





Data-driven density functional model for atomic nucleiZu-Xing Yang ^{1,2}, Xiao-Hua Fan ^{2,3,1}, Zhi-Pan Li ², and Haozhao Liang ^{3,4}¹*RIKEN Nishina Center, Wako, Saitama 351-0198, Japan*²*School of Physical Science and Technology, Southwest University, Chongqing 400715, China*³*Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan*⁴*RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program, Wako 351-0198, Japan*

(Received 14 March 2024; accepted 23 May 2024; published 10 June 2024)

Through ensemble learning with multitasking and complex connection neural networks, we aggregated nuclear properties, including ground-state density distributions, charge radii, and binding energies obtained from the Kohn-Sham auxiliary single-particle systems. The root-mean-squared (RMS) error in describing binding energies is reduced to about 500 keV, while the RMS error in describing charge radii is decreased to about 0.017 fm. In addition, by using the correlation between densities and binding energies, we discuss the impact of various densities and nuclear radii on the binding energy and find the neutron-skin thickness of ^{208}Pb to be around 0.220 fm. The present paper provides a new way to accelerate the integration of machine learning into nuclear density functional theory.

DOI: [10.1103/PhysRevC.109.064312](https://doi.org/10.1103/PhysRevC.109.064312)**I. INTRODUCTION**

Nuclear mass is a fundamental property for extracting various nuclear structure information, including nuclear pairing correlation, shell effect, deformation transition, and nuclear interactions [1,2]. In astrophysics research, nuclear mass also plays a crucial role in determining the composition on the surface of neutron stars [3], the astrophysical processes of nucleosynthesis, and the origin of elements in the Universe [4].

Considering the profound impact of mass in nuclear physics, a substantial amount of research has been devoted to enhancing the description and predictive accuracy. Traditional theoretical models, such as the Bethe-Weizsäcker mass formula [5], the finite-range droplet model [6], and the Weizsäcker-Skyrme model [7], as well as the Hartree-Fock-Bogoliubov mass model [8–10] and the relativistic mean-field model [11–15], typically exhibit an accuracy range between 0.3 and 3 MeV. Currently, the machine learning based research is gradually becoming one of the main forces on the path of achieving higher accuracy. Utama *et al.* [3] introduced the application of Bayesian neural networks to the residuals between theoretical and experimental data, achieving remarkable success with an improvement in mass accuracy of approximately 40%. The accuracy further reached an impressive 84 keV [16], through the incorporation of nuclear pairing and shell effects [17] together with meticulous design for multiple networks [16]. At the same time, machine learning approaches, such as the radial basis function [18,19], kernel ridge regression [20,21], the support vector machine [22], the Gaussian process [23–27], the decision tree [28,29], and others, have also been employed to describe the nuclear masses. From successful cases, another key insight is ensemble learning, which involves integrating multiple learning models to make

posterior predictions, also known as Bayesian model averaging [25–27,30] or world averaging [3].

With the description accuracy approaching the limits, the research emphasis should be reverted to the fundamental connections among observables for a deeper understanding of the physics behind the phenomena, for example, in the scheme of nuclear density functional theory (DFT). In a previous work [31], the Kohn-Sham network (KSN) was proposed to calibrate the nuclear charge radii based on the single-particle wave functions obtained by the Skyrme-Hartree-Fock theory with Bardeen-Cooper-Schrieffer (SHF + BCS) correlations. While enhancing the ability to describe the nuclear radius, it has to some extent lost the description of the binding energy. Therefore, it is necessary to reconnect the binding energy.

In this paper, based on the previous findings, we will establish a neural network mapping from nuclear spatial density, kinetic density, and spin-orbit density to nuclear binding energy. Different datasets will be employed to explore the network performance, where inputs include densities from SHF + BCS and those calibrated by KSN, and outputs comprise the binding energies from SHF + BCS and experimental data. The impact of the loss function on the training results will also be discussed. Furthermore, according to the correlation between densities and binding energies, the effects of various densities on binding energy, neutron skin thickness, and some relevant issues will also be investigated.

II. NEURAL NETWORK ARCHITECTURE

To aim at one of the most realistic mapping relationships, we focus on two main aspects for generating the inputs.

