelbaum, W. Glöckle, and U.-G. Meißner, The European Physical Journal A - Hadrons and Nuclei **19**, 125 (2004).

[50] E. Epelbaum, W. Glöckle, and U.-G. Meißner, The European Physical Journal A - Hadrons and Nuclei **19**, 401 (2004).

[51] E. Epelbaum, W. Glöckle, and U.-G. Meißner, Nuclear Physics A **747**, 362 (2005).

[52] E. Epelbaum, Progress in Particle and Nuclear Physics **57**, 654 (2006).

[53] V. Bernard, E. Epelbaum, H. Krebs, and U.-G. Meißner, Phys. Rev. C **77**, 064004 (2008).

[54] E. Epelbaum, H.-W. Hammer, and U.-G. Meißner, Rev. Mod. Phys. **81**, 1773 (2009).

[55] V. Bernard, E. Epelbaum, H. Krebs, and U.-G. Meißner, Phys. Rev. C **84**, 054001 (2011).

[56] E. Epelbaum, H. Krebs, D. Lee, and U.-G. Meißner, Phys. Rev. Lett. **106**, 192501 (2011).

[57] V. Bernard, E. Epelbaum, H. Krebs, and U.-G. Meißner, Phys. Rev. C **84**, 054001 (2011).

[58] H. Krebs, A. Gasparyan, and E. Epelbaum, Phys. Rev. C **85**, 054006 (2012).

[59] H. Krebs, A. Gasparyan, and E. Epelbaum, Phys. Rev. C **87**, 054007 (2013).

[60] E. Epelbaum, H. Krebs, and U.-G. Meißner, Phys. Rev. Lett. **115**, 122301 (2015).

[61] E. Epelbaum, H. Krebs, and U. G. Meißner, The European Physical Journal A **51**, 53 (2015).

[62] E. Epelbaum, J. Golak, K. Hebeler, T. Hüther, H. Kamada, H. Krebs, P. Maris, U.-G. Meißner, A. Nogga, R. Roth, R. Skibiński, K. Topolnicki, J. P. Vary, K. Vobig, and H. Witała (LENPIC Collaboration), Phys. Rev. C **99**, 024313 (2019).

[63] A. Nogga, P. Navrátil, B. R. Barrett, and J. P. Vary, Phys. Rev. C **73**, 064002 (2006).

[64] K. Hebeler, S. K. Bogner, R. J. Furnstahl, A. Nogga, and A. Schwenk, Phys. Rev. C **83**, 031301 (2011).

[65] P. Navrátil, V. G. Gueorguiev, J. P. Vary, W. E. Ormand, and A. Nogga, Phys. Rev. Lett. **99**, 042501 (2007).

[66] A. Gårdestig and D. R. Phillips, Phys. Rev. Lett. **96**, 232301 (2006).

[67] D. Gazit, S. Quaglioni, and P. Navrátil, Phys. Rev. Lett. **103**, 102502 (2009).

[68] D. Gazit, S. Quaglioni, and P. Navrátil, Phys. Rev. Lett. **122**, 029901 (2019).

[69] L. E. Marcucci, A. Kievsky, S. Rosati, R. Schiavilla, and M. Viviani, Phys. Rev. Lett. **108**, 052502 (2012).

[70] L. E. Marcucci, A. Kievsky, S. Rosati, R. Schiavilla, and M. Viviani, Phys. Rev. Lett. **121**, 049901 (2018).

[71] J. Simonis, S. R. Stroberg, K. Hebeler, J. D. Holt, and A. Schwenk, Phys. Rev. C **96**, 014303 (2017).

[72] T. D. Morris, J. Simonis, S. R. Stroberg, C. Stumpf, G. Hagen, J. D. Holt, G. R. Jansen, T. Papenbrock, R. Roth, and A. Schwenk, Phys. Rev. Lett. **120**, 152503 (2018).

[73] C. Drischler, K. Hebeler, and A. Schwenk, Phys. Rev. Lett. **122**, 042501 (2019).

[74] J. Hoppe, C. Drischler, K. Hebeler, A. Schwenk, and J. Simonis, Phys. Rev. C **100**, 024318 (2019).

