

Predictions of unknown masses using a feedforward neural network*

Cheng-wei Dai (戴程威)¹ Hui Jiang (姜慧)^{1†} Yang Lei (雷杨)²

¹School of Science, Shanghai Maritime University, Shanghai 201306, China

²School of Nuclear Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China

Abstract: In this study, a feedforward neural network (FNN) approach is employed to optimize three local mass models (GK, GKs, and GK+J). We find that adding physical quantities related to the pairing effect in the input layer can effectively improve the prediction accuracy of local models. For the known masses in AME2012, the FNN reduces the root-mean-square deviation between theory and experiment for the three mass models by 11 keV, 32 keV, and 623 keV. Among them, the improvement effect of the light mass region with mass number between 16 and 60 is better than that of medium and heavy mass regions. The approach also has good optimization results when extrapolating AME2012 to AME2020 and the latest measured masses after AME2020. Based on the improved mass data, the separation energies for single- and two-proton (neutron) emissions and α -decay energies are obtained, which agree well with the experiment.

Keywords: atomic mass, local mass model, feedforward neural network

DOI: 10.1088/1674-1137/add10a **CSTR:** 32044.14.ChinesePhysicsC.49094111

I. INTRODUCTION

Atomic mass (or mass for short) is one of the fundamental quantities of a nucleus. Atomic mass data are crucial in nuclear physics, astrophysics, and nuclear technology [1, 2]. They reveal the interaction mechanisms between nucleons, including strong, weak, and electromagnetic interactions [3], as well as the resulting shell effects [4, 5] and deformations [6, 7]. Although significant advances have been made in the measurement of mass [8–11], the masses of many unstable nuclei far from the β -stability line are unknown, such as most nuclei involved in the rapid neutron capture process (r-process) and many short-lived neutron-rich radioactive nuclei [12, 13].

Mass prediction has been one of the popular topics in nuclear structure theory. Generally, theoretical mass models can be divided into global and local types. Global mass models describe masses by considering the macroscopic and/or microscopic properties of the nucleus, such as the liquid drop model (LDM) [14], Bethe-Weizsäcker (BW) model [15, 16], relativistic mean-field (RMF) model [17–19], Duflo-Zuker (DZ) model [20], finite-range droplet model (FRDM) [21], Skyrme-Hartree-Fock-Bogoliubov (SHFB) theory [22–24], and modified Weizsäcker-Skyrme (WS) mass formula [25–28]. Local mass models are characterized by having fewer model parameters and simpler calculations, enabling them to ac-

curately describe and predict the masses of nuclei near the β -stability line, such as the Audi-Wapstra extrapolation [8–11], Garvey-Kelson (GK) mass relation [29–31], generalized GK mass relations (GKs) [32], improved Jänecke mass formula (GK+J) [33], and mass relations based on proton-neutron interactions [34, 35]. For a comprehensive review, see Ref. [36].

The use of machine learning (ML) to analyze and predict nuclear data is one of the focuses in the field of nuclear physics, which is useful for further understanding nuclear structures and reaction mechanisms [37–40]. In literature, neural networks have significantly progressed in optimizing global mass models [41–52], as well as predicting α -decay half-life [53], level density [54], charge density [55], etc. For example, a feedforward neural network (FNN) reduced the root-mean-square deviations (RMSDs) of the LDM from 2.38 MeV to 196 keV using multiple hidden layers [41]. A Bayesian neural network (BNN) was used to improve the nuclear mass predictions of six global mass models, and better predictive performance can be achieved if more physical features are included [43]. By combining the global nuclear mass model and local features, a convolutional neural network (CNN) achieved good optimization results in both training sets and extrapolating new masses [47]. Because the RMSDs of local mass models are low (mostly tens to hundreds of keV), research on the optimization of local

Received 11 March 2025; Accepted 25 April 2025; Published online 26 April 2025

* Supported by National Natural Science Foundation of China (11875188) and China Scholarship Council (202409390020)

† E-mail: huijiang@shmtu.edu.cn (Corresponding author)

©2025 Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

mass models using neural networks is insufficient. We remark that the back-propagation (BP) neural network has been applied to improve the local mass relations that connect with the proton-neutron interactions for nuclei with $A \geq 100$ [56]. Although the improvement to the local mass model is modest, it has the advantage that the RMSD in a larger mass region does not increase [56].

The aim of this study is to improve the local mass models of GK, GKs, and GK+J within the framework of the FNN. Among various ML methods, the FNN has the advantages of simple architecture, strong ability to handle low-dimensional inputs and non-spatial features, high interpretability, and powerful function approximation capability [41–42, 53–54, 57–58]. Therefore, the FNN is selected to optimize the local mass models in this paper. By carefully designing the input, hidden, and output layers of the FNN, we have significantly improved the prediction accuracy of the three models. For known masses in AME2012, the RMSDs from experimental values of GK, GKs, and GK+J are reduced by 11 keV, 32 keV, and 623 keV, respectively. For the new masses from AME2012 to AME2020, the RMSD reductions for the three models are 44 keV, 20 keV, and 963 keV, respectively.

The remainder of this paper is organized as follows. In Sec. II, the FNN is briefly introduced. In Sec. III, four neural network structures are designed, and the optimal network structure of three local mass models is discussed. We investigate our improvements in both descriptions and predictions of mass excess. In Sec. IV, one/two-proton/neutron separation energies (S_p, S_{2p}, S_n, S_{2n}), and α -decay energies Q_α are investigated. In particular, the α -decay energy of nuclei with proton number Z from 82 to 108 and its odd-even staggering (OES) are discussed. The summary and conclusions are given in Sec. V.

II. THEORETICAL FRAMEWORK

A neural network is a type of computational model that emulate the operational principles of the human brain, consisting of multiple layers of nodes (neurons), which typically include an input layer, hidden layers, and an output layer. Each neuron receives input signals from the previous layer, applies an activation function after weighted summation, generates output signals and passes them to the next layer.