On the one hand, we derive features that empirically encompass known physical information based on proton number Z and neutron number N , including valence proton number

Z_v , valence neutron number N_v , proton hole number Z_h , neutron hole number N_h , proton shell number Z_s , neutron shell number N_s , shell effect parameter S , proton number parity Z_p , neutron number parity N_p , and parity parameter P . Taking the proton as an example, the relations among Z , Z_s , Z_v , Z_h , and Z_p satisfy

$$\begin{aligned}
 Z &= \mathcal{M}(Z_s) + Z_v, \\
 Z_h &= \mathcal{M}(Z_s + 1) - Z_v, \\
 Z_p &= Z \pmod{2}
 \end{aligned} \quad (1)$$

with the magic number list \mathcal{M} being $\{8, 20, 28, 50, 82, 126, 184\}$. The shell effect parameter S and the parity parameter P are defined as [17]

$$S = d_p d_n / (d_p + d_n) \quad \text{and} \quad P = [(-1)^Z + (-1)^N] / 2 \quad (2)$$

with d_p (d_n) representing the difference between the actual proton (neutron) number Z (N) and the nearest magic number. The Z , N , Z_v , N_v , Z_h , and N_h share the same dimension; we denote them as X_1 for uniform normalization in the neural network, while Z_s , N_s , S , Z_p , N_p , and P would be a supplement organized as X_2 . In addition to the clearly defined shell effects and the odd-even staggering, this part of the network also implicitly supplements some beyond-mean-field physics that is challenging to describe within the Kohn-Sham framework, such as nucleon correlations forming on the nuclear surface, including the α cluster [32,33].

On the other hand, we obtain several crucial densities from KSN single-particle wave functions φ_i and occupancy weights w_i calibrated by experimental charge radii. They include the nuclear spatial density

$$\rho = \sum_i \frac{d_i (\sqrt{w_i} \varphi_i)^2}{4\pi}, \quad (3)$$

kinetic density

$$\tau = \sum_i \frac{d_i}{4\pi} \left[(\partial_r \sqrt{w_i} \varphi_i)^2 + \frac{l_i(l_i + 1)}{r^2} (\sqrt{w_i} \varphi_i)^2 \right], \quad (4)$$

and spin-orbit density

$$J = \sum_i \frac{d_i}{4\pi} \left[j_i(j_i + 1) - l_i(l_i + 1) - \frac{3}{4} \right] \frac{2}{r} (\sqrt{w_i} \varphi_i)^2. \quad (5)$$

Here $i \in \{1s_{1/2}, 1p_{3/2}, 1p_{1/2}, \dots\}$ indicates the single-particle states, while d_i , l_i , and j_i respectively represent the degeneracy, the orbital angular momentum, and the total angular momentum of a state i . In the Skyrme-Hartree-Fock theory, the three aforementioned densities determine the kinetic, potential, and spin-orbit terms of nuclear interactions, thereby determining the nuclear binding energy via the Kohn-Sham equations.

In practice, when employing charge radii to calibrate densities [31], 640 nuclei data with $Z > 40$ were utilized, encompassing the majority of deformed nuclei. In this sense, all densities should be considered as angularly averaged. This also implies that the deformation-induced changes in binding energy are also characterized by the features X_1 and X_2 .

The mapping network from the above inputs to the nuclear binding energy is referred to as a density-to-energy network

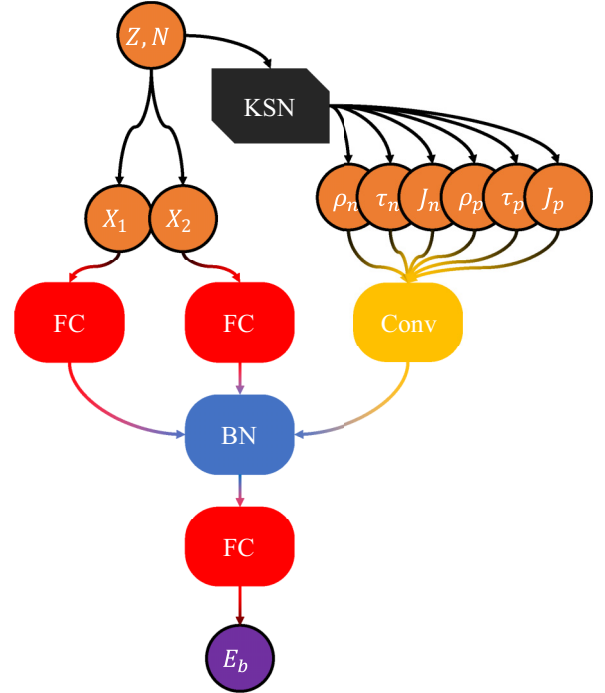


FIG. 1. Schematic diagram of the structure the of density-to-energy network. See the text for details.