[75] A. Ekström, G. Baardsen, C. Forssén, G. Hagen, M. Hjorth-Jensen, G. R. Jansen, R. Machleidt, W. Nazarewicz, T. Papenbrock, J. Sarich, and S. M. Wild, Phys. Rev. Lett. **110**, 192502 (2013).

[76] A. Ekström, G. R. Jansen, K. A. Wendt, G. Hagen, T. Papenbrock, B. D. Carlsson, C. Forssén, M. Hjorth-Jensen, P. Navrátil, and W. Nazarewicz, Phys. Rev. C **91**, 051301 (2015).

[77] B. D. Carlsson, A. Ekström, C. Forssén, D. F. Strömberg, G. R. Jansen, O. Lilja, M. Lindby, B. A. Mattsson, and K. A. Wendt, Phys. Rev. X **6**, 011019 (2016).

[78] H. Hergert, J. M. Yao, T. D. Morris, N. M. Parzuchowski, S. K. Bogner, and J. Engel, Journal of Physics: Conference Series **1041**, 012007 (2018).

[79] A. M. Shirokov, A. I. Mazur, S. A. Zaytsev, J. P. Vary, and T. A. Weber, Phys. Rev. C **70**, 044005 (2004).

[80] A. Shirokov, J. Vary, A. Mazur, S. Zaytsev, and T. Weber, Physics Letters B **621**, 96 (2005).

[81] A. Shirokov, J. Vary, A. Mazur, and T. Weber, Physics Letters B **644**, 33 (2007).

[82] A. Shirokov, I. Shin, Y. Kim, M. Sosonkina, P. Maris, and J. Vary, Physics Letters B **761**, 87 (2016).

[83] Y. Kim, I. J. Shin, A. M. Shirokov, M. Sosonkina, P. Maris, and J. P. Vary, Daejeon16 nn interaction, arXiv:1910.04367 (2019).

[84] J. W. Holt, N. Kaiser, and W. Weise, Phys. Rev. C **81**, 024002 (2010).

[85] K. Hebeler and A. Schwenk, Phys. Rev. C **82**, 014314 (2010).

[86] M. Kohno, Phys. Rev. C **88**, 064005 (2013).

[87] M. Kohno, Phys. Rev. C **96**, 059903 (2017).

[88] S. Yoshida, Nuclear shell model based on chiral effective field theory, Master's thesis, The University of Tokyo (2017).

[89] L. Coraggio, J. W. Holt, N. Itaco, R. Machleidt, L. E. Marcucci, and F. Sammarruca, Phys. Rev. C **89**, 044321 (2014).

[90] K. Minomo, M. Toyokawa, M. Kohno, and M. Yahiro, Phys. Rev. C **90**, 051601 (2014).

[91] K. Minomo, M. Toyokawa, M. Kohno, and M. Yahiro, Phys. Rev. C **96**, 059904 (2017).

[92] R. J. Furnstahl, N. Klco, D. R. Phillips, and S. Wesolowski, Phys. Rev. C **92**, 024005 (2015).

[93] R. J. Furnstahl, D. R. Phillips, and S. Wesolowski, Journal of Physics G: Nuclear and Particle Physics **42**, 034028 (2015).

[94] S. Wesolowski, N. Klco, R. J. Furnstahl, D. R. Phillips, and A. Thapaliya, Journal of Physics G: Nuclear and Particle Physics **43**, 074001 (2016).

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

[96] S. Wesolowski, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips, Journal of Physics G: Nuclear and Particle Physics **46**, 045102 (2019).

[97] J. A. Melendez, R. J. Furnstahl, D. R. Phillips, M. T. Pratola, and S. Wesolowski, Phys. Rev. C **100**, 044001 (2019).

[98] R. B. Wiringa, V. G. J. Stoks, and R. Schiavilla, Phys. Rev. C **51**, 38 (1995).

[99] S. D. Głazek and K. G. Wilson, Phys. Rev. D **48**, 5863 (1993).

[100] F. Wegner, Annalen der Physik **506**, 77 (1994).

[101] S. R. White, The Journal of Chemical Physics **117**, 7472 (2002).

[102] H. Hergert, S. Bogner, T. Morris, A. Schwenk, and K. Tsukiyama, Physics Reports **621**, 165 (2016), memorial Volume in Honor of Gerald E. Brown.

[103] W. Li, E. R. Anderson, and R. J. Furnstahl, Phys. Rev. C **84**, 054002 (2011).