The FNN constructed in this paper contains an input layer, one hidden layer, and an output layer. The activation function is selected to be the hyperbolic tangent function. Let $x = \{x_i\}$ be the input of the neural network, h_j be the output value of the hidden layer node j , and $\hat{y}$ be the output of the neural network. Thus, we have [58]

$$\hat{y}(x; w) = a + \sum_{j=1}^H w_j^{(2)} h_j, \quad (1)$$

with

$$h_j = \tanh \left(b_j + \sum_{i=1}^I w_{ji}^{(1)} x_i \right).$$

Here, $w = \{a, b_j, w_j^{(2)}, w_{ji}^{(1)}\}$ denotes the neural network parameter. a is a constant term that aids in adjusting the output layer's prediction; b_j is the bias term for the hidden layer neurons, which is used to adjust the activation values; $w_{ji}^{(1)}$ and $w_j^{(2)}$ are the weight matrices representing the connections from the input layer to the hidden layer and from the hidden layer to the output layer, respectively, and determine the strength of the connections between layers. H and I are the numbers of hidden layer neurons and inputs, respectively. During the initialization process of the parameter w , a and b_j are set as zero vectors, and $w_{ji}^{(1)}$ and $w_j^{(2)}$ are generated through Xavier initialization based on a normal distribution. To ensure reproducibility of the results, we adopt a random seed strategy to maintain the weight matrices generated during each initialization consistent.

Given a training set $P = \{(x_i, y_i) \mid i \in \{1, 2, \dots, N\}\}$ consisting of N data points and a loss function, the parameters w of the neural network can be trained. For the neural network regression prediction problem, the RMSD function is typically selected as the loss function, which is defined as follows.

$$L(y, \hat{y}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}. \quad (2)$$

The neural network optimizes the parameter w by minimizing the value of the loss function, learning to reduce the gap between the predicted and actual values to improve prediction accuracy. Here, the Adam algorithm [59] is adopted to optimize the neural network. It effectively improves the training efficiency and performance of the neural network by avoiding gradient issues, accelerating convergence, and utilizing prior knowledge to enhance model performance. During the training process, a back-propagation algorithm is used to calculate the error and propagate the error from the output layer back to the hidden and input layers layer by layer, gradually adjusting the weights to optimize the performance of the network.

III. MASS PREDICTION

To study the description and prediction of mass excess in different neural networks, we design four neural network structures. Because the neural network with a single hidden layer is easier to understand, more economical in terms of computational resources, and also reduces the risk of overfitting, a single hidden layer neural

network is adopted. Specific structural parameters, including the input variables, number of hidden layer neurons H , output, and total number of neural network parameters $\mathbb{N}$ are presented in Table 1. The H values of the four structures are changed to ensure that total number of parameters in the neural network is consistent. Here, N and Z represent the neutron and proton numbers, respectively. $P = N_p N_n / (N_p + N_n)$, with the valence proton number N_p and valence neutron number N_n , is the average number of interactions of each valence nucleon with those of the other type, which is a useful indicator of the pairing versus p-n competition [60]. Z_{eo} and N_{eo} are related to the pairing effect. If Z (N) is even, Z_{eo} (N_{eo}) equals 1; otherwise, it is zero. N_{shell} and Z_{shell} represent the shell model orbitals of the last proton and neutron [50], associated with the shell effects. The values of N_{shell} and Z_{shell} are defined as 1, 2, 3, 4, or 5, depending on whether the proton or neutron number falls within the specified ranges [8, 28], [29, 50], [51, 82], [83, 126], and [127, 184], respectively. The output is $\Delta M = M^{\text{expt}} - M^{\text{th}}$, that is, the deviation between the experimental data and the theoretical prediction of mass excess.

The neural network is trained based on the AME2012 experimental database, for nuclei with $N \geq 8$ and $Z \geq 8$. The distribution of the training and test sets on the nuclear chart is presented in Fig. 1. Here, the ratio of the number of known masses contained in the training set to the test set is 9 : 1. The partitioning of these two data sets is

Table 1. Input variables, number of hidden layer neurons H , output, and total number of parameters $\mathbb{N}$ in four neural network structures. Here, one hidden layer is used.

Structure	Input Variables	H	Output	$\mathbb{N}$
A	N, Z	30	ΔM	121
B	N, Z, P	24	ΔM	121
C	$N, Z, N_{\text{eo}}, Z_{\text{eo}}$	20	ΔM	121
D	$N, Z, N_{\text{shell}}, Z_{\text{shell}}$	20	ΔM	121

random, but all data must cover all regions of the nuclear chart. The nuclei in the test set are distributed both at the edges and within the interior of the nuclear chart, which guarantees the effectiveness and reliability of the testing process. Because the number of known masses in AME2012 described by GK, GKs, and GK+J models is 2265, 2318, and 2310 respectively, the training and test sets of the three models are not completely consistent, but they all satisfy the requirement of 9 : 1. The sum of the test and training sets is called the full dataset. To evaluate the extrapolation capability of the neural network, we construct one extrapolation set (denoted by "12-20") based on AME2020, including experimental masses that appear in AME2020 but are absent in AME2012. Because the vast majority of nuclei in the extrapolation set spanning from AME2012 to AME2016 are already contained in the 12–20 dataset, only one 12–20 extrapolation set is constructed in this paper. We also experiment with partitioning ratios of 7 : 3 and 8 : 2. The results indicate that while different ratios yield similar RMSDs on the training and test sets, the 9 : 1 ratio performs optimally on the extrapolation. Because the nuclei in the extrapolation set are primarily distributed at the edge of the nuclear chart, the performance of extrapolation is related to the distribution of the training set in the boundary region. Compared with 7 : 3 and 8 : 2, the 9 : 1 ratio makes the training set contain more edge samples, which can help the model learn the boundary behavior better, thus strengthening the extrapolation ability.

The steps to predict mass excess with the FNN are as follows. First, we construct a neural network using the training set. The input variables are shown in Table 1. The output of the neural network is uniformly given by $\Delta M(N, Z) = M^{\text{expt}}(N, Z) - M_0^{\text{th}}(N, Z)$. Here, M_0^{th} represents the theoretical predicted values of local mass model. Second, the optimal hyperparameters are determined through a grid-based parameter tuning process. For the GK, GKs, and GK+J models, the learning rates are set to 0.0094, 0.0095, and 0.0115 respectively, with correspond-