(DTEN), the structure of which is shown in Fig. 1. The X_1 and X_2 are input into two separate four-layer fully connected (FC) neural network cells (C_1 and C_2), while the six densities (ρ_n , τ_n , J_n , ρ_p , τ_p , J_p) as continuous variables are fed into a five-layer convolutional (Conv) neural network cell with six channels (C_3). Specifically, a max-pooling layer is connected after each convolutional layer to reduce the parameter and expedite convergence. Subsequently, the outputs from these branch cells are concatenated and uniformly batch normalized (BN) to align the feature distributions with a normal distribution. Afterward, passing through another FC cell (C_4), the features are finally mapped to the binding energy E_b . An essential point that must be emphasized is that the network consists of 23 layers with complex connections, whose complexity can lead to some neurons being trapped in the negative range and becoming deactivated under the commonly used ReLU ($=\max\{0, x\}$) activation function [34] for nonlinearity, especially after multiple iterations. To address this issue and improve convergence, we adopt the LReLU ($=\max\{0.01x, x\}$) activation function [35] to avoid gradient vanishing. Furthermore, to mitigate the influence of absolute data magnitudes, all input features are subject to min-max scaling, ensuring they fall within the range of 0 to 1. Due to the constrained range of values, we employed the Sigmoid ($=1/(1 + e^{-x})$) activation function in the final layer. The meticulously designed hyperparameter set for the DTEN architecture is listed in Table I. In the table, the initial row and the concluding row of each cell respectively represent its input and output, with the symbol “ \cup ” meaning concatenating two vectors, i.e., $[a, b, \dots] \cup [c, d, \dots] = [a, b, \dots, c, d, \dots]$. The network contains a total of 1 173 281 trainable parameters. During the training process, a dynamically decreasing

TABLE I. The hyperparameter set of the DTEN structure. The “ D ” represents the output dimension of the layer, the “ C_{in} ” the input channels, the “ C_{out} ” the output channels, the “ K_s ” the size of the kernel which includes both convolutional and pooling dimensions, the “ S_r ” the stride used during convolution or pooling operations, and the “ $g(x)$ ” the nonlinear activation function.

C_1 or C_2								
L	Type	D					$g(x)$	
	X_1 or X_2						6	
1	FC						32	LReLU
2	FC						64	LReLU
3	FC						128	LReLU
4	FC						256	LReLU
	Output ₁ or output ₂						256	
C_3								
L	Type	D	C_{in}	C_{out}	K_s	S_r	$g(x)$	
	Densities	(6,150)						
1	Conv.	(32,150)	6	32	3	1	LReLU	
2	Pooling	(32,75)	32	32	2	2		
3	Conv.	(64,75)	32	64	3	1	LReLU	
4	Pooling	(64,25)	64	64	3	3		
5	Conv.	(128,25)	64	128	3	1	LReLU	
6	Pooling	(128,8)	128	128	3	3		
7	Conv.	(256,8)	128	256	3	1	LReLU	
8	Pooling	(256,2)	256	256	4	4		
9	Conv.	(512,1)	256	512	2	1	LReLU	
	Output ₃	512						
C_4								
L	Type	D					$g(x)$	
	Output ₁ \cup output ₂ \cup output ₃						1024	
1	BN						1024	
2	FC						512	LReLU
3	FC						256	LReLU
4	FC						128	LReLU
5	FC						32	LReLU
6	FC						1	Sigmoid
	Binding energy						1	

learning rate, which reduces as the loss function converges, is applied with the Adaptive Momentum Estimation (Adam) optimizer [36]. The aforementioned definitions and concepts are entirely consistent with the patterns used in PYTORCH [37].

In the current design, the non-model-dependent inputs (X_1 and X_2) and the model-dependent inputs (Densities) are blended. The advantage is to retain accuracy as much as possible while exploring the impact of densities on binding energy. In this manner, the search for a parametrized energy density functional is replaced by neural network mapping, but the insights from traditional functionals are not aborted, which will produce profound significance for describing complex nuclear systems.

III. RESULTS AND ANALYSIS

In this paper, the proton single-particle wave functions generated by KSN have undergone the calibrations with the

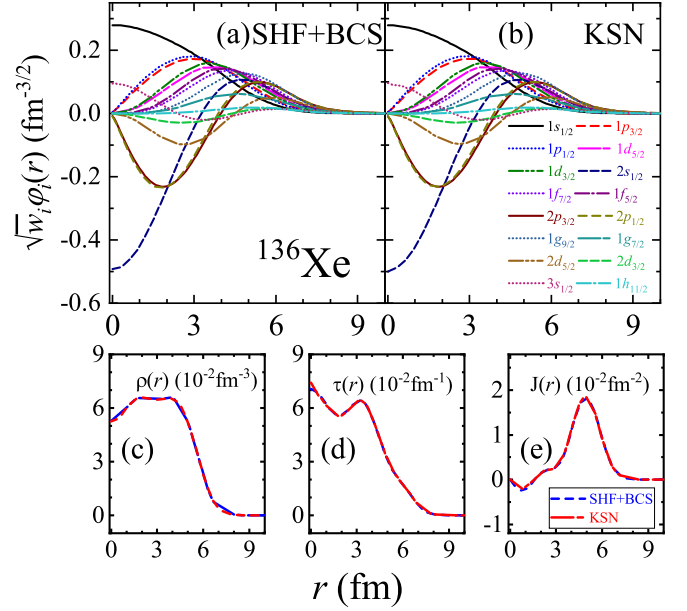


FIG. 2. The proton $\sqrt{w_i}\phi_i(r)$ in coordinate space for each single-particle orbital in the nucleus ^{136}Xe with (a) SHF + BCS and (b) calibrated KSN, and the comparisons for the corresponding (c) spatial densities, (d) kinetic densities, and (e) spin-orbit densities.