[104] N. M. Dicaire, C. Omand, and P. Navrátil, Phys. Rev. C **90**, 034302 (2014).

[105] J. Dawson, I. Talmi, and J. Walecka, Annals of Physics **18**, 339 (1962).

[106] G. F. Bertsch, Nuclear Physics **74**, 234 (1965).

[107] T. Kuo and G. Brown, Nuclear Physics **85**, 40 (1966).

[108] T. Kuo, Nuclear Physics A **90**, 199 (1967).

[109] T. Kuo and G. Brown, Nuclear Physics A **114**, 241 (1968).

[110] S. Cohen and D. Kurath, Nuclear Physics **73**, 1 (1965).

[111] B. Wildenthal, Progress in Particle and Nuclear Physics **11**, 5 (1984).

[112] B. A. Brown and B. H. Wildenthal, Annual Review of Nuclear and Particle Science **38**, 29 (1988).

[113] E. K. Warburton, J. A. Becker, and B. A. Brown, Phys. Rev. C **41**, 1147 (1990).

[114] E. Caurier, G. Martínez-Pinedo, F. Nowacki, A. Poves, J. Retamosa, and A. P. Zuker, Phys. Rev. C **59**, 2033 (1999).

[115] B. A. Brown and W. A. Richter, Phys. Rev. C **74**, 034315 (2006).

[116] M. Honma, T. Otsuka, B. A. Brown, and T. Mizusaki, Phys. Rev. C **65**, 061301 (2002).

[117] M. Honma, T. Otsuka, B. A. Brown, and T. Mizusaki, Phys. Rev. C **69**, 034335 (2004).

[118] M. Honma, T. Otsuka, B. A. Brown, and T. Mizusaki, The European Physical Journal A - Hadrons and Nuclei **25**, 499 (2005).

[119] P. J. Ellis and E. Osnes, Rev. Mod. Phys. **49**, 777 (1977).

[120] M. Hjorth-Jensen, T. T. Kuo, and E. Osnes, Physics Reports **261**, 125 (1995).

[121] N. Tsunoda, T. Otsuka, K. Tsukiyama, and M. Hjorth-Jensen, Phys. Rev. C **84**, 044322 (2011).

[122] J. D. Holt, T. Otsuka, A. Schwenk, and T. Suzuki, Journal of Physics G: Nuclear and Particle Physics **39**, 085111 (2012).

[123] J. D. Holt, J. Menéndez, and A. Schwenk, Phys. Rev. Lett. **110**, 022502 (2013).

[124] J. D. Holt, J. Menéndez, and A. Schwenk, The European Physical Journal A **49**, 39 (2013).

[125] N. Tsunoda, K. Takayanagi, M. Hjorth-Jensen, and T. Otsuka, Phys. Rev. C **89**, 024313 (2014).

[126] L. Coraggio, A. Covello, A. Gargano, N. Itaco, and T. T. S. Kuo, Phys. Rev. C **91**, 041301 (2015).

[127] L. Coraggio, A. Gargano, and N. Itaco, Phys. Rev. C **93**, 064328 (2016).

[128] J. Simonis, K. Hebeler, J. D. Holt, J. Menéndez, and A. Schwenk, Phys. Rev. C **93**, 011302 (2016).

[129] N. Tsunoda, T. Otsuka, N. Shimizu, M. Hjorth-Jensen, K. Takayanagi, and T. Suzuki, Phys. Rev. C **95**, 021304 (2017).

[130] S. Yoshida, M. Kohno, T. Abe, T. Otsuka, N. Tsunoda, and N. Shimizu, Shell-model hamiltonians from chiral forces, in Proceedings of the Ito International Research Center Symposium "Perspectives of the Physics of Nuclear Structure" (JPS Conference Proceedings, 2018) p. 013014.

[131] T. Fukui, L. De Angelis, Y. Z. Ma, L. Coraggio, A. Gargano, N. Itaco, and F. R. Xu, Phys. Rev. C **98**, 044305 (2018).

[132] Y. Z. Ma, L. Coraggio, L. De Angelis, T. Fukui, A. Gargano, N. Itaco, and F. R. Xu, Phys. Rev. C **100**, 034324 (2019).

[133] G. R. Jansen, J. Engel, G. Hagen, P. Navratil, and A. Signoracci, Phys. Rev. Lett. **113**, 142502 (2014).

