Study of yrast and yrare low-lying excited states using machine learning approaches*

Zhi Long Li (李志龙)[1,2] Bing Feng Lv (吕冰锋)[3†] Yong Jia Wang (王永佳)[1‡] C. M. Petrache[4]

[1]School of Science, Huzhou University, Huzhou 313000, China
[2]China Institute of Atomic Energy, Beijing 102413, China
[3]Institute of Modern Physics, Chinese Academy of Science, Lanzhou 730000, China
[4]Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Abstract: The low-lying excitation energies of the $2_1^+,4_1^+,2_2^+,0_2^+,3_1^-,0_3^+$ states in even-even nuclei are studied using two modern machine learning algorithms: the Light Gradient Boosting Machine (LightGBM) and Sparse Variational Gaussian Process (SVGP). The obtained results demonstrate that both LightGBM and SVGP perform well on the training and validation datasets when informed by a physics-based feature space. A detailed comparison of the results obtained for 2_1^+ and 2_2^+ states using the Hartree-Fock-Bogoliubov theory extended by the generator coordinate method and mapped onto a five-dimensional collective quadrupole Hamiltonian shows that both ML algorithms outperform this model in terms of accuracy. The extrapolation capabilities of these algorithms were further validated using newly measured 12 data points of 2_1^+ and 2_2^+ states, which were not included in the training set. In addition, the partial dependence plot method and the Shapley additive explanations method are used as interpretability tools to analyze the relationship between the input features and model predictions. These tools provide in-depth insights into how the input features influence the prediction of low-lying excitation energies and help identify the most important features that drive the prediction, which are valuable for understanding the low-lying excitation energies.

Keywords: machine learning, LightGBM, SVGP, $E(2_1^+)$, $E(2_2^+)$

DOI: 10.1088/1674-1137/adfe54 **CSTR:** 32044.14.ChinesePhysicsC.50014107

I. INTRODUCTION

The investigation of the properties of nuclear excited states is crucial for elucidating the complex interactions of protons and neutrons in atomic nuclei. However, although substantial progress has already been made in characterizing low-lying yrast and yrare states of nuclei near the stability line, experimental measurements remain particularly scarce for nuclides that are close to or beyond the drip lines [1−13]. The measurement of excited states unveils the wide diversity of nuclear phenomena, *e.g.*, shell evolution, pairing correlation, shape coexistence, octupole deformation, clustering, continuum effects [14−18]. Currently, a vast number of exotic nuclei have yet to be measured because of the challenges posed by various experimental techniques and the considerable investment of beam time required, which are primarily a result of their low production cross section. Among the excited states, the 2_1^+ state garners significant attention because of its fundamental importance, carrying crucial benchmark information for investigating the evolution of shell structures and collectivity. As the proton and neutron numbers approach the magic numbers, the energy of the 2_1^+ state, $E(2_1^+)$, increases sharply, reaching a maximum in double magic nuclei. As one moves away from the magic numbers along isotope chains, the presence of low 2_1^+ excitation energies reflects a deformed ground state stemming from the polarizing effect of added nucleons, which induces deformation. The non-yrast 2_2^+ states are regarded as indicators of collectivity, showing simpler characteristics compared to those of their 2_1^+ counterparts. The physical properties of 2_2^+ excited states are always of interest because of their potential role as initial levels of collective bands, which are traditionally referred to as γ-vibrational bands or quasi-γ bands.

Various theoretical methods have been developed to reproduce and predict the energies of low-lying excited states and to further explore their underlying physical

Received 20 May 2025; Accepted 19 August 2025; Published online 20 August 2025

* This work was supported by the National Natural Science Foundation of China (12305128), by the Hubert Curien Partnership (PHC) Cai Yuanpei Project, by the International Partnership Program of Chinese Academy of Sciences for Future Network (016GJHZ2023024FN), by the Gansu Natural Science Foundation (24JRRA038), by the Major Science and Technology Projects in Gansu Province (24GD13GA005), by National Key R&D Program of China (2023YFA1606402)

† E-mail: lvbingfeng@impcas.ac.cn
‡ E-mail: wangyongjia@zjhu.edu.cn

mechanisms over the Segrè chart, *e.g.*, shell model [19, 20], nuclear energy density functional theory [21−24]. However, these models often demand a large amount of computational resources, and notable discrepancies between the model calculations and experimental data have been observed for nuclei located far from the stability line [25−27]. These may be attributed to the complexity of the mapping function, which is non-trivial and, in some cases, cannot be properly defined. In recent years, machine learning (ML) algorithms have been used to simulate known data and predict unmeasured ones [28−31]. ML achieves significantly shorter computation times than traditional methods while maintaining comparable accuracy. Promising results have been obtained in using ML to address critical issues and predict unknown properties of nuclei [28−30], *e.g.*, β-decay half-lives and energy [32−36], α-decay [37−41], mass and charge radius of atomic nuclei [42−45], nuclear density distribution [46−49], and heavy ion collisions [50−53]. Recently, the energies of the 2_1^+ states were studied using the Bayesian neural network (BNN) approach, artificial neural networks method, and machine learning approach for even-even nuclei across the nuclide chart, showing superior performance compared to these of the shell model and the five-dimensional collective Hamiltonian model [54−57].

In this study, two types of algorithms are employed to investigate the excitation energies of 2_1^+ and 2_2^+ states: the Light Gradient Boosting Machine, which is based on decision trees, and Sparse Variational Gaussian Process, which is a Bayesian method. Both algorithms are supervised tasks that require ML algorithms along with a set of labeled data consisting of input and output variables for regression predictions. Our goal is to assess the performance of the two ML algorithms in predicting $E(2_1^+)$ and $E(2_2^+)$, and to examine their interpretability. In addition, the present study aims to extrapolate these observables to unseen data, rather than merely optimizing the fit to known data by minimizing the root-mean-square (rms) error. A comparison with the results of calculations based on the Hartree-Fock-Bogoliubov theory extended by the generator coordinate method and mapped onto a five-dimensional collective quadrupole Hamiltonian (HFB+5DCH) is also presented [25]. It is noteworthy that although $E(2_1^+)$ has already been studied using other machine learning approaches, it is intriguing to explore whether new algorithms can refine the predictions, thereby facilitating more reliable explorations of unmeasured nuclei. In addition, the present work is the first attempt to explore $E(2_2^+)$ within the framework of machine learning. Finally, the performance of these ML models in predicting $E(2_1^+)$ and $E(2_2^+)$ was examined using 12 newly measured data points that were not included in the training set.

The article is organized as follows: Section II describes the LightGBM and SVGP algorithms, training and testing datasets, and their input features. Section III presents the results obtained from the two algorithms and provides relevant discussions. Finally, Sec. IV summarizes our findings and conclusions from the present work.

II. METHODOLOGY

In this section, we introduce the two algorithms employed in our study: Light Gradient Boosting Machine and Sparse Variational Gaussian Process. We also outline the parameter sets and training and validation datasets used for these models.

A. Machine learning algorithms

LightGBM is an efficient gradient boosting framework developed by Microsoft that has been widely used for various machine learning tasks [58]. It employs a histogram-based learning algorithm that significantly accelerates training and reduces memory usage compared to those obtained with traditional boosting methods. LightGBM handles sparse data and categorical features effectively, supports parallel and distributed computing, and performs well in tasks such as classification, regression, and ranking. A key advantage of LightGBM lies in its optimized leaf-wise tree growth strategy, which improves accuracy by reducing loss more aggressively than the level-wise methods. Combined with its strong scalability and low computational cost, LightGBM has become a popular choice in both machine learning competitions and real-world applications.