experimental charge radii of over 600 nuclei. However, due to the lack of neutron information from laboratories, the original SHF + BCS neutron densities with SkM* interaction [38] are still employed. To ensure that the variation in density is within physically permissible limits, we need to examine the calibrated proton densities.

The comparisons in Fig. 2 depict $\sqrt{w_i}\phi_i(r)$ in coordinate space for each single-particle orbital, spatial densities, kinetic densities, and spin-orbit densities between SHF + BCS and calibrated KSN. It is evident that there are only minimal changes for the various densities. This is understandable: during the calibration process of KSN [31], the original $\sqrt{w_i}\phi_i(r)$ from SHF + BCS still retains a certain weight, and the calibration for charge radii often only minimally changes on the typical order of a few 10^{-2} fm. This implies that the loss of self-consistency in the Kohn-Sham equation is limited at a minimal level, hence it is feasible to roughly explore changes in the kinetic, potential, and spin-orbit terms of nuclear interactions on the basis of the SkM* parameters. However, as illustrated in Ref. [39], a neural network trained with density and binding energy data from Skyrme DFT was applied to discuss Ca isotopes, encountering a failure in describing ^{48}Ca with an overbinding phenomenon. Such failure can be attributed to an indispensable beyond-mean-field effect appearing near ^{48}Ca [40,41]. To capture the descriptive capacity of the beyond-mean-field effect, it is a natural approach to establish a mapping relation that targets experimental data. Turning to the training processes of DTENs, the changes in the loss value as a function of training epochs are shown in Fig. 3(a). In approximately 2400 nuclei with proton numbers greater than 40, for which binding energies have been measured with high precision, we utilize an 8:2 ratio for dividing the data into the training and validation sets. For each epoch,

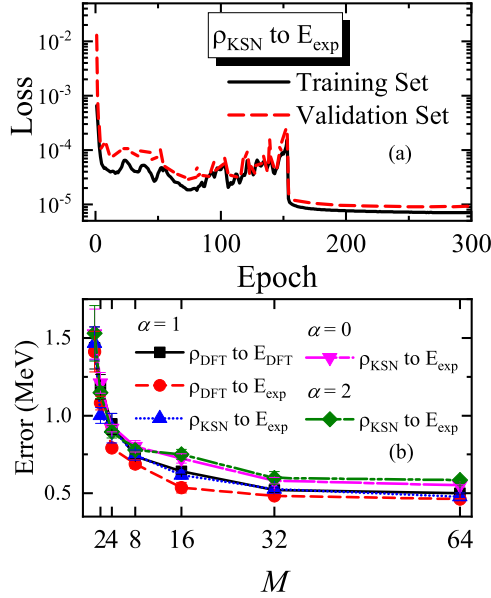


FIG. 3. (a) Loss values on the training set and validation set as a function of training epochs. (b) The root-mean-squared errors in integrating M networks with different datasets and loss functions. See text for the legends.

all nuclei in the training set are trained once, and then the loss value on the validation set (L_v) is calculated and recorded. If a certain L_v is lower than all previously records, then the model parameters will be saved and overwrite the previous ones. After experiencing 50 epochs without model parameters being overwritten, the model will reload the most recently saved parameters, and the learning rate will be reduced by a factor of 10. When the learning rate falls below 10^{-8} , the training stops.

Here, the loss function is constructed as a weighted mean squared error

$$\text{Loss} = C \times \langle (E_{b,\text{pre}} - E_{b,\text{tar}})^2 \times A^\alpha \rangle, \quad (6)$$

where $E_{b,\text{pre}}$ and $E_{b,\text{tar}}$ represent the predicted and experimental values of the binding energy, respectively. The constant $C = 2.78 \text{ MeV}^{-2}$ is caused by max-min scaling mentioned before. Weight A corresponds to the mass number with the exponent α emphasizing the importance of mass, where $\alpha = 0$ implies that the binding energies per nucleon are equally significant for all nuclei, while an α of 2 means that the masses of all nuclei are equally important. By setting $\alpha = 1$, a compromise is achieved between the binding energy per nucleon and mass to some degree. In Fig. 3(a), α is set to 1. It can be observed that the training set and validation set have essentially converged after 150 epochs, indicating basically no overfitting occurs and emphasizing the generalization capability.