[134] G. R. Jansen, M. D. Schuster, A. Signoracci, G. Hagen, and P. Navrátil, Phys. Rev. C **94**, 011301 (2016).

[135] Z. H. Sun, T. D. Morris, G. Hagen, G. R. Jansen, and T. Papenbrock, Phys. Rev. C **98**, 054320 (2018).

[136] S. K. Bogner, H. Hergert, J. D. Holt, A. Schwenk, S. Binder, A. Calci, J. Langhammer, and R. Roth, Phys. Rev. Lett. **113**, 142501 (2014).

[137] S. R. Stroberg, H. Hergert, J. D. Holt, S. K. Bogner, and A. Schwenk, Phys. Rev. C **93**, 051301 (2016).

[138] S. R. Stroberg, A. Calci, H. Hergert, J. D. Holt, S. K. Bogner, R. Roth, and A. Schwenk, Phys. Rev. Lett. **118**, 032502 (2017).

[139] K. Tsukiyama, S. K. Bogner, and A. Schwenk, Phys. Rev. Lett. **106**, 222502 (2011).

[140] G. R. Jansen, Phys. Rev. C **88**, 024305 (2013).

[141] A. Ekström, G. R. Jansen, K. A. Wendt, G. Hagen, T. Papenbrock, S. Bacca, B. Carlsson, and D. Gazit, Phys. Rev. Lett. **113**, 262504 (2014).

[142] K. Suzuki and R. Okamoto, Progress of Theoretical Physics **92**, 1045 (1994).

[143] T. Miyagi, T. Abe, R. Okamoto, and T. Otsuka, Phys. Rev. C **96**, 054312 (2017).

[144] T. Miyagi, T. Abe, M. Kohno, P. Navrátil, R. Okamoto, T. Otsuka, N. Shimizu, and S. R. Stroberg, Phys. Rev. C **100**, 034310 (2019).

[145] V. Somà, C. Barbieri, and T. Duguet, Phys. Rev. C **87**, 011303 (2013).

[146] H. Hergert, S. Binder, A. Calci, J. Langhammer, and R. Roth, Phys. Rev. Lett. **110**, 242501 (2013).

[147] H. Hergert, S. K. Bogner, T. D. Morris, S. Binder, A. Calci, J. Langhammer, and R. Roth, Phys. Rev. C **90**, 041302 (2014).

[148] K. Tsukiyama, S. K. Bogner, and A. Schwenk, Phys. Rev. C **85**, 061304 (2012).

[149] G. Hagen, T. Papenbrock, M. Hjorth-Jensen, and D. J. Dean, Reports on Progress in Physics **77**, 096302 (2014).

[150] L. Huth, V. Durant, J. Simonis, and A. Schwenk, Phys. Rev. C **98**, 044301 (2018).

[151] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, Handbook of Markov Chain Monte Carlo, Chapman & Hall/CRC Handbooks of Modern Statistical Methods (CRC press, 2011).

[152] C. Andrieu and J. Thoms, Statistics and Computing **18**, 343 (2008).

[153] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, Physics Letters B **195**, 216 (1987).

[154] Stan, state-of-the-art platform for statistical modeling and high-performance statistical computation `https://mc-stan.org`.

[155] PyMC, python module that implements Bayesian statistical models and fitting algorithms, including Markov chain Monte Carlo. `http://pymc-devs.github.io/pymc/`.

[156] R. M. Neal, ArXiv e-prints (2012), arXiv:1206.1901 .

[157] A. Arima and M. Ichimura, Unpublished.

[158] C. M. Bishop, Pattern Recognition and Machine Learning (Springer, 2006).

[159] B. A. Brown, Progress in Particle and Nuclear Physics **47**, 517 (2001).

[160] E. Caurier, G. Martínez-Pinedo, F. Nowacki, A. Poves, and A. P. Zuker, Rev. Mod. Phys. **77**, 427 (2005).

[161] T. Otsuka, A. Gade, O. Sorlin, T. Suzuki, and Y. Utsuno, Evolution of nuclear structure in exotic nuclei and nuclear forces (2018), arXiv:1805.06501 .

[162] S. R. Stroberg, H. Hergert, S. K. Bogner, and J. D. Holt, Annual Review of Nuclear and Particle Science **69**, 307 (2019).

