Data-driven approaches in nuclear shell-model calculations

Noritaka SHIMIZU[†1], Sota YOSHIDA[2], and Takahiro MIZUSAKI[3]

[1]Center for Nuclear Study, the University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[2]Institute for Promotion of Higher Academic Education, Utsunomiya University, Tochigi, Japan
[3] Institute of Natural Sciences, Senshu University, 3-8-1 Kanda-Jinbocho, Chiyoda-ku, Tokyo 101-8425, Japan
[†]Email: shimizu@cns.s.u-tokyo.ac.jp

Abstract

We briefly review some applications of machine learning and its related techniques to nuclear shell-model calculations. As an example, we quantified uncertainty caused by the parameters of the shell-model Hamiltonian utilizing Bayesian statistics. It enables us to quantify the uncertainty of the theoretical predictions and their agreement with experimental data in a statistical manner. Moreover, we point out that a large deviation of the confidence interval for the energy in shell-model calculations from the corresponding experimental data can be used as an indicator of some exotic property, *e.g.*, alpha clustering. Besides, we briefly introduce a recent effort to use the restricted Boltzmann machine to describe nuclear shell-model wave functions.

1 Introduction

As computational power has been increasing recently, machine learning and its peripheral techniques are intensively developed in the prevalence of artificial-intelligence (AI) techniques. The applications of these techniques are expected to accelerate the progress of natural science, including nuclear structure physics. Among the theoretical models in nuclear structure physics, nuclear shell model calculation is one of the most powerful tools to investigate the ground and low-lying excited states of nuclei, since it can describe any many-body correlations inside the valence shell on equal footing by configuration mixing. In this paper, we briefly review two applications of these techniques to the shell-model calculations as follows. The uncertainty quantification utilizing the Bayesian statistics in shell-model calculations [1] is discussed in Sect. 2. The introduction of the restricted Boltzmann machine (RBM) to the Variational Monte Carlo (VMC) formulation in shell-model calculations is discussed in Sect. 3. This paper is summarized and some other applications are mentioned in Sect. 4.

2 Uncertainty quantification by Bayesian analysis

2.1 Bayesian analysis

We here show the uncertainty quantification of the shell-model study of p-shell nuclei by applying the Bayesian analysis. In this study, we take the $0p_{3/2}$ and $0p_{1/2}$ single-particle orbits with ^{4}He

being an inert core. While the traditional shell-model Hamiltonian by S. Cohen and D. Kurath [2] is well-known for this model space, we construct the ensemble of the parameters of the Hamiltonian to estimate the uncertainty.

We consider the probability distribution of the interaction parameters which reproduce the experimental data with reasonable uncertainty. In the Bayesian analysis, it is described by the posterior distribution, *i.e.*, the conditional probability under the observation of data. This posterior probability is obtained exploiting Bayes' theorem,

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \propto P(d|\theta)P(\theta). \quad (1)$$

In this work, θ is a set of 17 parameters, which consist of the 2 single-particle energies and 15 two-body matrix elements and defines the shell-model Hamiltonian in the p-shell model space. The D is a set of the physical observables, which are taken as 33 data of energies and excitation energies in this work. In the present work, the likelihood function is taken as

$$P(D|\theta) = \exp[-\chi^2(\theta)/2] \quad (2)$$

with the squared errors,

$$\chi^2(\theta) = \sum_{n=1}^{N_D} \left(\frac{\mathcal{O}_n^{\rm expt} - \mathcal{O}_n^{\rm th}[\theta]}{\Delta\mathcal{O}} \right)^2 . \quad (3)$$

where the $\mathcal{O}_n^{\rm expt}$ and $\mathcal{O}_n^{\rm th}$ denote the n-th experimental data and theoretical data, respectively. The N_D is the number of data and, in this case, we take $N_D = 33$ binding and excitation energies of the p-shell nuclei. $\Delta\mathcal{O}$ denotes the typical error, 0.35 MeV, containing experimental and theoretical ones although the experimental error is negligible for these energies.

For simplicity, the prior probability is taken as the uniform distribution, namely $P(\theta) \propto 1$. In this case, the maximum a posteriori (MAP) estimation becomes with the minimization of the χ^2 fit. Note that the denominator in Eq.(1) is absorbed to the normalization factor and does not need to be considered.