SVGP is a scalable probabilistic modeling approach that integrates variational inference with sparse approximations [59, 60]. By introducing a small set of inducing points, SVGP reduces the computational complexity of standard Gaussian Processes (GP) from $O(N^3)$ to $O(M^3)$, where $M \ll N$, while preserving core GP properties such as uncertainty quantification. Its main strengths include: (1) Efficiency, through sparse matrix operations and evidence lower bound (ELBO) optimization; (2) Flexibility, supporting non-Gaussian likelihoods, high-dimensional features, and customizable kernels; and (3) Interpretability, as the variational distribution $q(u)$ directly conveys the model's confidence in predictions. Compared to traditional GPs and deep neural networks, SVGP offers a balanced trade-off between expressiveness, uncertainty estimation, and computational feasibility, serving as a bridge between Bayesian inference and large-scale machine learning.

To ensure fair and effective model evaluation, key hyperparameters for both models were systematically configured, with an emphasis on balancing predictive accuracy and computational efficiency. For LightGBM, the main hyperparameters are set as follows: the maximum number of leaves per tree is 10, the maximum depth is set to –1 (indicating no depth limit), and the number of

boosting iterations is fixed at 50,000. All other parameters remain at their default values to maintain standard model behavior. For SVGP, the number of inducing points (n_z) is set to 100 to enable sparse variational approximation. During training, 20 particles (n_particles = 20) are used for Monte Carlo estimation of the ELBO, and 100 particles (n_particles_test = 100) are used during testing to improve stability. The model is trained using a batch size of 300 (batch_size = 300) over 8000 epochs (n_epoch = 8000), with a learning rate of 0.01 (lr = 0.01).

B. Data and input features

The dataset used in this study comprises a total of 660 and 437 experimental data points corresponding to the excitation energies of the 2_1^+ and 2_2^+ states, respectively, in even-even nuclei with proton numbers ranging from $Z = 10$ to $Z = 100$. The data are obtained from the National Nuclear Data Center [61]. To train the machine learning models, the dataset is randomly divided into training and test sets in various ratios. In our previous study, we initially investigated the effect of the ratio between the training and validation sets on the prediction performance of the LightGBM model [56]. The results showed that the rms deviation decreases as the proportion of training data increases. However, an excessively large training set reduces the number of data points in the test set, thereby increasing the uncertainty of the performance evaluation. Therefore, following the set in the earlier studies, we adopt a split ratio of 80% for training and 20% for validation in this study. This 4:1 division, along with other commonly used ratios such as 7:3 or 3:1, is a standard strategy widely applied in machine learning model design. In total, we randomly divide the data points into training and validation datasets 500 times. To study $E(2_1^+)$ and $E(2_2^+)$, we incorporate 16 relevant features of nuclei into our algorithm, referred to as $M16$. $M16$ consists of 16 inputs: one-proton separation energy (S_p), one-neutron separation energy (S_n), two-proton separation energy (S_{2p}), two-neutron separation energy (S_{2n}), and the separation energy of ^{4}He (S_{He}); deformation parameters derived from FRDM calculations [62] (β_1) and from WS4 [63] (β_2); proton number (Z), neutron number (N), mass number (A), experimental binding energy (B), binding energy predicted from the liquid drop model ($B_{\rm LDM}$), the difference between experimental B and the liquid drop model ($B - B_{\rm LDM}$), Casten factor (P) [64], the valence number of neutrons as measured from the nearest closed shell (v_n), and the valence number of protons as measured from the nearest closed shell (v_p). These features are recognized as fundamental characteristics of nuclei and impact the performance of ML models. The performance of the ML algorithms on both the training and test datasets were quantified using the standard rms deviation between the machine learning predictions and experimental data, defined as follows:

$$rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\log_{10}E^{\rm Exp} - \log_{10}E^{\rm ML})^2},$$

where $\log_{10}E^{\rm ML}$ represents the ML prediction, $\log_{10}E^{\rm Exp}$ denotes the corresponding experimental value, and N is the total number of nuclei.

III. RESULTS AND DISCUSSION

Performance of LightGBM and SVGP in the study of $E(2_1^+)$. In this work, to ensure the quality the models, we initially examined the learning curves of LightGBM and SVGP for the excitation energies of 2_1^+ and 2_2^+ states under the 4:1 train-validation split, as shown in Fig. 1. The learning curves show how the loss function evolves with the number of training iterations (or trees). As training proceeds, the loss gradually decreases and converges, indicating that the model is approaching an optimal state. The difference between the training and validation losses reflects the model's fitting behavior: the two curves should be close under proper training, while a large difference between them suggests overfitting or underfitting. One can infer from Fig. 1 that the loss values for validation and training decrease with the increase in the number of trees (or epoches) and saturate at approximately 20000 for LightGBM and approximately 500 for SVGP in the study of $E(2_1^+)$. For $E(2_2^+)$, the saturation occurs at approximately 25000 for LightGBM and approximately 500 for SVGP. In addition, one can observe in Fig. 1 that the loss functions for both models decrease steadily on both training and validation curves lying on top of each other, indicating no fitting issues. These results demonstrate that the selected hyperparameter configurations effectively ensure training stability, good generalization, and computational efficiency for both models.

The excitation energies of the 2_1^+ states were previously explored in our work using the LightGBM algorithm [56]. The rms deviation of the LightGBM approach with respect to the experimental $\log_{10}E$ was determined to be 0.030(1) for $E(2_1^+)$. In addition, $E(2_1^+)$ was analyzed using LightGBM in Ref. [57], albeit with only five physical features. Nevertheless, it still revealed that the average difference between the LightGBM predictions and the experimental data was 18 times smaller than that obtained with the shell model and only 70% that of the BNN prediction results [57]. In this study, we incorporated deformation parameters derived from FRDM calculations [62] (β_1) as an additional feature to account for the collectivity of nuclei. Finally, rms values of 0.032(3) and 0.049(5) were obtained on the training and validation sets, respectively. By utilizing the same 16 input ($M16$) features within the framework of SVGP, we also