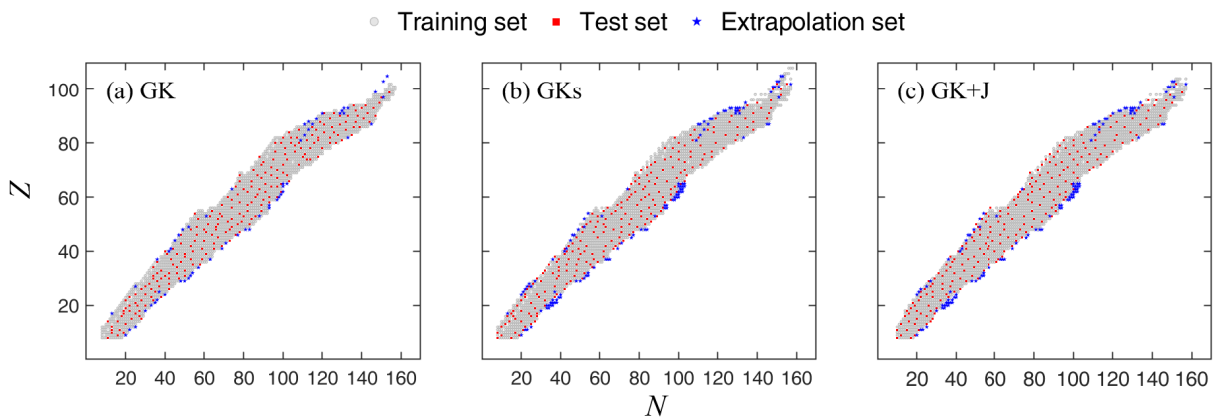


Fig. 1. (color online) Training, test, and extrapolation sets on the nuclear chart for the GK, GKs, and GK+J models.

ing training iteration counts of 50000, 60000, and 50000. Third, when the input, output, and hyperparameters are determined, neural network parameters $w = \{a, b_j, w_j^{(2)}, w_{ji}^{(1)}\}$ can be obtained using Eq. (1). Finally, the predicted deviation $\Delta M'(N, Z)$ of mass excess by the FNN is reconstructed by Eq. (1). The predicted value of the FNN is expressed as $M_{\text{FNN}}^{\text{th}}(N, Z) = \Delta M'(N, Z) + M_0^{\text{th}}(N, Z)$.

The performance of four neural network structures in predicting nuclear masses are evaluated on the training set, test set, and full dataset for three local mass models, as shown in Fig. 2. Here, σ_0 and σ represent the mass RMSDs between theory and experiment initially and after FNN training, respectively. A ratio σ/σ_0 less than 1 indicates that the prediction of the theoretical model has improved after FNN training. We observe that, based on the known masses in the AME2012, the FNN neural network performs optimally with the C structure for the GK, GKs, and GK+J models. Furthermore, Fig. 3 shows the extrapolation of these neural networks over the mass from AME2012 to AME2020, where the optimal network structure coincides with those identified in Fig. 2. This indicates that the pairing effect is the most effective correction in improving the GK, GKs, and GK+J models. It is known that the GK mass relations are constructed such that the neutron-neutron, proton-proton, and neutron-proton interactions are canceled at the first order. The GK mass relations are not strictly zero. Refs. [34, 61] observed that, for different parity combinations of neutrons and protons, the GK relations have an odd-even staggering pattern with respect to the average deviation from zero, which dominantly originates from the pairing interaction. Therefore, the result that the "pairing effect" as the physical input is the optimal network structure is consistent with the above conclusion.

Table 2 shows mass RMSDs of the three local mass models under the network structure C for the training, test, and full sets in AME2012, as well as the results of extrapolation set from AME2012 to AME2020 (denoted by "12–20"). Here, $\Delta = \sigma_0 - \sigma$ represents the improvement of the RMSD after FNN training. $\mathcal{N}$ is the number

of nuclei in the corresponding dataset. We observe that, for the known nuclei in AME2012, the RMSD of full dataset is reduced by 11, 32, and 623 keV for the GK, GKs, and GK+J models, respectively; for the extrapolation set from AME2012 to AME2020, $\Delta = 44, 20$, and 963 keV, respectively. These results demonstrate that the FNN can significantly improve the prediction accuracy of local mass models. For the GK model, the improvement achieved by the FNN on the test set is minor ($\Delta = 2$ keV). We have experimented with altering the division between the training and test sets. Although this adjustment can appropriately increase the improvement on the test set, the optimal improvement for the full dataset remains at approximately 11 keV, and the extrapolation performance is inferior to the existing results. As shown in Table 2 (and Fig. 4), the original predictive accuracy of the GK model is already quite high, with an RMSD of 142 keV on the test set. Achieving an overall improvement of 11 keV based on such a precise dataset is our current best result. To further validate our results, we replace the mass values in the AME2012 full dataset with the updated values from AME2020 and recalculate the RMSDs in Table 2. The improvements Δ obtained using the FNN for the three models are very close to those in Table 2. This indicates that the optimization effect of the FNN is almost

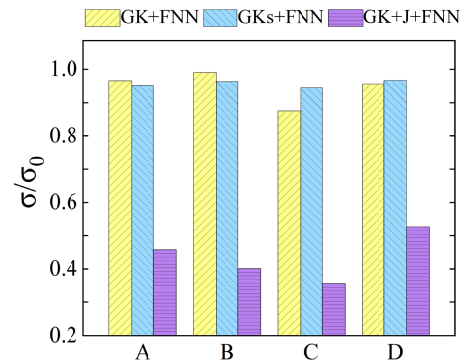


Fig. 3. (color online) Same as Fig. 2 but for the extrapolation set from AME2012 to AME2020.

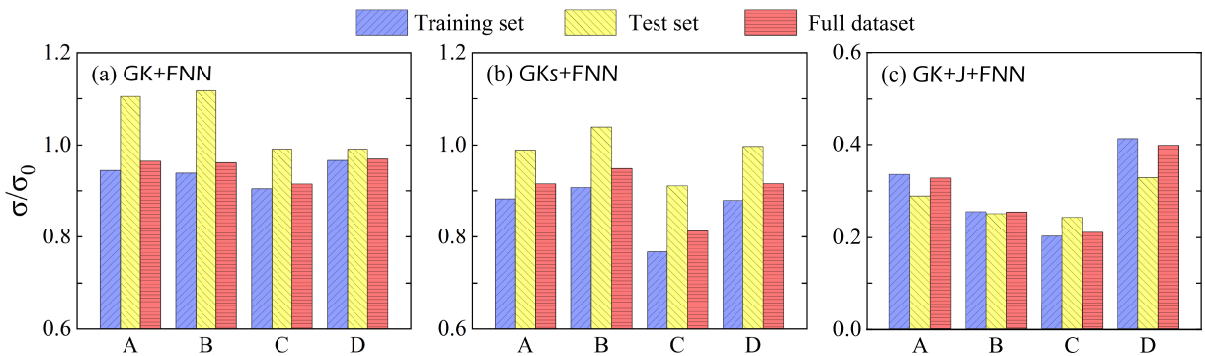


Fig. 2. (color online) Mass RMSD ratios (σ/σ_0) of four network structures of local mass model GK, GKs, and GK+J. Results of the training set, test set, and full dataset are presented by blue, yellow, and red, respectively.