[163] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization (2014), arXiv:1412.6980 .

[164] ENSDF database, `http://www.nndc.bnl.gov/ensdf/`.

[165] E. K. Warburton and B. A. Brown, Phys. Rev. C **46**, 923 (1992).

[166] M. Chernykh, H. Feldmeier, T. Neff, P. von Neumann-Cosel, and A. Richter, Phys. Rev. Lett. **98**, 032501 (2007).

[167] Y. Kanada-En'yo, Progress of Theoretical Physics **117**, 655 (2007).

[168] Y. Kanada-En'yo, Progress of Theoretical Physics **121**, 895 (2009).

[169] P. Maris, J. P. Vary, and A. M. Shirokov, Phys. Rev. C **79**, 014308 (2009).

[170] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, Rev. Mod. Phys. **87**, 1067 (2015).

[171] H. T. Fortune, M. E. Cobern, S. Mordechai, G. E. Moore, S. Lafrance, and R. Middleton, Phys. Rev. Lett. **40**, 1236 (1978).

[172] A. A. Wolters, A. G. M. van Hees, and P. W. M. Glaudemans, Phys. Rev. C **42**, 2062 (1990).

[173] P. Navrátil, J. P. Vary, and B. R. Barrett, Phys. Rev. Lett. **84**, 5728 (2000).

[174] D. Arthur and S. Vassilvitskii, In Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (2007) 1027-1035 .

[175] J. Dohet-Eraly, P. Navrátil, S. Quaglioni, W. Horiuchi, G. Hupin, and F. Raimondi, Physics Letters B **757**, 430 (2016).

[176] J. M. R. Fox and C. W. Johnson, Uncertainty quantification of an empirical shell-model interaction using principal component analysis (2019), arXiv:1911.05208 .

[177] S. Yoshida, N. Shimizu, T. Togashi, and T. Otsuka, Phys. Rev. C **98**, 061301 (2018).

[178] T. Suzuki, R. Fujimoto, and T. Otsuka, Phys. Rev. C **67**, 044302 (2003).

[179] C. Yuan, T. Suzuki, T. Otsuka, F. Xu, and N. Tsunoda, Phys. Rev. C **85**, 064324 (2012).

[180] P. Maris, J. P. Vary, P. Navrátil, W. E. Ormand, H. Nam, and D. J. Dean, Phys. Rev. Lett. **106**, 202502 (2011).

[181] H. Morita and Y. Kanada-En'yo, Phys. Rev. C **96**, 044318 (2017).

[182] S. Yoshida, Non-parametric Bayesian approach to extrapolation problems in configuration interaction methods (2019), arXiv:1907.04974 .

[183] B. R. Barrett, P. Navrátil, and J. P. Vary, Progress in Particle and Nuclear Physics **69**, 131 (2013).

[184] P. Navrátil, S. Quaglioni, I. Stetcu, and B. R. Barrett, Journal of Physics G: Nuclear and Particle Physics **36**, 083101 (2009).

[185] I. J. Shin, Y. Kim, P. Maris, J. P. Vary, C. Forssén, J. Rotureau, and N. Michel, Journal of Physics G: Nuclear and Particle Physics **44**, 075103 (2017).

[186] M. K. G. Kruse, E. D. Jurgenson, P. Navrátil, B. R. Barrett, and W. E. Ormand, Phys. Rev. C **87**, 044301 (2013).

[187] P. Maris, J. P. Vary, and A. M. Shirokov, Phys. Rev. C **79**, 014308 (2009).

[188] H. Zhan, A. Nogga, B. R. Barrett, J. P. Vary, and P. Navrátil, Phys. Rev. C **69**, 034302 (2004).

[189] S. A. Coon, M. I. Avetian, M. K. G. Kruse, U. van Kolck, P. Maris, and J. P. Vary, Phys. Rev. C **86**, 054002 (2012).

[190] R. J. Furnstahl, G. Hagen, and T. Papenbrock, Phys. Rev. C **86**, 031301 (2012).

[191] S. N. More, A. Ekström, R. J. Furnstahl, G. Hagen, and T. Papenbrock, Phys. Rev. C **87**, 044326 (2013).

[192] K. A. Wendt, C. Forssén, T. Papenbrock, and D. Sääf, Phys. Rev. C **91**, 061301 (2015).