It can be observed that the loss value has converged to a very low level. Further examination reveals that the root-mean-squared (RMS) error of binding energy corresponding to this loss is 1.6 MeV. Based on this, to achieve higher accuracy, we further employ ensemble learning, combining multiple DTENs with identical structures for a final

prediction. For each DTEN, the model initialization is different, which leads to variations in the convergence paths during the model training process using stochastic gradient descent, resulting in different predictive capabilities. The predictive capability of a DTEN can be described by the RMS error σ , serving as a prior according to the Bayesian principle. Therefore, the final prediction can be expressed as

$$E_{b,\text{final}} = \sum_i^M w_i \times E_{b,\text{pre},i} \quad (7)$$

with

$$w_i = \frac{1/\sigma_i^2}{\sum_i^M 1/\sigma_i^2}, \quad (8)$$

where M is the total number of models in the ensemble. The blue scatter line for the mapping from the KSN-calibrated densities to the experimental masses (labeled as “ ρ_{KSN} to E_{exp} ”) in Fig. 3(b) describes the corresponding variation of the final σ concerning M . It is clear that with the increase in the number of models, the RMS error is substantially reduced, ultimately converging to about 500 keV. There is almost no improvement with $M = 64$ compared with $M = 32$; therefore, further increasing the number of models in the ensemble is not essential. The final precision, compared with the adopted SHF + BCS theory [42] with the RMS error being about 10 MeV, has been improved by orders of magnitude.

The comparison of different α values for ρ_{KSN} to E_{exp} is shown in Fig. 3(b). It is found that the case of $\alpha = 1$ leads to the smallest error, compared with the cases of $\alpha = 0$ (pink curve) and $\alpha = 2$ (green curve). This might be because $\alpha = 0$ leads the network to focus on lighter nuclei, while $\alpha = 2$ causes the network to pay more attention to heavier nuclei, resulting in a partial loss of global perspective in both cases.

According to the past experience, the learning capability of a network often depends on the dataset. We further discuss this issue on the basis of $\alpha = 1$. Two additional datasets are conducted: one mapping from the DFT densities to the corresponding DFT energies (labeled as “ ρ_{DFT} to E_{DFT} ”), and the other mapping from the DFT densities to the experimental energies [43] (labeled as “ ρ_{DFT} to E_{exp} ”). The corresponding comparisons are presented in Fig. 3(b). Upon comparing the training results of three datasets, we observe that under the final ensemble ($M = 64$), the loss values of all three are very similar, with differences of only a few tens of keV. Although the ρ_{DFT} to E_{exp} case seems to exhibit superior accuracy in describing experimental values, the difference is not significant. Meanwhile, to balance the accurate description of radii, we still consider the ρ_{KSN} to E_{exp} case to be more reasonable.

In principle, the data derived from laboratory measurements should add greater complexity to the dataset. However, from the above discussion, we find that the descriptive capability of the present network is consistent for the descriptions of both the computed data and the experimental data. This supports the notion that neural networks based on the simulated data are also effective in the description of the real world.

Based on the strong correlation between densities and binding energies, inferring the nuclear neutron-skin thickness is an interesting and crucial problem. The KSN-calibrated

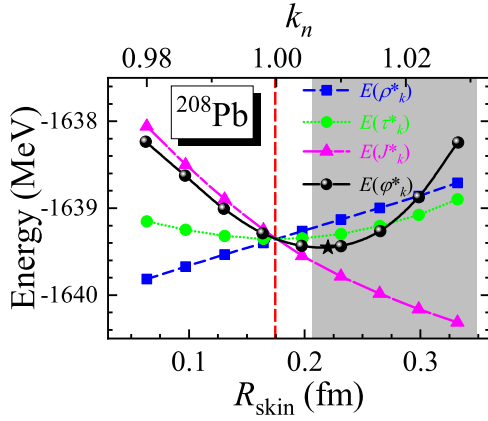


FIG. 4. The energy of ^{208}Pb as a function of compression ratio k_n , where $E(\rho^*)$, $E(\tau^*)$, and $E(J^*)$ denote the change in energy due to the compressed space density, kinetic density, and spin-orbit density, respectively. The $E(\varphi^*)$ represents the effect caused by the compressed wave function, which compresses all three densities simultaneously. The corresponding neutron-skin thickness for ρ^* and φ^* is labeled with the lower x axis. The red dashed line represents the calculated values on the training set, while the shadow indicates the experimental value [44].

proton densities have reflected the true information about proton radii. By observing the correlation between neutron density variations and the binding states of a nucleus, we can further infer the neutron-skin thickness. To this end, ^{208}Pb , as a prominent spherical nucleus, is taken for examination. Physically, the variation in spatial density will also lead to the changes in kinetic density and spin-orbit density. To maintain the self-consistency, based on the Kohn-Sham auxiliary single-particle system, we apply a compression operator to the neutron single-particle wave function, i.e.,

$$\varphi_n(r) \rightarrow C_{k_n} \varphi_n(k_n r), \quad (9)$$

where the parameter k_n controls the compression ratio, while C_{k_n} maintains the normalization. Simultaneously, to qualitatively understand the impact of each density on binding energy, they are also individually compressed for comparison.