Table 2. Mass RMSDs (in keV) of the GK, GKs, and GK+J models in the neural network structure C for the training, test, and full sets in AME2012, and the extrapolation set from AME2012 to AME2020 (denoted by "12-20"). Here, σ_0 and σ represent the RMSDs between theory and experiment initially and after FNN training, respectively. $\Delta = \sigma_0 - \sigma$. N is the number of nuclei in the corresponding dataset.

Data	GK+FNN				GKs+FNN				GK+J+FNN			
	Train	Test	Full	12-20	Train	Test	Full	12-20	Train	Test	Full	12-20
σ_0	131	142	132	355	154	300	174	348	750	1073	788	1496
σ	118	140	121	311	118	274	142	328	151	261	165	533
Δ	13	2	11	44	36	26	32	20	599	812	623	963
N	2039	226	2265	61	2087	231	2318	118	2079	231	2310	113

unaffected by the version update of experimental data.

To evaluate the performance of the FNN in optimizing local mass models in different mass regions, the RMSDs in light ($16 \leq A < 60$), medium ($60 \leq A < 120$), and heavy ($A \geq 120$) mass regions of the full dataset in the AME2012 are given in Table 3. The initial mass deviation $\Delta M_0 = M^{\text{expt}} - M_0^{\text{th}}$ of the three models and the mass deviation $\Delta M_1 = M^{\text{expt}} - M_{\text{FNN}}^{\text{th}}$ after FNN improvement are also shown in Fig. 4. In the light mass region, the FNN significantly improves the three mass models, reducing the RMSDs between the predicted and experimental values by 44, 108, and 1802 keV, respectively. In the medium and heavy mass regions, because the local mass models themselves perform well (the mean RMSD is approximately 151 keV in the medium region and about 76 keV in the heavy region), the improvement of the FNN is not as significant as in the light mass region.

Finally, the accuracy of our predictions for the new measured masses after AME2020 is worth examining. Since the release of AME2020, the masses of about 100 more atomic nuclei have been measured experimentally. GKs and GK+J models can predict 20 of these [62–72], whereas the GK model can predict seven. The mass RMSDs after FNN improvement are shown in Table 4. The improved ability of the FNN is robust and can reduce the predictive RMSDs of the three models by 16, 27, and 408 keV, respectively.

In Table 5 and Fig. 5, the specific predicted values and corresponding mass deviations ΔM of the three models for 20 new masses are respectively presented and compared with the predictions in AME2020. Here, for the same nucleus, the predicted value with the smallest deviation from the experiment is expressed by "FNN-Best" in Table 5. According to Table 5 and Fig. 5, the three local mass models optimized using the FNN exhibit excellent performance in extrapolation. For these 20 nuclei, the RMSD between the predicted and experimental values of AME2020 is 195 keV. Of the three models, the GKs model has the best performance, with an RMSD of 181 keV between its predicted and experimental values, which is 14 keV lower than that of the AME2020. Furthermore, if the optimal prediction value "FNN-best" is

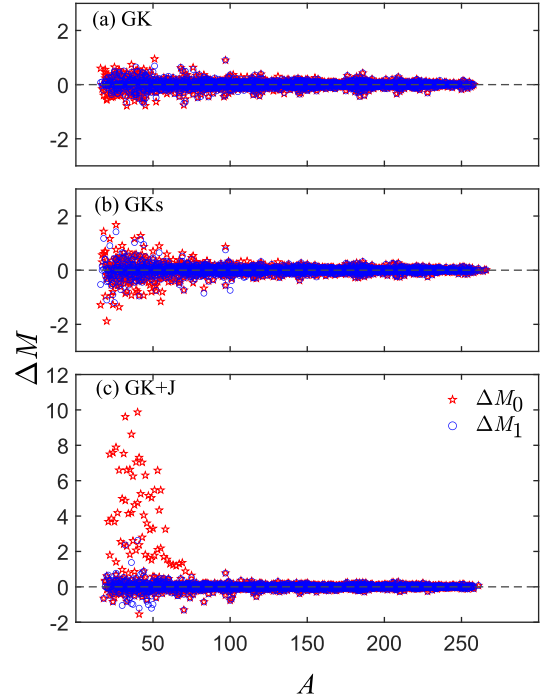


Fig. 4. (color online) Mass deviation ΔM (in MeV) of the GK, GKs, and GK+J models in the full dataset of AME2012. Here, $\Delta M_0 = M^{\text{expt}} - M_0^{\text{th}}$ represents the deviation of the model itself; $\Delta M_1 = M^{\text{expt}} - M_{\text{FNN}}^{\text{th}}$ refers to the results after the FNN improvement.

taken, the RMSD between the theoretical and experimental value is reduced to 159 keV.

IV. APPLICATION OF PREDICTED MASSES

The mass excess (M) is closely related to various forms of nuclear energy, such as single- and two-proton/neutron separation energies (S_p , S_{2p} , S_n , and S_{2n}), and α -decay energy Q_α , defined by

$$S_{ip}(Z, N) = M(Z - i, N) + iM_H - M(Z, N),$$

Table 3. Mass RMSDs (in keV) for different mass regions of the AME2012 full dataset. The definitions of σ_0 , σ , Δ , and N are the same as in Table 2.

Region	σ_0	σ	Δ	N
GK+FNN				
$16 \leq A < 60$	257	213	44	297
$60 \leq A < 120$	133	133	0	634
$A \geq 120$	80	79	2	1334
GKs+FNN				
$16 \leq A < 60$	414	306	108	300
$60 \leq A < 120$	134	131	3	632
$A \geq 120$	75	74	1	1386
GK+J+FNN				
$16 \leq A < 60$	2183	381	1802	295
$60 \leq A < 120$	185	142	42	636
$A \geq 120$	73	73	0	1379

Table 4. Same as Table 2 except for the RMSDs (in keV) for the new measured masses after AME2020.

Data	GK+FNN	GKs+FNN	GK+J+FNN
σ_0	199	208	770
σ	183	181	362
Δ	16	27	408
N	7	20	20

$$S_{in}(Z, N) = M(Z, N - i) + iM_n - M(Z, N),$$

$$Q_\alpha(Z, N) = M(Z, N) - M(Z - 2, N - 2) - M_{\text{He}}.$$

Here, S_{ip} and S_{in} represent the i -proton and i -neutron separation energies, respectively.