We generate sets of the Hamiltonian parameters θ whose frequency distribution obeys the probability $P(D|\theta)$ to estimate the uncertainty of the theoretical results caused by the parameter fitting. For that purpose, we adopted the Laplace approximation [1], since we found it numerically difficult to obtain the parameter sets using the Markov Chain Monte Carlo method without this approximation.

2.2 Results

We prepared 50,000 samples of θ whose frequency distribution obeys the posterior probability, Eq.(1), and performed shell-model calculations to estimate the uncertainty of the shell-model results. Figures 1(a) and (b) show the excitation energies and energy expectation values and their uncertainties of the low-lying states of ^{12}C. In the figures, the violin plots with error bars show the theoretical results and their uncertainties. The shell-model results well reproduce the experimental values, which means that the experimental values are inside the 1σ uncertainty ranges except for the 0_2^+ state. It indicates that the 0_2^+ state cannot be described by p-shell shell-model calculation, which is reasonable since this state is known as three α-cluster state, or called the Hoyle state and its structure is considered to be far from shell-model wave functions. The result of the valence-shell in-medium similarity renormalization group (VS-IMSRG), in which the effective interaction is given in an *ab initio* way, is also shown for comparison.

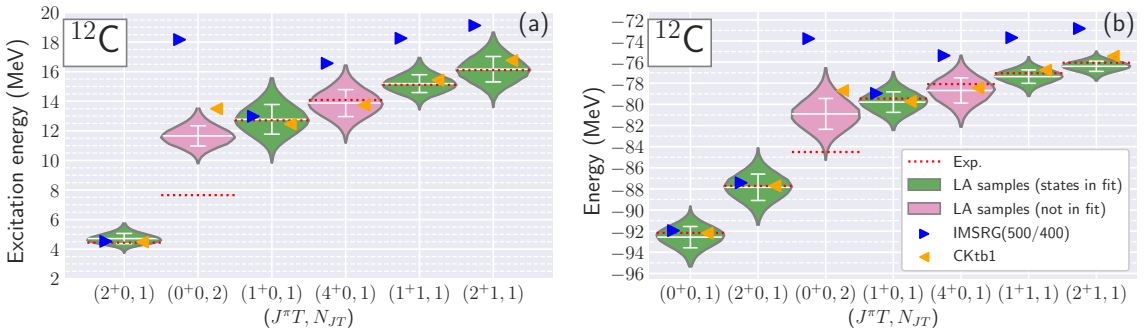

Figure 1: (a) Excitation energies and (b) energy expectation values of the ground and low-lying states of ^{12}C. The labels (J^π, T, N_{JT}) denote the total angular-momentum and parity J^π, isospin T, and N-th lowest state of (J^π, T). The violin plots with error bars show the theoretical values and their uncertainties. Red dotted lines, orange triangles, and blue triangles denote the experimental values, the traditional shell-model results [2], and the VS-IMSRG results [3]. Taken from Ref. [1].

3 Variational Monte Carlo in nuclear shell-model calculations and the restricted Boltzmann machine

In general, the shell-model calculation is difficult to be applied to heavy-mass nuclei since the dimension of the Hamiltonian matrix often becomes too huge to be diagonalized. As an attempt to solve this problem, we propose to introduce the Restricted Boltzmann Machine (RBM), one of the artificial neural networks, to the Variational Monte Carlo (VMC) framework. The introduction of the artificial neural network to solve quantum many-body problems was firstly succeeded in condensed matter physics [4, 5].

3.1 Formulation

We start with the framework of the VMC in shell-model calculations [6]. The VMC trial wave function with even A particles, $|\phi\rangle$, is defined as

$$\langle m|\phi\rangle \quad = \quad N(m)\langle m|\left(f_{ij}c_i^\dagger c_j^\dagger\right)^{A/2}|-\rangle \tag{4}$$

where i and j denote single-particle states, and f_{ij} is a set of variational parameters to describe pair correlation. $|m\rangle = c_{m_1}^\dagger c_{m_2}^\dagger \cdots c_{m_A}^\dagger|-\rangle$ is the M-scheme configuration [6, 7] specified by the A-particle occupation of the single-particle states $m = (m_1, m_2, \cdots m_A)$.