With the variation of k_n , the corresponding changes occur in the binding energy, as shown in Fig. 4. At a neutron-skin thickness of 0.220 fm for the compressed wave function φ^* , the nuclear binding strength is maximized. The obtained value is greater than the initial training value (red dashed line) and coincidentally falls on the edge of the experimental measurement range [44]. Another noteworthy point is that near the minimum point, the energy changes relatively softly, less than 100 keV, suggesting that the neutron-skin thickness is not so sensitive to its binding energy. Furthermore, when compressing each density individually, it can be observed that the kinetic density τ has only a relatively minor effect, whereas the spin-orbit term J and the matter density term ρ have strong contributions to the binding energy. In this regard, artificial intelligence recognizes that with the compression of the nucleus, the matter density contributes more energy, whereas the energy from the spin-orbit density diminishes, a phenomenon aligned with the logic of physics.

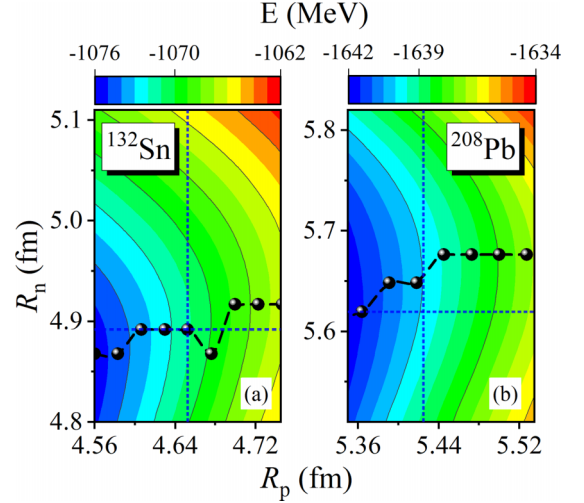


FIG. 5. The binding energies of (a) ^{132}Sn and (b) ^{208}Pb as 2D functions of proton and neutron radii. The blue dashed line represents $k_{n(p)} = 1$. The black spherical dotted line depicts the variation of the binding energy peak for each R_p .

The impact of variations in the proton radius on the system is illustrated in Fig. 5, where the minimum energy neutron radius has been searched under different proton radii. According to the black spherical dotted line, which depicts the variation of the binding energy peak with changes in the proton radius, it is noticeable that an increase in the proton radius necessitates an increase in the neutron radius. However, the increase in the neutron radius is always less than that of the proton radius ($\Delta R_n < \Delta R_p$) from the ^{132}Sn and ^{208}Pb , implying that the increasing proton radius will reduce the skin thickness. It is worth emphasizing that the current conclusions are quite stable. The energy surfaces under other conditions presented in Fig. 3(b) have also been examined, yielding remarkably similar results.

Finally, as supplementary information, the RMS deviation of the charge radii of 630 nuclei with $Z > 40$ from experimental data [45] is further compared among the SHF + BCS calculation, neural network calibration (Ref. [31]), and the ensemble prediction results using 50 calibrated neural networks (labeled as ‘‘This paper’’), which is shown in Table II. By comparing the SHF + BCS calculation with the calibration of a single network [31], it can be noted that the single network has significantly improved the description of charge radii and has limited overfitting (the RMS deviation in the training set is 0.0167 fm and in the validation set is 0.0172 fm). Unlike the prediction of binding energies, further ensemble learning for charge radii is highly inefficient. This may be due to the composite loss function based on both theory and experiment

TABLE II. The root-mean-squared deviation (fm) of the charge radii of 630 nuclei with $Z > 40$ from experimental data.

SHF+BCS	Ref. [31]	This paper
0.0456	0.0168	0.0163

during calibration having already sufficiently constrained the charge radii.

The current research, in conjunction with Ref. [31], has essentially completed the construction of neural networks based on the Kohn-Sham scheme for enhancing DFT. Regarding observables, it achieves high accuracy in describing both binding energies and nuclear radii. In terms of physical details, it can also be employed to explore the contributions of single-particle states and shell structure, as well as the impact of various densities on binding energy.