[193] C. Forssén, B. D. Carlsson, H. T. Johansson, D. Sääf, A. Bansal, G. Hagen, and T. Papenbrock, Phys. Rev. C **97**, 034328 (2018).

[194] C. Stumpf, T. Wolfgruber, and R. Roth, Electromagnetic strength distributions from the ab initio no-core shell model (2017), arXiv:1709.06840 .

[195] T. Abe, P. Maris, T. Otsuka, N. Shimizu, Y. Utsuno, and J. P. Vary, Phys. Rev. C **86**, 054301 (2012).

[196] L. Liu, T. Otsuka, N. Shimizu, Y. Utsuno, and R. Roth, Phys. Rev. C **86**, 014302 (2012).

[197] T. Togashi, Y. Tsunoda, T. Otsuka, and N. Shimizu, Phys. Rev. Lett. **117**, 172502 (2016).

[198] G. A. Negoita, J. P. Vary, G. R. Luecke, P. Maris, A. M. Shirokov, I. J. Shin, Y. Kim, E. G. Ng, C. Yang, M. Lockner, and G. M. Prabhu, Phys. Rev. C **99**, 054308 (2019).

[199] W. G. Jiang, G. Hagen, and T. Papenbrock, Phys. Rev. C **100**, 054326 (2019).

[200] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning) (The MIT Press, 2005).

[201] L. Neufcourt, Y. Cao, W. Nazarewicz, and F. Viens, Phys. Rev. C **98**, 034318 (2018).

[202] R. A. Vargas-Hernández, J. Sous, M. Berciu, and R. V. Krems, Phys. Rev. Lett. **121**, 255702 (2018).

[203] S. Das, G. Siopsis, and C. Weedbrook, Phys. Rev. A **97**, 022315 (2018).

[204] D. Wee, J. Kim, S. Bang, G. Samsonidze, and B. Kozinsky, Phys. Rev. Materials **3**, 033803 (2019).

[205] L. Neufcourt, Y. Cao, W. Nazarewicz, E. Olsen, and F. Viens, Phys. Rev. Lett. **122**, 062502 (2019).

[206] S. Zhou, P. Giulani, J. Piekarewicz, A. Bhattacharya, and D. Pati, Phys. Rev. C **99**, 055202 (2019).

[207] R. M. Neal, Bayesian Learning for Neural Networks, Ph.D. thesis, University of Toronto (1995).

[208] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein, Deep neural networks as gaussian processes (2017), arXiv:1711.00165 .

[209] D. Mochihashi and S. Oba, Gaussian process and machine learning (In Japanese) (2019).

[210] B. Matérn, Spatial variation, Lecture notes in statistics (Springer-Verlag, 1986).

[211] M. L. Stein, Interpolation of spatial data, Springer Series in Statistics (Springer-Verlag, New York, 1999).

[212] Z. Noumir, P. Honeine, and C. Richard, 2012 IEEE International Workshop on Machine Learning for Signal Processing 10.1109/MLSP.2012.6349753 (2012).

[213] S. Yoshida, GitHub repository:IssueOnGPs `https://github.com/SotaYoshida/IssueOnGPs` .

[214] N. Gordon, D. Salmond, and A. Smith, IEEE Proceedings F, Radar and Signal Processing **140**, 107 (1993).

[215] D. Frame, R. He, I. Ipsen, D. Lee, D. Lee, and E. Rrapaj, Phys. Rev. Lett. **121**, 032501 (2018).

[216] S. König, A. Ekström, K. Hebeler, D. Lee, and A. Schwenk, Eigenvector continuation as an efficient and accurate emulator for uncertainty quantification (2019), arXiv:1909.08446 .

[217] A. Ekström and G. Hagen, Phys. Rev. Lett. **123**, 252501 (2019).

[218] S. R. Stroberg, GitHub repository:PRL118032502 `https://github.com/ragnarstroberg/PRL118032502` .

[219] S. Yoshida, Y. Utsuno, N. Shimizu, and T. Otsuka, Phys. Rev. C **97**, 054321 (2018).

[220] Julia language `https://julialang.org`.

[221] J. Fairbrother, C. Nemeth, M. Rischard, J. Brea, and T. Pinder, (2018), arXiv:1812.09064 .