Using the predicted mass excess, we evaluate the separation energies for single- and two-proton (neutron) emissions, as well as the α -decay energies. The RMSDs between predicted values of three local mass models and the experimental values are listed in Table 6, for nuclei with $N \geq 8$ and $Z \geq 8$. Meanwhile, the results of FRDM12 [21] and WS4 [28] are also listed in the table for comparison. Here, the definitions of σ_0 , σ , Δ , and N are the same as in Table 2. We observe that, after the optimization using the FNN, the three local mass models are effectively improved. The RMSDs of the five energies of GK, GKs, and GK+J are reduced by $\Delta = 12.8, 49.8$, and 265.6 keV on average, respectively. In addition, the prediction accuracy of the three local models improved by the FNN is significantly higher than that of the FRDM and WS4 models. This indicates that the FNN can effectively improve the local mass models, and our prediction

results have certain competitiveness.

Next, we focus on the α -decay energy of heavy nuclei. The predicted Q_α values for nuclei with proton number Z ranging from 82 to 108 are presented in Fig. 6. Because "GKs+FNN" has the highest prediction accuracy for α -decay energy among the three local models, we only plot the results of GKs+FNN in the figure. Meanwhile, the results of the FRDM and WS4 model as well as the experimental data are also compared. As shown in the figure, the Q_α values predicted by the GKs+FNN, FRDM, and WS4 models are in excellent agreement with the experimental values (with RMSDs of 97, 276, and 269 keV, respectively). For experimentally unknown nuclei, the three theoretical models (GKs+FNN, FRDM, and WS4) predict roughly the same trend. That is, for nuclei with $N < 126$, Q_α of each isotopic chain increases as the number of neutrons decreases; for nuclei with $N > 126$, Q_α decreases as the number of neutrons increases. Notably, for the isotopes with $Z = 82, 84 - 87$, the predicted values of the three theoretical models have slightly different trends for the lightest and/or heaviest mass regions.

Recently, an interesting phenomenon has been observed: Q_α energies of nuclei with $Z > 82$ and $N < 126$ exhibit a distinct odd-even staggering (OES), which is caused by a combination of pairing correlations and blocking effect of unpaired nucleons [73]. The OES of Q_α with the number of neutrons (denoted by ΔQ_n) or protons (denoted by ΔQ_p) can be quantitatively studied by the following formula [74–76]:

$$\Delta Q_n = \frac{1}{2} [2Q_\alpha(N, Z) - Q_\alpha(N - 1, Z) - Q_\alpha(N + 1, Z)],$$

$$\Delta Q_p = \frac{1}{2} [2Q_\alpha(N, Z) - Q_\alpha(N, Z - 1) - Q_\alpha(N, Z + 1)].$$

Our predicted ΔQ_n for Rn, Fr, Ra, Ac, Th, Pa, and U isotopes and ΔQ_p for $N = 115 - 125$ isotones are presented with hollow symbols in Fig. 7 (a) and (b), respectively. The experimental values in the figure are represented by corresponding solid symbols. To avoid overlapping, each chain is translated upward by a certain amount, and the value of the translation is shown in the number beside each chain. Our results reproduce both neutron OES and proton OES patterns very well. The RMSDs between the theoretical predictions and the experimental values are 51 keV (for 70 known nuclei) in panel (a) and 31 keV (for 62 known nuclei) in panel (b), respectively. Ref. [73] indicates that the typical amplitude of OES for α -decay energy is approximately 70 keV for nuclei with $Z > 82$ and $N < 126$. The mean values of $|\Delta Q_n|$ and $|\Delta Q_p|$ predicted by the GKs+FNN are 86 keV and 68 keV, respectively. Our results agree well with the experimental observations.

Table 5. Theoretical predicted values of the newly measured mass excess. The predictions in AME2020 and experimental values are listed for comparison. Here, for the same nucleus, the value with the smallest deviation from the experimental value in the three models is labeled "FNN-Best."

Nuclei	GK+FNN	GKs+FNN	GK+J+FNN	FNN-Best	AME2020	Expt.
^{65}Cr		-28245	-28993	-28245	-28310	-28208 (45) [62]
^{69}Fe	-39424	-39208	-38936	-39424	-39199	-39504 (11) [63]
^{70}Fe		-36845	-36855	-36855	-36890	-37053 (12) [63]
^{74}Ni	-48712	-48527	-48544	-48527	-48700	-48451 (3.5) [64]
^{75}Ni		-44068	-43746	-44068	-44240	-44056 (14.7) [64]
^{60}Ca		-40019	-39947	-40019	-39590	-40005 (30) [65]
^{103}Rb		-33286	-33716	-33286	-33160	-33049 (32) [65]
^{103}Sr		-47237	-47371	-47237	-47280	-47220 (29) [66]
^{104}Sr		-43698	-44030	-43698	-43760	-43411 (33) [66]
^{105}Sr		-38037	-37970	-37970	-38190	-37886 (44) [66]
^{120}Rh	-58870	-58636	-58789	-58636	-58620	-58614 (58) [67]
^{99}In	-61410	-61267	-61124	-61410	-61376	-61429 (77) [68]
^{133}In	-57403	-57390	-57198	-57403	-57690	-57678 (41) [69]
^{134}In		-51390	-51328	-51390	-51970	-51855 (44) [69]
^{152}Ce	-58969	-58824	-58806	-58824	-58980	-58878 (23) [70]
^{153}Ce		-54764	-54860	-54764	-54910	-54712 (24) [70]
^{154}Ce		-52059	-52044	-52059	-52220	-52069 (24) [70]
^{150}Yb		-38727	-38738	-38727	-38830	-38635 (44) [71]
^{153}Yb	-47209	-47212	-47143	-47143	-47160	-47102 (46) [71]
^{251}No		82827	82780	82827	82849	82851 (23) [72]