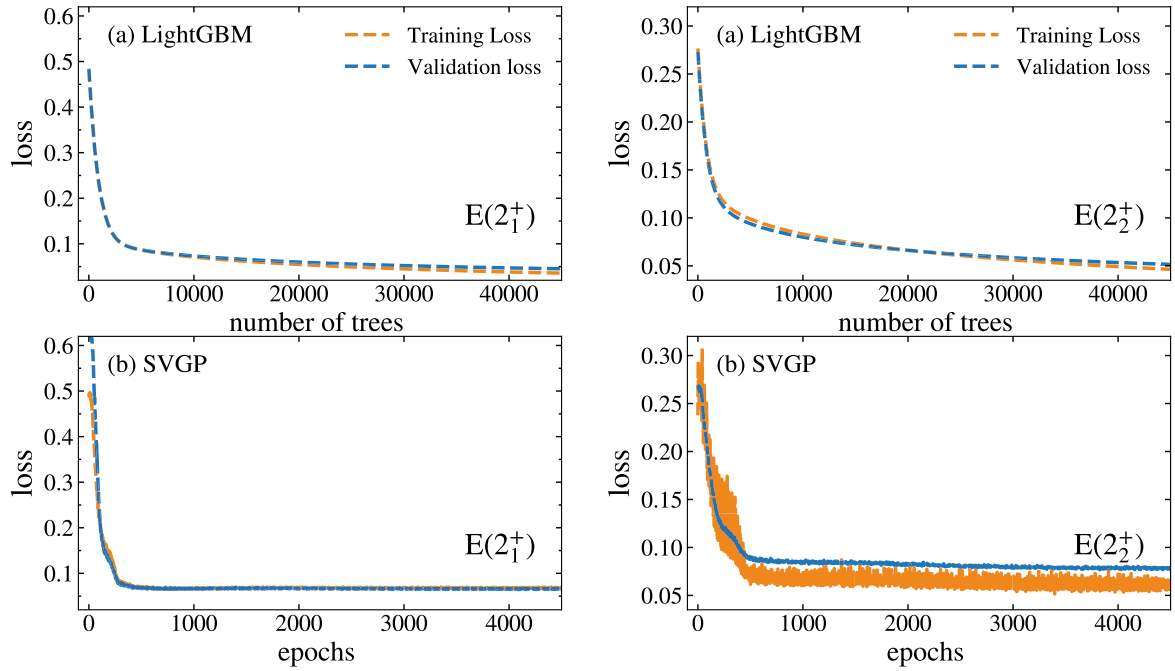

Fig. 1. (color online) Learning curves of the LightGBM and SVGP models for the prediction of the excitation energies of 2_1^+ and 2_2^+ states illustrate the evolution of the loss function with the number of training iterations (*i.e.*, number of decision trees or training epochs).

investigated the $E(2_1^+)$ of even-even nuclei. This implementation yielded rms values of 0.066(2) and 0.070(6) on the training and validation sets, respectively, which are a factor of two inferior to those obtained using LightGBM, suggesting that the LightGBM more effectively captures the excitation energies of 2_1^+ states.

Performance of SVGP and LightGBM in the study of $E(2_2^+)$. Prior to this work, the excitation energies of 2_2^+ states remained unexplored by machine learning methodologies. In the current study, we employed LightGBM and SVGP to analyze $E(2_2^+)$ using the $M16$ feature space. Figure 2 illustrates the density distribution of the rms values for the validation dataset comprising the excitation energies of the 2_2^+ states using LightGBM and SVGP. The rms values for the validation data points predicted by LightGBM and SVGP were 0.058(5), and 0.077(6), respectively. This indicates that LightGBM and SVGP reproduce the experimental data within a factor of $10^{0.058} = 1.14$ and $10^{0.077} = 1.19$, respectively. These results are better than those obtained with most traditional theoretical models. Similar to that observed during the study of $E(2_1^+)$, the decision tree-based LightGBM outperforms SVGP in predicting $E(2_2^+)$. For the reader's convenience, we have summarized the current results in Table 1. It is worth noting that although the predictions from LightGBM are deterministic under fixed training data and parameters, the reported validation rms error (*e.g.*, 0.058 ± 0.005) results from repeated random partitioning of the training data. Specifically, we performed 500 independent training runs, each using a randomly selected

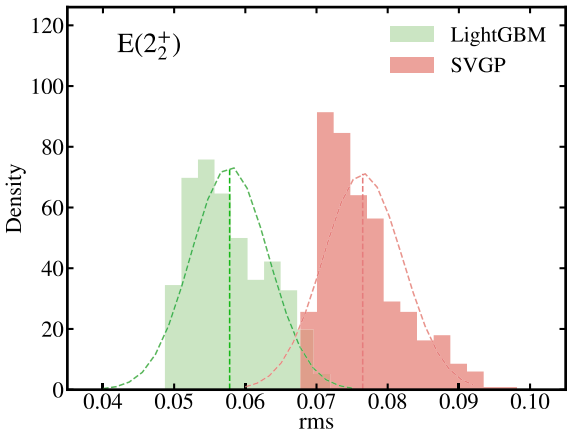

Fig. 2. (color online) Density distribution of the rms values of the validation dataset from LightGBM and SVGP for $E(2_2^+)$ using M16 feature space. The results for 500 runs are displayed. Dashed lines denote a Gaussian fit to the distribution. In each run, the data points were randomly split into training and validation sets at a ratio of 4:1.

80% of the data for training and the remaining 20% for validation. The reported mean value of rms and standard deviation are calculated across these runs to assess the model's stability and generalization under different data splits.

Figure 3 displays the excitation energies of the 2_2^+ states as derived from LightGBM, SVGP, and HFB+5DCH calculations, as a function of their corresponding experimental values. This comparison facilitates

Table 1. Average rms values on the training and validation sets.

	Algorithms	Training	Validation
$E(2_1^+)$	LightGBM	0.032(4)	0.049(5)
$E(2_1^+)$	SVGP	0.066(2)	0.070(6)
$E(2_2^+)$	LightGBM	0.040(3)	0.058(5)
$E(2_2^+)$	SVGP	0.067(7)	0.077(6)

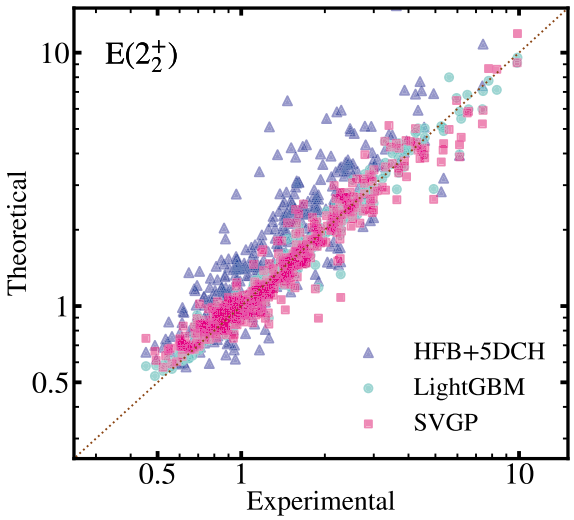

Fig. 3. (color online) Theoretical results for $E(2_2^+)$, obtained using HFB+5DCH, LightGBM, and SVGP, compared with the corresponding experimental data.

a comprehensive evaluation of the predictive power of each methodology across an extensive energy scale, spanning over an order of magnitude from approximately 0.3 MeV to 10 MeV. The analysis reveals that, on average, the HFB+ 5DCH model tends to overestimate the excitation energies $E(2_2^+)$, predicting values that are approximately 25% higher than the experimentally observed ones [25]. By contrast, both SVGP and LightGBM exhibit reduced variance in their predictions, indicating greater consistency. Notably, the LightGBM predictions demonstrate a high degree of concordance with the experimental data, underscoring its superior accuracy compared to those of the HFB+5DCH and SVGP approaches.