Nevertheless, the current model still deserves further refinement and optimization. Primarily, when considering the nonmagic nuclei, the valence-nucleon-induced deformation effects do not directly manifest in the densities. The present densities are assumed to be angularly averaged, which may impact the accuracy of describing binding energies and further compromise the capability to characterize multipole deformation potential surfaces. Therefore, three-dimensionalizing the current model is imperative. Secondly, the current model still falls short of providing a self-consistent description of the system. Embedding the present network into the Kohn-Sham equations represents a significantly important direction. Finally, the critical role played by spin-orbit density suggests that a relativistic framework may be more effective in describing the neutron skin.

IV. SUMMARY

With the aid of the nuclear single-particle wave functions generated by the experimental charge radius calibrated Kohn-Sham network, we computed three essential densities in DFT, i.e., spatial density, kinetic density, and spin-orbital density. Through an elaborated neural network, the densities are further mapped to the experimental binding energies. By employing a weighted ensemble of multiple models, the RMS error in describing binding energies is decreased to 455 keV, while the RMS error in describing charge radii is decreased

to 0.0163 fm. There has been a noticeable improvement compared to the initial calculations based on SHF + BCS.

The relation between binding energy and neutron-skin thickness is further explored. Considering the self-consistency among densities, a contraction operator is applied to the neutron single-particle wave functions to establish the correlation between neutron skin and binding energy. By searching for the minimum point, the estimated neutron-skin thickness of ^{208}Pb is obtained as approximately 0.220 fm. When discussing the impacts of various densities independently, artificial intelligence identifies that both spatial and spin-orbit densities play significant roles. Further investigation of the energy variations in the R_n - R_p two-dimensional plane indicates that changes in the proton radius are more sensitive than those in the neutron radius. We look forward to further experimental validation of these conclusions.

This paper aggregates the charge radius data over 600 nuclei, the binding energy data for more than 2400 nuclei, and the single-particle state data based on DFT. Ultimately, it bypasses the many-body interaction potential and establishes the correlations among observables, whose descriptive performance for the nuclear ground state has surpassed that of the majority of existing density functional models. In the future, by three-dimensionalizing the model, incorporating adversarial neural networks, as well as introducing more experimental data, the neural network for enhancing DFT will possess stronger descriptive capabilities.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under Grants No. 12005175 and No. 12375126 and a Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (S) under Grant No. 20H05648. This work is also partially supported by the RIKEN Pioneering Project: Evolution of Matter in the Universe.

-
- [1] D. Lunney, J. M. Pearson, and C. Thibault, *Rev. Mod. Phys.* **75**, 1021 (2003).
 - [2] M. Bender, P.-H. Heenen, and P.-G. Reinhard, *Rev. Mod. Phys.* **75**, 121 (2003).
 - [3] R. Utama, J. Piekarewicz, and H. B. Prosper, *Phys. Rev. C* **93**, 014311 (2016).
 - [4] E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle, *Rev. Mod. Phys.* **29**, 547 (1957).
 - [5] H. A. Bethe and R. F. Bacher, *Rev. Mod. Phys.* **8**, 82 (1936).
 - [6] P. Möller, W. D. Myers, H. Sagawa, and S. Yoshida, *Phys. Rev. Lett.* **108**, 052501 (2012).
 - [7] N. Wang, M. Liu, X. Wu, and J. Meng, *Phys. Lett. B* **734**, 215 (2014).
 - [8] S. Goriely, N. Chamel, and J. M. Pearson, *Phys. Rev. Lett.* **102**, 152503 (2009).
 - [9] S. Goriely, S. Hilaire, M. Girod, and S. Péru, *Phys. Rev. Lett.* **102**, 242501 (2009).
 - [10] S. Goriely, N. Chamel, and J. M. Pearson, *Phys. Rev. C* **93**, 034337 (2016).
 - [11] K. Zhang, M.-K. Cheoun, Y.-B. Choi, P. S. Chong, J. Dong, Z. Dong, X. Du, L. Geng, E. Ha, X.-T. He *et al.* (DRHBc Mass Table Collaboration), *At. Data Nucl. Data Tables* **144**, 101488 (2022).
 - [12] Y. L. Yang, Y. K. Wang, P. W. Zhao, and Z. P. Li, *Phys. Rev. C* **104**, 054312 (2021).
 - [13] S. E. Agbemava, A. V. Afanasjev, D. Ray, and P. Ring, *Phys. Rev. C* **89**, 054320 (2014).
 - [14] A. Taninah and A. V. Afanasjev, *Phys. Rev. C* **107**, L041301 (2023).
 - [15] H. Liang, N. Van Giai, and J. Meng, *Phys. Rev. Lett.* **101**, 122502 (2008).
 - [16] Z. M. Niu and H. Z. Liang, *Phys. Rev. C* **106**, L021303 (2022).
 - [17] Z. Niu and H. Liang, *Phys. Lett. B* **778**, 48 (2018).
 - [18] Z. M. Niu, Z. L. Zhu, Y. F. Niu, B. H. Sun, T. H. Heng, and J. Y. Guo, *Phys. Rev. C* **88**, 024325 (2013).
 - [19] Z. M. Niu, J. Y. Fang, and Y. F. Niu, *Phys. Rev. C* **100**, 054311 (2019).
 - [20] X. H. Wu and P. W. Zhao, *Phys. Rev. C* **101**, 051301(R) (2020).