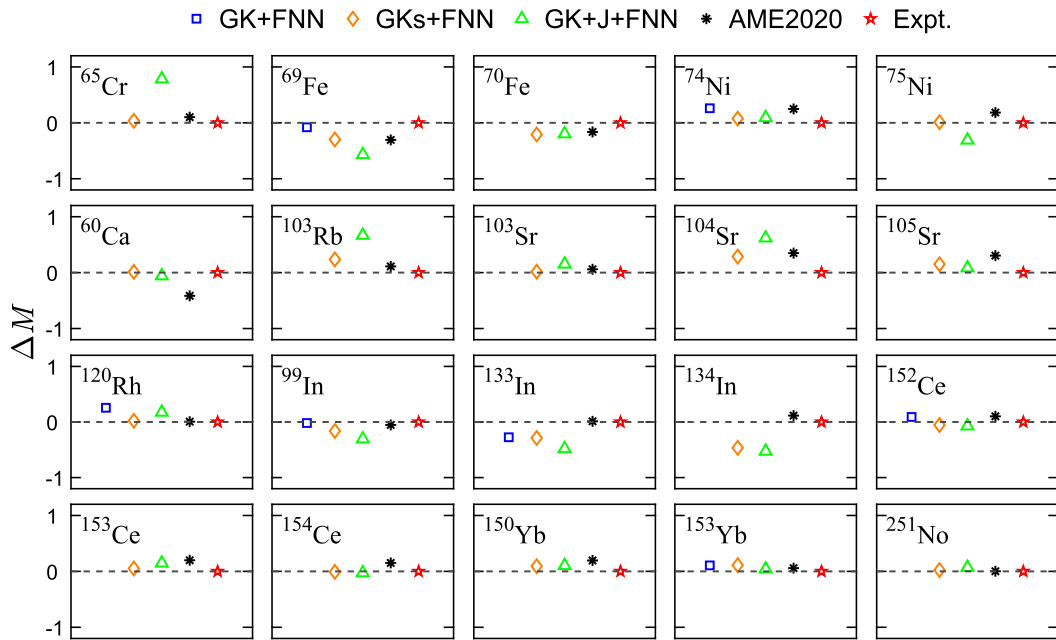


Fig. 5. (color online) Mass deviation $\Delta M = M^{\text{expt}} - M^{\text{th}}$ (in MeV) of the GK, GKs, and GK+J models for 20 newly measured masses. The AME2020 prediction results are also plotted for comparison.

Table 6. RMSDs (in keV) of single-neutron (S_n), single-proton (S_p), two-neutron (S_{2n}), and two-proton (S_{2p}) separation energies and α -decay energy Q_α with respect to the experimental data in AME2020. Here, the definitions of σ_0 , σ , Δ , and N are the same as in Table 2.

Energy	GK+FNN				GKs+FNN				GK+J+FNN				FRDM12		WS4	
	σ_0	σ	Δ	N	σ_0	σ	Δ	N	σ_0	σ	Δ	N	σ_0	N	σ_0	N
S_n	207	188	19	2181	249	198	51	2281	433	226	207	2282	350	2309	257	2309
S_p	202	187	15	2147	264	201	63	2252	416	215	201	2254	365	2281	273	2281
S_{2n}	177	168	9	2098	277	212	65	2203	688	233	455	2201	455	2223	266	2223
S_{2p}	192	180	12	2008	292	227	65	2122	683	237	446	2144	470	2134	321	2134
Q_α	213	204	9	2198	188	183	5	2337	223	204	19	2298	552	2342	344	2342

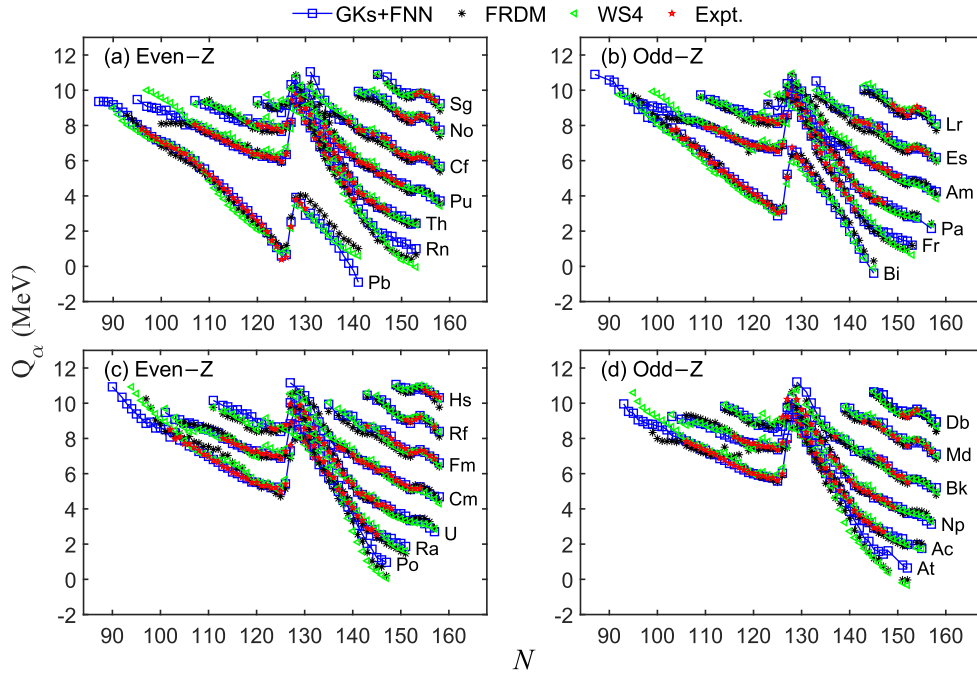


Fig. 6. (color online) α -decay energies Q_α (in MeV) for nuclei with proton number Z from 82 to 108. Panels (a,c) and (b,d) show the results of even- Z and odd- Z nuclei, respectively. Experimental data are from AME2020.

V. SUMMARY

We employed a feedforward neural network (FNN) to improve three local mass models (GK, GKs, and GK+J). The constructed neural network consists of an input layer, a hidden layer, and an output layer. By comparing different combinations of input features, we find that the physical quantities Z_{co} and N_{co} related to the pairing effect can effectively improve the predictive ability of the neural network, thus improving the prediction accuracy of the three local mass models. Our improvements in both description and prediction of mass excess are investigated. For known masses in AME2012, the FNN-improved GK, GKs, and GK+J models achieve better agreement with experimental values, reducing the RMSDs to 121 keV, 142 keV, and 165 keV, respectively. However, note that the improvement to the GK model by the FNN is minor

on the test set. For the AME2012 to AME2020 extrapolation, the FNN approach reduced the mass RMSDs to 311 keV, 328 keV, and 533 keV, respectively. In addition, this approach successfully predicts 20 new masses measured experimentally after 2020, reducing the RMSDs predicted by the three local mass models to 183 keV, 181 keV, and 362 keV, respectively, showing strong predictive power. Based on the improved mass data, one- and two-proton/neutron separation energies and α -decay energies are studied, and good optimization results are obtained. Finally, note that the quantification and representation of model uncertainty are important. Compared with the FNN, the BNN can naturally provide the uncertainties in mass predictions [43]. Therefore, we will consider using the BNN to improve local mass models and evaluate model uncertainty in the future research.