To more clearly illustrate the discrepancies between the calculations and experimental results, the differences between the machine learning predictions and experimental data for $E(2_2^+)$ on the nuclear chart are quantified in Fig. 4. It is evident that for both machine learning algorithms, the predictions exhibit excellent agreement with the experimental data for nearly all nuclei, including triaxial, transitional, and magic nuclei. However, deviations are observed mainly in light nuclei with complex structures, presenting halos and clusters, and in those nuclei that exhibit dominant single-particle characteristics. Additionally, significant deviations are noticeable in

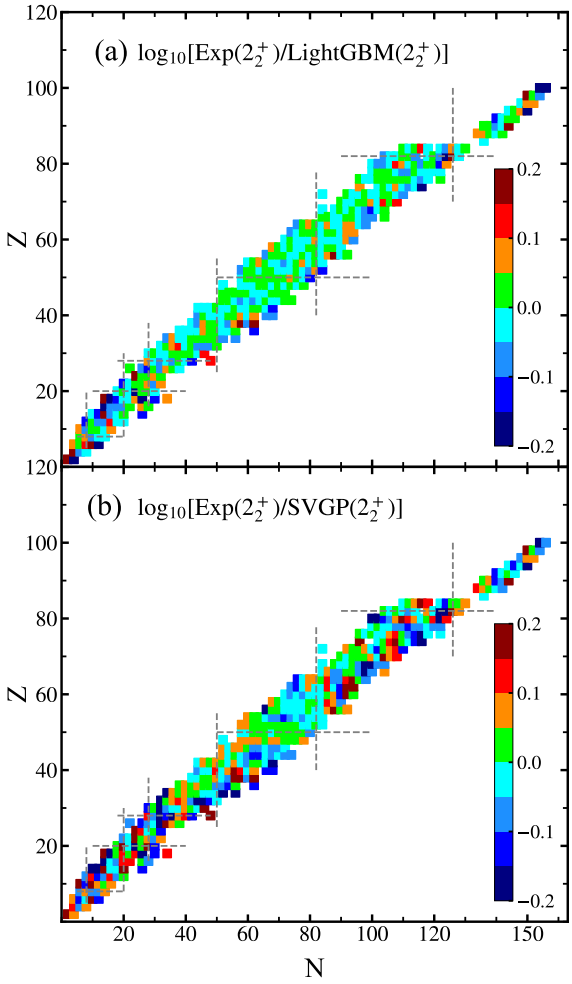

Fig. 4. (color online) Panels (a) and (b) show the differences between the experimental data and machine learning predictions using LightGBM and SVGP for the excitation energies of 2_2^+ states, respectively.

the neutron-rich nuclei located in the medium-heavy mass regions, likely as a result of the relative scarcity of training data in these areas.

Physically interpretablity. The implementation of machine learning algorithms is often constrained by their "black box" nature, which obscures the reasoning behind their predictions. Recently, explainable AI techniques have gained significant attention as critical tools for addressing this interpretability gap. Among the various methods, SHapley Additive exPlanations (SHAP) has become widely recognized for its effectiveness in providing clear, comprehensible insights into model behavior, while also quantifying feature importance through cooperative game theory frameworks [65−68]. In this work, we apply the SHAP technique to gain a deeper understanding of the results obtained using LightGBM. Figure 5 presents a ranking of the features in the $M16$ feature space used for predicting $E(2_2^+)$ using LightGBM. In these plots, the features are ordered along the y-axis ac-

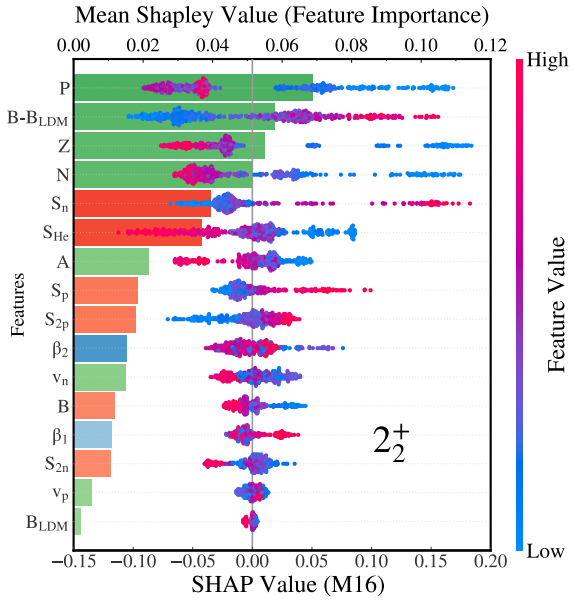

Fig. 5. (color online) Ranking of feature importance measured by SHAP value for the excitation energies of 2_2^+ states. The color of each bar indicates the direction of the feature's influence on the prediction: blue indicates a decrease, with lower feature values, whereas red indicates an increase, with higher feature values, with the intensity of the color denoting the magnitude of the feature's value.

cording to their impact, with the most impactful feature at the top and the least influential at the bottom. The x-axis represents the relative importance of each feature, as determined by its SHAP value. Features with larger absolute SHAP values have a greater impact on the model's predictions, while those with smaller values have a smaller effect. The analysis reveals that the feature P is the most critical because it reflects nuclear collectivity, to which $E(2_2^+)$ has a high sensitivity. A similar behavior of P was also observed during the study of $E(2_1^+)$ [56, 57]. The difference in binding energy between the experimental data and the predictions of the liquid drop model (denoted as $B-B_{\rm LDM}$) occupies a secondary position in the ranking, reflecting shell and pairing effects, as well as nuclear deformation [56]. Following this, the top six features also include Z, N, S_n, and S_{He}. Beyond these, all other features have relatively smaller, albeit indispensable, contributions to the prediction of E_2^+. This suggests that using all 16 basic features appears to be a reasonable tactic for making accurate predictions.

To identify feature redundancy, we constructed a simplified feature set consisting of the top six SHAP-ranked features (denoted as M6) and used it as input to train both the LightGBM and SVGP models. On the $E(2_2^+)$ validation set, the resulting rms errors were 0.070 and 0.090, respectively—higher than the 0.058 (LightGBM) and 0.077 (SVGP) achieved with the complete M16 feature set. Compared to those of the M16 model, the rms values of the M6 models increased by approximately 20% and 16%, respectively, suggesting that non-linear interactions may exist among input features, and that removing certain features could weaken the model's ability to leverage other information, ultimately impairing predictive performance. Therefore, although WS4's β_2 ranked lower than FRDM's β_1 in the SHAP analysis, we retained FRDM's β_1 in the initial feature set to preserve any potential synergistic effects. Additionally, both LightGBM and SVGP have strong feature selection capabilities and employ regularization (LightGBM) and variational inference (SVGP) during training, making them robust to redundant features. This robustness underscores the reliability of our methodology and modeling choices.

Partial Dependence Plot (PDP) analysis is another effective method for investigating the functional relationship between input features and model predictions in the LightGBM algorithm [69, 70]. Within this framework, if the curve for a particular feature is nearly constant or shows random fluctuations, it suggests that the feature may be insignificant or uninformative. Conversely, a steep PDP curve or one that exhibits significant changes indicates that the feature has a substantial contribution to the model's predictions. Figure 6 quantifies the impact of the six most salient input features on the prediction of $E(2_2^+)$. Although PDP and SHAP utilize different techniques to characterize importance, the identified six most important features are the same. Among the input features, proton number (Z) has the most pronounced influence on the prediction of $E(2_2^+)$. It has a negative impact; an increase in its value from 20 to 80 resulted in a decrease in the PDP value from 3.34 to 3.10, corresponding to a reduction of approximately 7.1%. The Casten factor P also exhibits a negative impact; an increase in its value from 0 to 4 resulted in a decrease in the PDP value from 3.29 to 3.08, corresponding to a reduction of approximately 6.2%. Similar behaviors were observed for the features N and S_{He}. However, as shown in Fig. 6, the predicted $E(2_2^+)$ increases with increasing values of $B-B_{\rm LDM}$ and S_n. The effect becomes particularly significant when the difference in the binding energy ($B-B_{\rm LDM}$) increases from –3 to 3 MeV. As shown in Fig. 7, the PDP curve for $B_{\rm LDM}$ is nearly flat, with a variation of only about 0.16%, indicating a negligible average impact on the model's output. This finding is consistent with its low importance as revealed by the SHAP analysis.