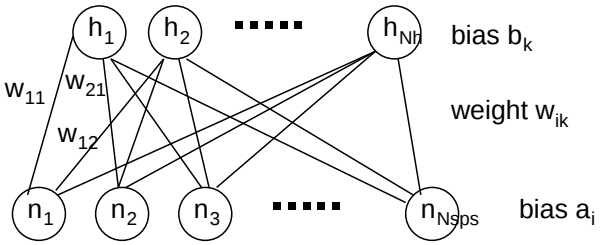

Figure 2: Schematic view of the restricted Boltzmann machine.

The $N(m)$ is given by the RBM and defined as [5]

$$\begin{aligned} N(m) &= \sum_{\{h_k=\pm 1\}} \exp\left(\sum_i a_i n_i + \sum_{i,k} w_{ik} n_i h_k + \sum_k b_k h_k\right) \\ &= \prod_k 2\cosh(b_k + \sum_i w_{ik} n_i) \exp(\sum_i a_i n_i) \end{aligned} \tag{5}$$

where a_i, b_k, and w_{ik} denote the bias of visible nodes, the bias of hidden nodes, and weights between the visible node i and the hidden node k, respectively. $n_i = 1$ for the occupied states ($n_{m_a} = 1$ for $a = 1, 2, ..., A$) and $n_i = 0$ for the unoccupied states. The schematic view of the RBM is shown in Fig. 2. Note that the weights of the RBM are restricted only between the visible and hidden nodes, from which the second line of Eq.(5) is deduced.

The energy expectation values of this trial wave function is evaluated statistically by

$$\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{1}{\sum_m |\langle m|\psi\rangle|^2} \sum_m |\langle m|\psi\rangle|^2 \frac{\langle m|H|\psi\rangle}{\langle m|\psi\rangle} = \sum_m p(m) E_l(m), \tag{6}$$

where the local energy is defined as $E_l(m) = \langle m|H|\psi\rangle/\langle m|\psi\rangle$. The summation $\sum_m p(m)$ in Eq. (6) is computed statistically by preparing a set of the M-scheme configurations $|m\rangle$ whose frequency distribution obeys the probability $p(m) \propto |\langle m|\psi\rangle|^2$ exploiting the Markov Chain Monte Carlo method [6]. Thus, we avoid to store the whole possible $|m\rangle$, the number of which may be too huge. The variational parameters, f_{ij}, a_i, b_k and w_{ik} are determined to minimize the energy expectation value by the stochastic reconfiguration method [8], which can be considered as one of the machine-learning techniques.

3.2 Benchmark test

As a benchmark test, we performed the VMC calculation to evaluate the ground-state energy of ^{28}Si with the *sd*-shell model space and the USD interaction [9]. In this case, 24 visible nodes are used to describe each occupation of the single-particle states in *sd* shell for the RBM.

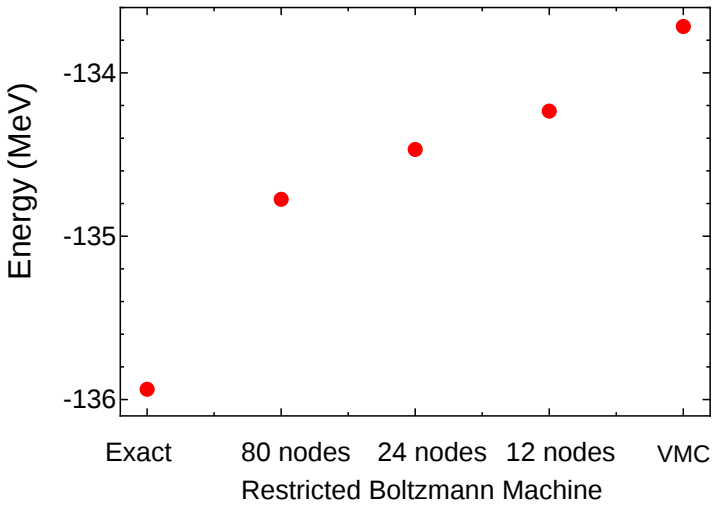

Figure 3: Shell-model energies of the ground state of ^{28}Si obtained by the exact diagonalization, the VMC results with the 80 RBM hidden nodes, 24 RBM hidden nodes, 12 RBM hidden nodes, and the VMC result without the RBM.