Currently, research on improving local mass models

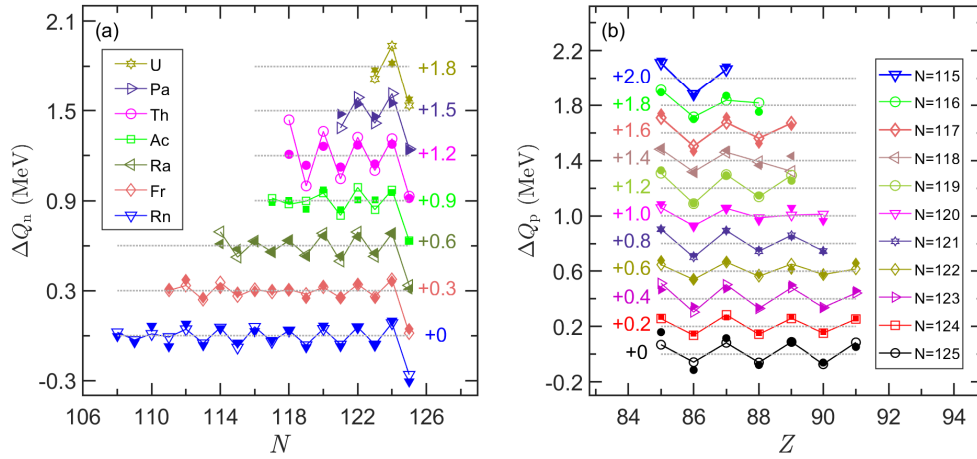


Fig. 7. (color online) Odd-even staggering (OES) pattern of α -decay energies: (a) neutron OES (ΔQ_n in MeV) versus neutron number N and (b) proton OES (ΔQ_p in MeV) versus proton number Z . The theoretical values of GKs+FNN and experimental data are represented by hollow and solid symbols, respectively. Most experimental data are obtained from the AME2020 [10–11], except for $^{207,208}\text{Th}$ [73] and $^{214,217}\text{U}$ [77]. The Q_α energy of ^{209}Th has no relevant experimental value, and the solid symbol plotted in the figure is from the theoretical prediction [78]. To avoid overlapping, a small translation is added to the $\Delta Q_n/\Delta Q_p$ values of each isotopes/isotones (the corresponding value is indicated next to each chain).

based on neural networks is insufficient, and the study reported in this paper is a useful attempt. Our results indicate that neural networks provide a new method to further improve the accuracy of local mass models in describing and predicting atomic mass related data in unknown regions. Although this paper achieves high accuracy by applying the FNN to improve three local mass models, a gap remains between local and global mass models in terms of describing and predicting masses far from known regions. Based on this, an important future direction is to integrate the advantages of local and global models (for instance, incorporating global features into local mass models or embedding local features into glob-

al mass models) and apply neural networks for improvement to further enhance the predictive capability for unknown masses.

ACKNOWLEDGMENTS

We thank Prof. Meng Wang (Institute of Modern Physics, Chinese Academy of Sciences) for providing the bibliography of new mass measurements after AME2020. We also thank Man Bao from University of Shanghai for Science and Technology and Chang Ma from Yanshan University for providing the data of GKs and GK+J models, respectively.

References

- [1] M. R. Mumpower, R. Surman, G. C. McLaughlin *et al.*, *Prog. Part. Nucl. Phys.* **86**, 86 (2016)
- [2] W. P. Liu, B. Guo, X. X. Bai *et al.*, *Nucl. Phys. Rev.* **34**, 284 (2017)
- [3] X. K. Dong and B. S. Zou, *Phys.* **51**, 611 (2022)
- [4] E. M. Ramirez, D. Ackermann, K. Blaum *et al.*, *Science* **337**, 1207 (2012)
- [5] F. Wienholtz, D. Beck, K. Blaum *et al.*, *Nature* **498**, 346 (2013)
- [6] B. A. Brown, *Phys. Rev. C* **58**, 220 (1998)
- [7] G. A. Lalazissis, T. Nikšić, D. Vretenar *et al.*, *Phys. Rev. C* **71**, 024312 (2005)
- [8] M. Wang, G. Audi, A. H. Wapstra *et al.*, *Chin. Phys. C* **36**, 1603 (2012)
- [9] M. Wang, G. Audi, F. G. Kondev *et al.*, *Chin. Phys. C* **41**, 030003 (2017)
- [10] W. J. Huang, M. Wang, F. G. Kondev *et al.*, *Chin. Phys. C* **45**, 030002 (2021)
- [11] M. Wang, W. J. Huang, F. G. Kondev *et al.*, *Chin. Phys. C* **45**, 030003 (2021)
- [12] Y. G. Ma, *Sci. Technol. Rev.* **41**, 14 (2023)
- [13] Z. D. An and C. C. Guo, *Nucl. Phys. Rev.* **37**, 643 (2020)
- [14] G. Gamow, *Proc. R. Soc. Lond. A* **126**, 632 (1930)
- [15] C. F. Weizsäcker and Z. T. D. Kernmassen, *Z. Physik* **96**, 431 (1935)
- [16] H. A. Bethe and R. F. Bacher, *Rev. Mod. Phys.* **8**, 82 (1936)
- [17] L. Geng, H. Toki, and J. Meng, *Prog. Theor. Phys.* **113**, 785 (2005)
- [18] X. M. Hua, T. H. Heng, Z. M. Niu *et al.*, *Sci. China-Phys. Mech. Astron.* **55**, 2414 (2012)
- [19] D. Peña-Arteaga, S. Goriely, and N. Chamel, *Eur. Phys. J. A* **52**, 320 (2016)
- [20] J. Duflo and A. P. Zuker, *Phys. Rev. C* **52**, R23 (1995)
- [21] P. Möller, W. D. Myers, H. Sagawa *et al.*, *Phys. Rev. Lett.* **108**, 052501 (2012)
- [22] S. Goriely, M. Samyn, and J. M. Pearson, *Phys. Rev. C* **75**, 064312 (2007)
- [23] S. Goriely, N. Chamel, and J. M. Pearson, *Phys. Rev. Lett.*