Performance of SVGP and LightGBM on the newly reported data. In addition to best reproducing the known $E(2_1^+)$ and $E(2_2^+)$ data across the nuclear landscape, one of the important objectives of this work is to apply the algorithms in the extrapolated region where experimental measurements are absent or currently inaccessible. In the present work, we initially evaluate the extrapolation abilities of the LightGBM and SVGP algorithms by venturing beyond the training area using new data points. Fig-

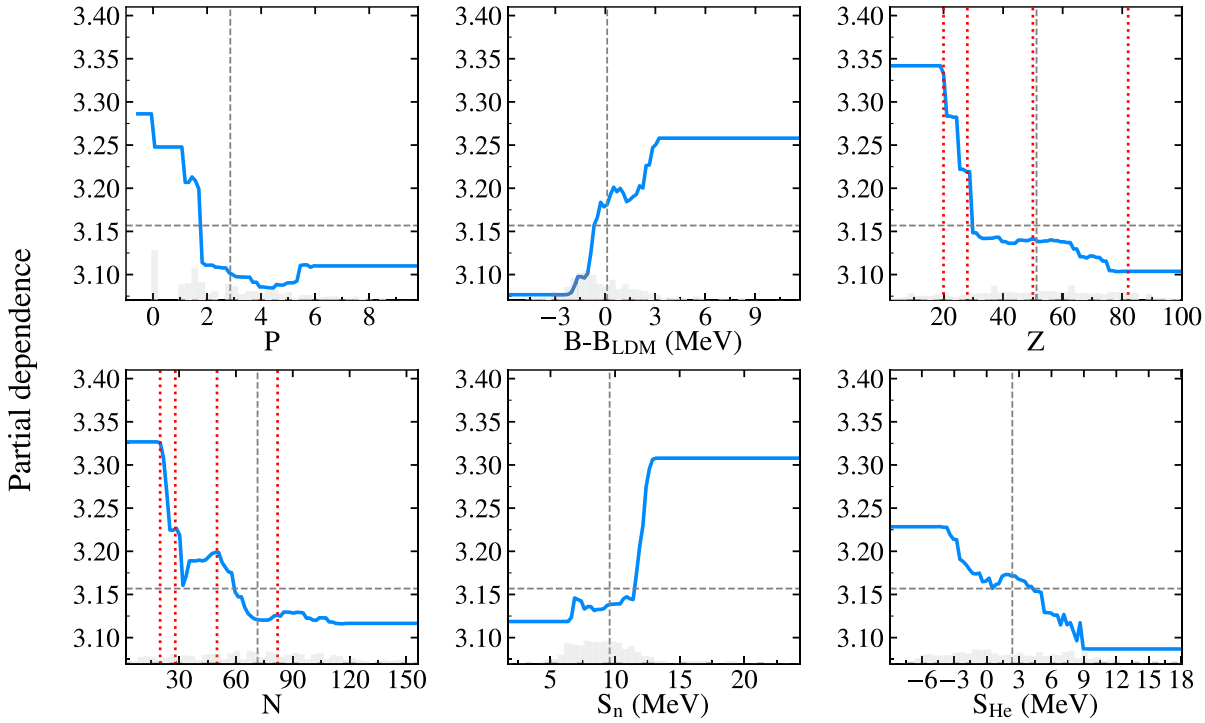

Fig. 6. (color online) PDP analysis of the effect of input features on the excitation energies of 2_2^+ states. The blue lines represent the partial dependence value, while the gray columns represent the data point distribution for each input feature at a certain value. The red dashed lines indicate the positions of neutron or proton magic numbers.

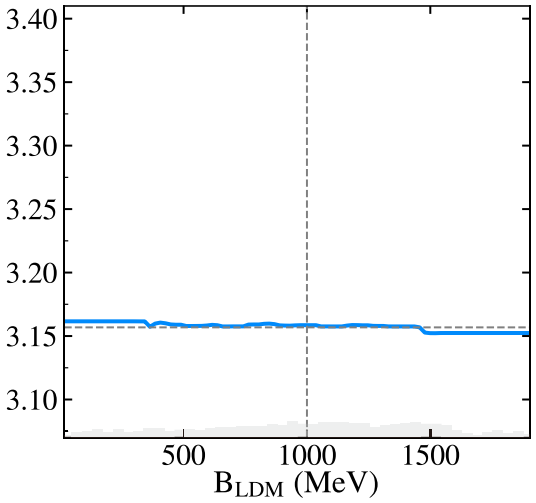

Fig. 7. (color online) Similar to Fig. 6, but for $B_{\rm LDM}$.

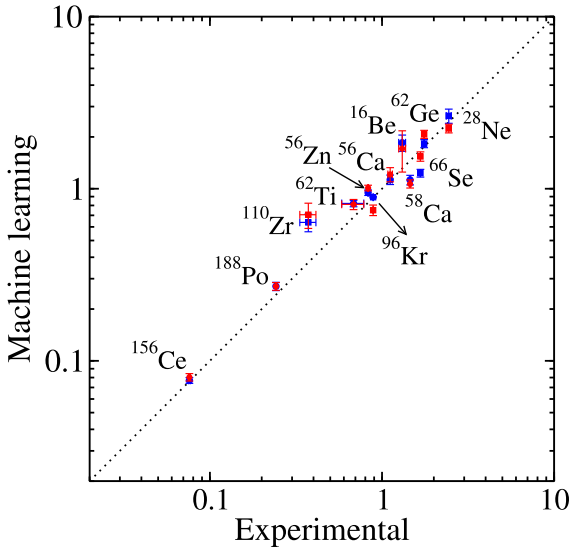

Fig. 8. (color online) Comparing the new experimental data for $E(2_1^+)$ and $E(2_2^+)$ in MeV with the prediction results of LightGBM and SVGP; the blue symbols denote predictions from SVGP, while the red symbols denote predictions from LightGBM. The $E(2_1^+)$ and $E(2_2^+)$ are shown as circles and squares, respectively.

ure 8 compares a total of 12 newly reported data from 2020 to the present with predictions from LightGBM and SVGP [71−81]. These newly measured data points were not included in the training process. It is evident that LightGBM and SVGP demonstrate reasonable predictive capabilities outside the training region for these nuclei, showing good agreement within the uncertainty. This favorable outcome could be attributed to the newly measured nuclei, which differ by up to 2 nucleons from their nearest neighbors in the (Z, N) space of the training set, with a maximum difference of 4 for the 2_1^+ states. Additionally, although present results indicate that the employed machine learning models maintain strong general-

ization capability and stable predictive performance for the nuclei that are not so close to the drip lines and to where neighbors exist, it is challenging to draw conclusions regarding the robustness of drip-line nuclei, where no close neighbors are present. This difficulty arises as a result of the weakly bound or unbound nature of these nuclei, which may lead to abnormal behavior. Such strong performance in these extrapolation tests instills confidence in the model's predictions for nearby nuclei that have not yet been experimentally investigated. However, as also indicated in Fig. 8, there are slight discrepancies between the ML predictions and the experimental data for ^{110}Zr. These differences may arise from the tentative assignment of spin and parity from the experimental side, or the enhanced triaxial deformation that the ML may not have adequately captured [75]. An attempt to include the triaxial deformation (γ) information from the sophisticated nuclear models in the feature space may enhance the prediction capability. It should be mentioned that all target values used for training in this study were sourced from the National Nuclear Data Center [61]. However, given that experimental data on $E(2_2^+)$ and $E(2_1^+)$ are intrinsically limited (covering only a few hundred nuclei), we did not apply special filtering for cases with uncertain assignments during dataset construction. We will now explore how to effectively incorporate experimental uncertainty or systematic bias into machine learning models, for example by introducing confidence-based weighting or using Bayesian frameworks, to enhance model robustness against label noise.