- [21] X. Wu, L. Guo, and P. Zhao, *Phys. Lett. B* **819**, 136387 (2021).
- [22] J. W. Clark and H. Li, *Int. J. Mod. Phys. B* **20**, 5015 (2006).
- [23] A. Pastore, D. Neill, H. Powell, K. Medler, and C. Barton, *Phys. Rev. C* **101**, 035804 (2020).
- [24] M. Shelley and A. Pastore, *Universe* **7**, 131 (2021).
- [25] L. Neufcourt, Y. Cao, W. Nazarewicz, E. Olsen, and F. Viens, *Phys. Rev. Lett.* **122**, 062502 (2019).
- [26] L. Neufcourt, Y. Cao, S. A. Giuliani, W. Nazarewicz, E. Olsen, and O. B. Tarasov, *Phys. Rev. C* **101**, 044307 (2020).
- [27] L. Neufcourt, Y. Cao, S. Giuliani, W. Nazarewicz, E. Olsen, and O. B. Tarasov, *Phys. Rev. C* **101**, 014319 (2020).
- [28] M. Carnini and A. Pastore, *J. Phys. G* **47**, 082001 (2020).
- [29] Z.-P. Gao, Y.-J. Wang, H.-L. Lü, Q.-F. Li, C.-W. Shen, and L. Liu, *Nucl. Sci. Tech.* **32**, 118 (2021).
- [30] A. Hamaker, E. Leistenschneider, R. Jain, G. Bollen, S. A. Giuliani, K. Lund, W. Nazarewicz, L. Neufcourt, C. R. Nicoloff, D. Puentes, R. Ringle, C. S. Sumithrarachchi, and I. T. Yandow, *Nat. Phys.* **17**, 1408 (2021).
- [31] Z.-X. Yang, X.-H. Fan, Z.-P. Li, and H. Liang, *Phys. Lett. B* **840**, 137870 (2023).
- [32] J. Tanaka, Z. Yang, S. Typel, S. Adachi, S. Bai, P. van Beek, D. Beaumel, Y. Fujikawa, J. Han, S. Heil, S. Huang, A. Inoue, Y. Jiang, M. Knösel, N. Kobayashi, Y. Kubota, W. Liu, J. Lou, Y. Maeda, Y. Matsuda *et al.*, *Science* **371**, 260 (2021).
- [33] S. Typel, *Phys. Rev. C* **89**, 064321 (2014).
- [34] M. J. Villani and N. Schoots, [arXiv:2306.11827](https://arxiv.org/abs/2306.11827) (2023).
- [35] B. Xu, N. Wang, T. Chen, and M. Li, [arXiv:1505.00853](https://arxiv.org/abs/1505.00853) (2015).
- [36] D. P. Kingma and J. Ba, in *Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA*, edited by Y. Bengio and Y. LeCun (ICLR, 2015).
- [37] PyTorch documentation, <https://pytorch.org/docs/stable/index.html> (accessed on 2023-08-03).
- [38] J. Bartel, P. Quentin, M. Brack, C. Guet, and H.-B. Håkansson, *Nucl. Phys. A* **386**, 79 (1982).
- [39] Z.-X. Yang, X.-H. Fan, T. Naito, Z.-M. Niu, Z.-P. Li, and H. Liang, *Phys. Rev. C* **108**, 034315 (2023).
- [40] U. C. Perera, A. V. Afanasjev, and P. Ring, *Phys. Rev. C* **104**, 064313 (2021).
- [41] T. Naito, T. Oishi, H. Sagawa, and Z. Wang, *Phys. Rev. C* **107**, 054307 (2023).
- [42] P.-G. Reinhard, in *Computational Nuclear Physics* (Springer-Verlag, Berlin, 1991), Vol. 1, pp. 28–50.
- [43] M. Wang, W. Huang, F. Kondev, G. Audi, and S. Naimi, *Chin. Phys. C* **45**, 030003 (2021).
- [44] D. Adhikari, H. Albatineh, D. Androic, K. Aniol, D. Armstrong *et al.*, *Phys. Rev. Lett.* **126**, 172502 (2021).
- [45] I. Angeli and K. Marinova, *At. Data Nucl. Data Tables* **99**, 69 (2013).