Figure 3 shows the shell-model energies obtained by the VMC without the RBM, with the RBM 12 hidden nodes, the 24 hidden nodes, the 80 hidden nodes. The exact energy is also shown in the left-hand side of Fig. 3. The statistical errors of the VMC are small enough and not shown in the figure. While the VMC without the RBM shows 2.2-MeV deviation from the exact one, the introduction of the RBM factor fills this gap to some extent. As increasing the number of the RBM hidden nodes, the description power of the RBM is enhanced and the energy approaches the exact value, although the 1-MeV gap remains even with the 80 hidden nodes. Further study is expected to fill this gap by applying the angular-momentum projection [6].

4 Summary and perspectives

We briefly reviewed some applications of the machine-learning techniques to nuclear shell-model calculations. Bayesian analysis was applied to the shell-model calculations and we demonstrated that uncertainty quantification is feasible with p-shell nuclei as an example. The statistical analysis of the sd-shell nuclei is found in Ref. [10]. As another application, we introduced the RBM to the VMC approach in the framework of the shell-model calculations and performed the benchmark test, which shows promising features to overcome the numerical difficulty in shell-model calculations.

Many more promising attempts have been done to apply the AI-related methods to shell-model calculations, which cannot be described in this paper. For example, we exploit a clustering algorithm to divide the basis states of the no-core Monte Carlo shell model [11] into a small number of groups in order to discuss the α-clustering structure of ^{12}C. Besides, Gaussian Process and artificial neural network are exploited to the extrapolation of the energy eigenvalue to infinite model space in the no-core shell model approach [12, 13]. We expect that such AI-related approaches will drastically open the frontiers of nuclear structure physics.

Acknowledgment

Numerical calculations were performed on Oakforest-PACS supercomputer (hp200130, hp190160, hp180179, hp170230, hp160146, xg18i035) and CX400 supercomputer of Nagoya University (hp160146). NS acknowledge valuable supports by "Priority Issue on post-K computer" (Elucidation of the Fundamental Laws and Evolution of the Universe) and "Program for Promoting Researches on the Supercomputer Fugaku" (Simulation for basic science: from fundamental laws of particles to creation of nuclei), MEXT, Japan.

NS and SY acknowledge T. Togashi and T. Otsuka for our collaboration about the uncertainty quantification of shell-model studies.

References

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

[2] S. Cohen and D. Kurath, Nucl. Phys. **73**, 1 (1965).

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

[4] G. Carleo and M. Troyer, Science **355**, 602 (2017)

[5] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017); L. Huang, and L. Wang, Phys. Rev. B **95**, 035105 (2017); G. Carleo, Y. Nomura, and M. Imada, Nat. Comm. **9**, 5322 (2018).

[6] T. Mizusaki and N. Shimizu, Phys. Rev. C **85**, 021301(R) (2012); N. Shimizu and T. Mizusaki, Phys. Rev. C **98**, 054309 (2018).

[7] N. Shimizu, T. Mizusaki, Y. Utsuno, and Y. Tsunoda, Comp. Phys. Comm. **244**, 372 (2019).

[8] S. Sorella, Phys. Rev. B **64**, 024512 (2001).

[9] B. A. Brown and B. H. Wildenthal, Ann. Rev. Nucl. Part. Sci. **38**, 29 (1988).

[10] J. M. R. Fox, C. Johnson, and R. N. Perez, Phys. Rev. C **101**, 054308 (2020).

[11] T. Yoshida, N. Shimizu, T. Abe and T. Otsuka, J. Phys.: Conf. Ser. **569**, 012063 (2014).

[12] S. Yoshida, Phys. Rev. C **102**, 024305 (2020)

[13] W. G. Jiang, G. Hagen, and T. Papenbrock, Phys. Rev. C **100**, 054326 (2019) G. A. Negoita et al., Phys. Rev. C **99**, 054308 (2019).