Data-driven approaches should not only aim for predictive accuracy but also strive to uncover underlying physical patterns. In Fig. 6, the PDP curves for Z and N show pronounced slope changes near known magic numbers, indicating the model's sensitivity to shell effects. Figure 9(a) and (c) present the variation of $E(2_2^+)$ with Z and N, with red markers denoting cases where Z or N corresponds to a magic number. These results show that closed-shell nuclei generally have higher excitation energies. Figure 9(b) and (d) show how SHAP values vary with Z and N, reproducing the shell effects in many regions, although discrepancies remain in deformation-dominated areas —likely due to the complex nature of nuclear shapes. This indicates that the model's predictive accuracy in such regions is still lower than that in spherical nuclei. Nonetheless, these results demonstrate that

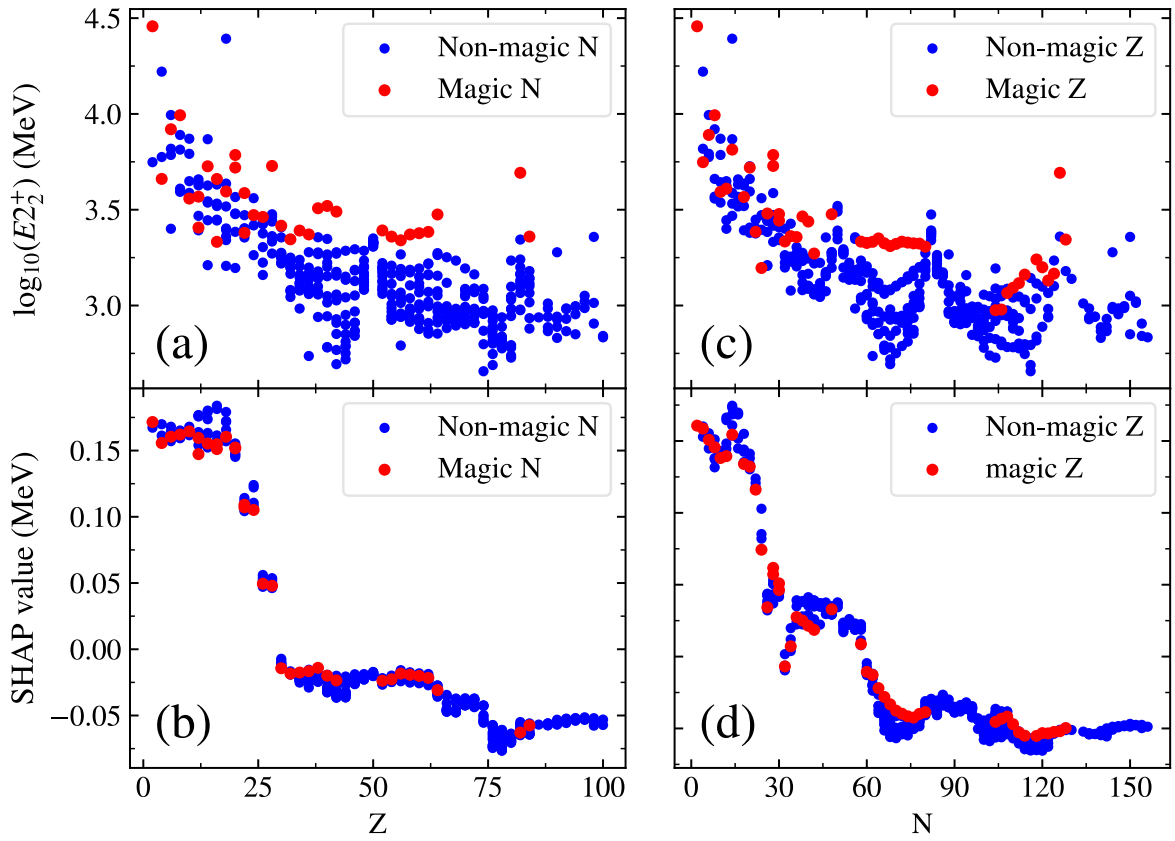

Fig. 9. (color online) Panels (a) and (b) show the variation of $E(2_2^+)$ energies and SHAP values with proton number Z, where red dots correspond to nuclei with neutron magic numbers, and blue dots to non-magic N. Panels (c) and (d) depict the variation of $E(2_2^+)$energies and SHAP values with neutron number N, where red dots correspond to nuclei with proton magic numbers and blue dots to non-magic Z.

the machine learning model captures not only statistical trends but also key physical features of the nuclear structure.

As demonstrated herein, the performance of machine learning models tends to deteriorate in regions for which experimental data are scarce because these algorithms are unable to learn any points in such areas. At present, we extrapolate to unmeasured nuclei for each isotope chain by considering two additional even-even nuclei. Given that ML algorithms generally exhibit reliable extrapolation abilities for nuclei that are relatively proximate to the training region, and have fast simulation speeds, it will be straightforward and easy to extend the predictions to unknown data once new experimental data are measured. In this study, even though the prediction results of LightGBM and SVGP can be tested with only a limited number of data points as a result of the challenges associated with measuring these points far from stability line nuclei, predictions for over 300 nuclei have already been made and are included in the supplementary material [82]. Further experimental data from rare-isotope facilities such as HIAF, GANIL, FAIR, FRIB, and RIKEN will be crucial for validating these extrapolated results. Additionally, these new data may provide a reference for future experimental measurements.

Finally, to conduct a comprehensive study on the excitation energies of low-lying excited sates, we also trained both models to study the excitation energies of $2_1^+,2_2^+,4_1^+,0_2^+,0_3^+$ states; the results are listed in Table 2. One can see that for all the states examined in this work, the typical rms values obtained on the validation set are approximately 0.02 to 0.07, which indicates that the differences between the experimental and theoretical results range from 1.05 to 1.17 times. This performance is significantly better than the results provided by theoretical nuclear models. When the findings are compared with those reported in Ref. [57], it can be observed that incorporating more features that contain physical information leads to smaller rms values. This is because only 5 features were used in Ref. [57], whereas our work employed more than 16 features. In addition, the rms of SVGP on the validation set is very close to that on the training set, while the rms of LightGBM on the validation set is larger than that on the training set. This is because LightGBM has a very deep structure, allowing it to achieve high accuracy on the training set. Overall, one can observe that we adopt state-of-the-art ML algorithms and consider a wide range of physically meaningful input features, achieving the best accuracy across multiple prediction tasks. A key strength of our approach lies in the high dimensionality of the input features, which essentially encode our current comprehensive understanding of nuclear excitation energies. The combination of high-dimensional inputs, measured experimental data, and advanced ML algorithms allows us to generate predictive results that can be compared with future experimental observations. Although the SVGP model performs slightly worse than LightGBM in terms of accuracy on certain tasks, as

Table 2. Average rms values for the excitation energies of $2_1^+,4_1^+,2_2^+,0_2^+,3_1^-,0_3^+$ states.