- 102**, 152503 (2009)
- [24] S. Goriely, N. Chamel, and J. M. Pearson, *Phys. Rev. C* **93**, 034337 (2016)
- [25] N. Wang, M. Liu, and X. Z. Wu, *Phys. Rev. C* **81**, 044322 (2010)
- [26] M. Liu, N. Wang, Y. Deng *et al.*, *Phys. Rev. C* **84**, 014333 (2011)
- [27] N. Wang and M. Liu, *Phys. Rev. C* **84**, 051303 (2011)
- [28] N. Wang, M. Liu, X. Z. Wu *et al.*, *Phys. Lett. B* **734**, 215 (2014)
- [29] G. T. Garvey and I. Kelson, *Phys. Rev. Lett.* **16**, 197 (1966)
- [30] G. T. Garvey, W. J. Gerace, R. L. Jaffe *et al.*, *Rev. Mod. Phys.* **41**(4), S1 (1969)
- [31] J. Barea, A. Frank, J. G. Hirsch *et al.*, *Phys. Rev. C* **77**, 041304 (2008)
- [32] M. Bao, Z. He, Y. Lu *et al.*, *Phys. Rev. C* **88**, 064325 (2013)
- [33] C. Ma, M. Bao, Z. M. Niu *et al.*, *Phys. Rev. C* **101**, 045204 (2020)
- [34] G. J. Fu, Y. Lei, H. Jiang *et al.*, *Phys. Rev. C* **84**, 034311 (2011)
- [35] H. Jiang, G. J. Fu, B. Sun *et al.*, *Phys. Rev. C* **85**, 054303 (2012)
- [36] M. Bao, H. Jiang, and Y. M. Zhao, *Nucl. Phys. Rev.* **40**, 141 (2023)
- [37] Y. G. Ma, L. G. Pang, R. Wang *et al.*, *Chin. Phys. Lett.* **40**, 122101 (2023)
- [38] W. B. He, Q. F. Li, Y. G. Ma *et al.*, *Sci. China Phys. Mech. Astron.* **66**, 282001 (2023)
- [39] R. D. Lasserri, D. Regnier, A. Penon *et al.*, *Phys. Rev. Lett.* **124**, 162502 (2020)
- [40] L. G. Pang, K. Zhou, and X. N. Wang, *Nucl. Phys. Rev.* **37**, 720 (2020)
- [41] X. K. Le, N. Wang, and X. Jiang, *Nucl. Phys. A* **1038**, 122707 (2023)
- [42] D. C. Tian, S. W. Chen, and Z. M. Niu, *Sci. Sin.-Phys. Mech. Astron.* **52**, 252007 (2022)
- [43] Z. M. Niu and H. Z. Liang, *Phys. Lett. B* **778**, 48 (2018)
- [44] T. M. Sprouse, R. Navarro Perez, R. Surman *et al.*, *Phys. Rev. C* **101**, 055803 (2020)
- [45] R. Utama, J. Piekarewicz, and H. B. Prosper, *Phys. Rev. C* **93**, 014311 (2016)
- [46] R. Utama and J. Piekarewicz, *Phys. Rev. C* **97**, 014306 (2018)
- [47] Y. Lu, T. Shang, P. Du *et al.*, *Phys. Rev. C* **111**, 014325 (2025)
- [48] T. Lu, T. Shang, P. Du, *et al.*, (2024), arXiv: 2404.14948
- [49] H. F. Zhang, L. H. Wang, J. P. Yin *et al.*, *J. Phys. G: Nucl. Part. Phys.* **44**, 045110 (2017)
- [50] E. Yüksel, D. Soydaner, and H. Bahtiyar, *Phys. Rev. C* **109**, 064322 (2024)
- [51] L. Neufcourt, Y. Cao, W. Nazarewicz *et al.*, *Phys. Rev. Lett.* **122**, 062502 (2019)
- [52] L. Neufcourt, Y. Cao, S. A. Giuliani *et al.*, *Phys. Rev. C* **101**, 044307 (2020)
- [53] X. D. Bu, D. Wu, and C. L. Bai, *Sci. Sin.-Phys. Mech. Astron.* **52**, 252005 (2022)
- [54] P. X. Du, T. S. Shang, K. P. Geng *et al.*, *Phys. Rev. C* **109**, 044325 (2024)
- [55] T. S. Shang, H. H. Xie, J. Li *et al.*, *Phys. Rev. C* **110**, 014308 (2024)
- [56] B. B. Jiao, *Int. J. Mod. Phys. E* **29**, 2050024 (2020)
- [57] J. Yang and J. Ma, *Expert Syst. Appl.* **116**, 255 (2019)
- [58] C. M. Bishop, *Pattern Recognition and Machine Learning*, (New York: Springer, 2006), p. 227
- [59] D. P. Kingma and J. Ba, (2014), arXiv: 1412.6980 [cs.LG]
- [60] R. F. Casten, D. S. Brenner, and P. E. Haustein, *Phys. Rev. Lett.* **58**, 658 (1987)
- [61] Z. He, M. Bao, Y. M. Zhao *et al.*, *Phys. Rev. C* **87**, 057304 (2013)
- [62] R. Silwal, C. Andreoiu, B. Ashrafkhani *et al.*, *Phys. Lett. B* **833**, 137288 (2022)
- [63] W. S. Porter, B. Ashrafkhani, J. Bergmann *et al.*, *Phys. Rev. C* **105**, L041301 (2022)
- [64] S. Giraud, L. Canete, B. Bastin *et al.*, *Phys. Lett. B* **833**, 137309 (2022)
- [65] S. F. Paul, J. Bergmann, J. D. Cardona *et al.*, *Phys. Rev. C* **104**, 065803 (2021)
- [66] M. I. Mukul, C. Andreoiu, J. Bergmann *et al.*, *Phys. Rev. C* **103**, 044320 (2021)
- [67] M. Hukkanen, W. Ryssens, P. Ascher *et al.*, *Phys. Rev. C* **107**, 014306 (2023)
- [68] M. Mougeot, D. Atanasov, J. Kartheim *et al.*, *Nat. Phys.* **17**, 1099 (2021)
- [69] C. Izzo, J. Bergmann, K. A. Dietrich *et al.*, *Phys. Rev. C* **103**, 025811 (2021)
- [70] R. Orford, N. Vassh, J. A. Clark *et al.*, *Phys. Rev. C* **105**, L052802 (2022)
- [71] S. Beck, B. Kootte, I. Dedes *et al.*, *Phys. Rev. Lett.* **127**, 112501 (2021)
- [72] O. Kaleja, B. Andelić, O. Bezrodnova *et al.*, *Phys. Rev. C* **106**, 054325 (2022)
- [73] H. B. Yang, Z. G. Gan, Z. Y. Zhang *et al.*, *Phys. Rev. C* **105**, L051302 (2022)
- [74] R. P. de Groote, J. Billowes, C. L. Binnersley *et al.*, *Nat. Phys.* **16**, 620 (2020)
- [75] W. Satula, J. Dobaczewski, and W. Nazarewicz, *Phys. Rev. Lett.* **81**, 3599 (1998)
- [76] J. Dobaczewski, P. Magierski, W. Nazarewicz *et al.*, *Phys. Rev. C* **63**, 024308 (2001)
- [77] <http://www.nndc.bnl.gov/ensdf/>
- [78] M. Bao, Z. He, Y. M. Zhao *et al.*, *Phys. Rev. C* **90**, 024314 (2014)