Energy	Number of Nuclei	Feature Number	Training Ratio	Algorithm	Training Set rms	Validation Set rms	Reference
	622	5	0.9	LightGBM		0.130	Ref. [57]
	629	3	0.9	BNN		0.048	Ref. [55]
$E(2_1^+)$	660	16	0.9	LightGBM	0.0079(4)	0.030(9)	Ref. [56]
	660	16	0.8	LightGBM	0.032(4)	0.049(5)	This work
	660	16	0.8	SVGP	0.066(2)	0.070(6)	This work
	594	3	0.9	BNN		0.035	Ref. [55]
$E(4_1^+)$	608	16	0.9	LightGBM	0.0071(1)	0.020(3)	Ref. [56]
	608	16	0.8	SVGP	0.050(2)	0.046(2)	This work
$E(2_2^+)$	437	16	0.8	LightGBM	0.040(3)	0.058(5)	This work
				SVGP	0.067(7)	0.070(6)	This work
	338	3	0.9	BNN		0.063	Ref. [55]
$E(0_2^+)$	323	16	0.8	LightGBM	0.046(3)	0.064(4)	Ref. [83]
	323	16	0.8	SVGP	0.075(4)	0.072(3)	This work
$E(3_1^-)$	316	16	0.8	LightGBM	0.035(1)	0.047(3)	Ref. [83]
				SVGP	0.050(3)	0.055(3)	This work
$E(0_3^+)$	225	16	0.8	LightGBM	0.050(2)	0.064(3)	Ref. [83]
				SVGP	0.062(3)	0.067(4)	This work

discussed in the manuscript, these two algorithms are fundamentally different in their design. SVGP, similar to BNN, estimates parameter distributions during training and thus provides predictive uncertainties. By contrast, LightGBM estimates uncertainty by repeatedly training on different randomly selected subsets of the data and computing the statistical distribution of the predictions. Exploring low-lying excitation energies using both types of models offers complementary insights and adds valuable robustness to our investigation of the systematic behavior of excitation energies. In addition, we place great importance on the verifiability of our research and have organized and uploaded the code and datasets used in this study to a public GitHub repository: https://github.com/lzl888-afk/Study-the-yrast-and-yrare-2-states-using-machine-learning-approaches.

IV. SUMMARY

In this study, the excitation energies of the 2_1^+ and 2_2^+ states for even-even nuclei were evaluated using two machine learning algorithms: LightGBM (a decision tree-based method) and SVGP (a Bayesian approach). This is the first attempt to apply machine learning algorithms to predict $E(2_2^+)$, marking the inaugural use of data-driven approaches for this observable. Based on the $M16$ feature space, the rms deviations of LightGBM and SVGP with respect to the experimental $\log_{10}E$ values are 0.049 ± 0.005 (1.12 times) and 0.070 ± 0.006 (1.18 times) for 2_1^+, respectively. For $E(2_2^+)$, the rms deviations of LightGBM and SVGP with respect to the experimental $\log_{10}E$ values are 0.058 ± 0.005 (1.14 times) and 0.077 ± 0.006 (1.19 times), respectively. Both methodologies substantially outperform the prediction results of the HFB+5DCH model [25], effectively capturing the behaviors of the excitation energies of the 2_1^+ and 2_2^+ states over nuclear the landscape. A comparative analysis between LightGBM and SVGP reveals that LightGBM consistently delivers superior performance compared to that of SVGP across both training and validation datasets. However, they exhibit very similar behaviors on the limited new dataset. These findings underscore the potential of machine learning as a formidable alternative to traditional theoretical models in predicting excitation energies, showcasing clear and substantial success. In conclusion, the results of our investigation of two machine learning frameworks not only corroborate the findings of previous studies but also establish LightGBM and SVGP as useful and effective tools for machine learning applications in nuclear physics.

ACKNOWLEDGMENTS

The authors are grateful to the C3S2 computing center in Huzhou University for calculation support.

References

[1] I. A. Egorova, R. J. Charity, L. V. Grigorenko *et al.*, Phys. Rev. Lett. **109**, 202502 (2012)
[2] M. Pfützner, M. Karny, L. V. Grigorenko *et al.*, Rev. Mod. Phys. **84**, 567 (2012)
[3] F. Wamers, J. Marganiec, F. Aksouh *et al.*, Phys. Rev. Lett. **112**, 132502 (2014)
[4] I. Mukha, K. Sümmerer, L. Acosta *et al.*, Phys. Rev. Lett. **99**, 182501 (2007)
[5] K. W. Brown, R. J. Charity, L. G. Sobotka *et al.*, Phys. Rev. Lett. **113**, 232501 (2014)
[6] Y. Jin, C. Y. Niu, K. W. Brown *et al.*, Phys. Rev. Lett. **127**, 262502 (2021)
[7] K. Tshoo, Y. Satou, H. Bhang *et al.*, Phys. Rev. Lett. **109**(2), 022501 (2012)
[8] A. Spyrou, Z. Kohley, T. Baumann *et al.*, Phys. Rev. Lett. **108**, 102501 (2012)
[9] S. I. Sidorchuk, A. A. Bezbakh, V. Chudoba *et al.*, Phys. Rev. Lett. **108**, 202502 (2012)
[10] P. Wady, J. Smith, E. Paul *et al.*, Phys. Rev. C—Nucl. Phys. **85**(3), 034329 (2012)
[11] D. Hodge, D. Cullen, M. Taylor *et al.*, Phys. Rev. C **94**(3), 034321 (2016)
[12] H. L. Tann, *Excited states in the highly deformed proton emitter 131Eu*, The University of Liverpool (United Kingdom), 2022
[13] D. Seweryniak, B. Blank, M. Carpenter *et al.*, Phys. Rev. Lett. **99**(8), 082502 (2007)
[14] T. Otsuka, A. Gade, O. Sorlin *et al.*, Rev. Mod. Phys. **92**, 015002 (2020)
[15] D. Steppenbeck, S. Takeuchi, N. Aoi *et al.*, Nature **502**(7470), 207 (2013)
[16] R. Taniuchi, C. Santamaria, P. Doornenbal *et al.*, Nature **569**(7754), 53 (2019)
[17] K. Heyde and J. L. Wood, Rev. Mod. Phys. **83**, 1467 (2011)
[18] P. E. Garrett, M. Zielińska, E. Clément *et al.*, Prog. Part. Nucl. Phys. **124**, 103931 (2022)
[19] M. G. Mayer, Phys. Rev. **75**, 1969 (1949)
[20] M. Hamermesh, Phys. Rev. **75**, 1766 (1949)
[21] T. Nikšić, D. Vretenar, and P. Ring, Prog. Part. Nucl. Phys. **66**(3), 519 (2011)
[22] P. Zhao, P. Ring, and J. Meng, Phys. Rev. C **94**(4), 041301 (2016)
[23] J. L. Egido, Phys. Scrip. **91**(7), 073003 (2016)
[24] J. Meng and P. Zhao, *Relativistic density-functional theories*, in: *Handbook of Nuclear Physics* (Springer, 2023) p. 2111
[25] J.-P. Delaroche, M. Girod, J. Libert *et al.*, Phys. Rev. C **81**(1), 014303 (2010)
[26] S. Quan, Q. Chen, Z. Li *et al.*, Phys. Rev. C **95**(5), 054321 (2017)
[27] P. Ring and P. Schuck, *The nuclear many-body problem* (Springer Science & Business Media, 2004)
[28] W.-B. He, Y.-G. Ma, L.-G. Pang *et al.*, Nucl. Sci. Tech. **34**(6), 88 (2023)

[29] A. Boehnlein, M. Diefenthaler, N. Sato *et al*., Rev. Mod. Phys. **94**(3), 031003 (2022)
[30] W. He, Q. Li, Y. Ma *et al*., Sci. Chin. Phys. Mech. Astron. **66**(8), 282001 (2023)
[31] G. Aarts, K. Fukushima, T. Hatsuda *et al*., Nat. Rev. Phys. **7**, 154 (2025)
[32] N. J. Costiris, E. Mavrommatis, K. A. Gernoth *et al*., Phys. Rev. C **80**, 044332 (2009)
[33] Z. M. Niu, H. Z. Liang, B. H. Sun *et al*., Phys. Rev. C **99**, 064307 (2019)
[34] Y.-F. Gao, B.-S. Cai, and C.-X. Yuan, Nucl. Sci. Tech. **34**(1), 9 (2023)
[35] J. M. Munoz, S. Akkoyun, Z. P. Reyes *et al*., Phys. Rev. C **107**(3), 034308 (2023)
[36] A. Jalili, F. Pan, Y. Luo *et al*., Phys. Rev. C **111**(3), 034321 (2025)
[37] C.-Q. Li, C.-N. Tong, H.-J. Du *et al*., Phys. Rev. C **105**(6), 064306 (2022)
[38] N.-N. Ma, T.-L. Zhao, W.-X. Wang *et al*., Phys. Rev. C **107**(1), 014310 (2023)
[39] D. F. Rojas-Gamboa, N. G. Kelkar, and O. L. Caballero, Phys. Rev. C **110**, 035804 (2024)
[40] Z. Jin, M. Yan, H. Zhou *et al*., Phys. Rev. C **108**, 014326 (2023)
[41] H.-Q. You, R.-H. Wu, H.-Z. Su *et al*., Phys. Rev. C **110**, 024319 (2024)
[42] X.-X. Dong, R. An, J.-X. Lu *et al*., Phys. Lett. B **838**, 137726 (2023)
[43] X.-X. Dong, R. An, J.-X. Lu *et al*., Phys. Rev. C **105**(1), 014308 (2022)
[44] M. Mumpower, T. Sprouse, A. Lovell *et al*., Phys. Rev. C **106**(2), L021301 (2022)
[45] Z.-P. Gao, Y.-J. Wang, H.-L. Lü *et al*., Nucl. Sci. Tech. **32**(10), 109 (2021)
[46] Z.-X. Yang, X.-H. Fan, P. Yin *et al*., Phys. Lett. B **823**, 136650 (2021)
[47] X. H. Wu, Z. X. Ren, P. W. Zhao, Phys. Rev. C **105**, L031303 (2022)
[48] T.-S. Shang, J. Li, and Z.-M. Niu, Nucl Sci. Tech. **33**(12), 153 (2022)
[49] N. Hizawa, K. Hagino, and K. Yoshida, Phys. Rev. C **108**, 034311 (2023)
[50] Y. Wang and Q. Li, Front. Phys. **18**(6), 64402 (2023)
[51] Y. Wang, F. Li, Q. Li *et al*., Phys. Lett. B **822**, 136669 (2021)
[52] N. Cox, X. Grundler, and B.-A. Li, Phys. Rev. C **110**, 044604 (2024)
[53] W. Fan, G. Vujanovic, S. A. Bass *et al*., Phys. Rev. C **109**, 064903 (2024)
[54] S. Akkoyun, H. Kaya, and Y. Torun, Indian J. Phys. **96**(6), 1791 (2022)
[55] Y. Wang, X. Zhang, Z. Niu *et al*., Phys. Lett. B **830**, 137154 (2022)
[56] B. Lv, Z. Li, Y. Wang *et al*., Phys. Lett. B **857**, 139013 (2024)
[57] H. Liu, X.-X. Li, Y. Yuan *et al*., Nucl. Sci. Tech. **36**(2), 21 (2025)
[58] G. Ke, Q. Meng, T. Finley *et al*., in: *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc., 2017
[59] W. J. Maddox, S. Stanton, and A. G. Wilson, *Advances in Neural Information Processing Systems*, vol. 34, p.6365 (2021)
[60] F. Leibfried, V. Dutordoir, and S. John, arXiv: 2012.13962
[61] National Nuclear Data Center, Brookhaven National Laboratory http://www.nndc.bnl.gov/.
[62] P. Möller, A. J. Sierk, T. Ichikawa *et al*., Atom. Data Nucl. Data Tables **109**, 1 (2016)
[63] M. Liu, N. Wang, Y. Deng *et al*., Phys. Rev. C **84**, 014333 (2011)
[64] R. Casten, *Nuclear structure from a simple perspective*, Vol. 23 (Oxford university press, 2000)
[65] S. M. Lundberg and S.-I. Lee, *Advances in neural information processing systems*, p30 (2017)
[66] F. Emmert-Streib, O. Yli-Harja, and M. Dehmer, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **10**(6), e1368 (2020)
[67] F. Xu, H. Uszkoreit, Y. Du *et al*., Explainable ai: *A brief survey on history, research areas, approaches and challenges*, in: *Natural Language Processing and Chinese Computing: 8th CCF International Conference*, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8 (Springer, 2019), p. 563
[68] A. Thampi, Interpretable AI: *Building explainable machine learning systems* (Simon and Schuster, 2022)
[69] J. Feng, X. Gao, K. Xu *et al*., J. Hazardous Materials **490**, 137787 (2025)
[70] Z.-Y. Feng, J.-L. Tian, T. Wu *et al*., Nucl. Sci. Tech. **35**(6), 93 (2024)
[71] H. Wang, M. Yasuda, Y. Kondo *et al*., Phys. Lett. B **843**, 138038 (2023)
[72] Z. Elekes, V. Panin, T. Rodríguez *et al*., Phys. Lett. B **844**, 138072 (2023)
[73] R.-B. Gerst, A. Blazhev, K. Moschner *et al*., Phys. Rev. C **105**(2), 024302 (2022)
[74] K. Wimmer, P. Ruotsalainen, S. Lenzi *et al*., Phys. Lett. B **847**, 138249 (2023)
[75] B. Moon, W. Korten, K. Wimmer *et al*., Phys. Lett. B **858**, 139047 (2024)
[76] S. Chen, F. Browne, P. Doornenbal *et al*., Phys. Lett. B **843**, 138025 (2023)
[77] B. Monteagudo, F. M. Marqués, J. Gibelin *et al*., Phys. Rev. Lett. **132**, 082501 (2024)
[78] A. Fernández, A. Jungclaus, P. Doornenbal *et al*., Phys. Lett. B **823**, 136784 (2021)
[79] M. Cortés, W. Rodriguez, P. Doornenbal *et al*., Phys. Lett. B **800**, 135071 (2020)
[80] W. Urban, T. Rząca-Urban, A. G. Smith *et al*., Phys. Rev. C **102**, 064321 (2020)
[81] W. Q. Zhang, A. N. Andreyev, Z. Liu *et al*., Phys. Rev. C **106**, 024317 (2022)
[82] Supplemental Material for studying the yrast and yrare 2+ states using machine learning approaches.
[83] B. Lv, Y. Wang, Z. Li *et al*., Phys. Rev. C **111**(6), 064324 (2025)