Calibration of nuclear charge density distribution by back-propagation neural networks

Zu-Xing Yang (杨祖星),[1] Xiao-Hua Fan (范小华),[2] Tomoya Naito (内藤 智也),[3,4] Zhong-Ming Niu (牛中明),[5] Zhi-Pan Li (李志攀),[2] and Haozhao Liang (梁豪兆)[4,3]

[1]*RIKEN Nishina Center, Wako 351-0198, Japan*

[2]*School of Physical Science and Technology, Southwest University, Chongqing 400715, China*

[3]*RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program, Wako 351-0198, Japan*

[4]*Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan*

[5]*School of Physics and Optoelectronic Engineering, Anhui University, Hefei 230601, China*

(Received 10 August 2022; revised 25 June 2023; accepted 13 July 2023; published 29 September 2023)

Based on the back-propagation neural networks and density functional theory, a supervised learning is performed firstly to generate the nuclear charge density distributions. The charge density is further calibrated to the experimental charge radii by a composite loss function. It is found that, when the parity, pairing, and shell effects are taken into account, about 96% of the nuclei in the validation set fall within 2 standard deviations of the predicted charge radii. Moreover, the kink in charge radii on Hg isotopes has been successfully reproduced. The calibrated charge density is then mapped to the matter density and further mapped to the binding energies according to the Hohenberg-Kohn theorem. It provides an improved description of some nuclei in both binding energies and charge radii. Moreover, the anomalous overbinding in ^{48}Ca implies that the segmental calibrations by neural networks for beyond-mean-field effects deserve further discussion.

DOI: 10.1103/PhysRevC.108.034315

I. INTRODUCTION

The charge density distribution, experimentally given by electron scattering experiments [1–3], is essential for extracting nuclear structure information, including the shell-structure evolution, shape coexistence, shape transition, and neutron-skin thickness [4–7]. The relative radii of neighboring nuclei can also be determined using muonic-atom spectra as well as isotope shifts of laser spectroscopy [1,8]. Since a strong connection exists between charge densities and matter densities, one generally calculates the charge density distribution by merging the nucleon form factors [9,10] with the matter densities determined by scattering experiments, the $(-1p)$ reaction [11], or even heavy-ion collisions [12]. Theoretically, the Fermi distributions [13,14] and the Fourier-Bessel expansion [14,15] were used to describe nuclear matter and charge densities approximately in the early stages of nuclear physics development. With the development of computational power, various branching models [16–25] derived from density functional theory (DFT) and shell-model calculation have become more popular among theorists. However, due to the complexity of nuclear many-body systems, the calculations of these theories still face challenges in describing the beyond-mean-field effects and nucleon-nucleon correlations [26–29].

For nuclear complex systems, back-propagation neural networks [30] have achieved a series of successes in various aspects [31–45]. In early research, machine learning was employed as a universal approximator. Typically, the input and the output consist of low-dimensional experimental data, such as nuclear binding energy and charge radius [46,47], and its performance is largely indistinguishable from that of regular polynomial fitting. Taking into account the fundamental physical essence behind experimental data, the research emphasis has been shifted towards integrating theories for forecasting experimental outcomes. During this period, machine-learning research dominated by Bayesian neural networks began to learn and predict the residuals between theoretical and experimental values, which has an advantage that the corrected predictions are more accurate than the existing theoretical model [31,43,48]. As various theoretical models further mature, it became urgent to study the dependence of predictions on the model framework. As early as in Ref. [48], the variation in predicting nuclear masses with Bayesian neural networks based on several theoretical models was discussed. The model dependence of β-decay half-lives was also extensively discussed in Ref. [43]. In order to eliminate the dependencies, the "world average" scheme was initially adopted in Ref. [48], where the root-mean-square errors between experimental values and neural network predictions are taken to weight each model so as to obtain an average prediction. This later evolved into the Bayesian model averaging (BMA) [49,50], where the reliability of mass predictions is determined through separation energies under the Bayes principle. That is to say, the BMA allows a more targeted and physical determination of weights.

On the other hand, generative models have gained widespread acceptance across diverse fields, with their research focus transitioning from low-dimensional data to large-scale data such as images and texts. In accordance with the trend of popular pretraining generative models [51], and in order to further enhance the interpretability of network predictions and the correlation among different observables, we

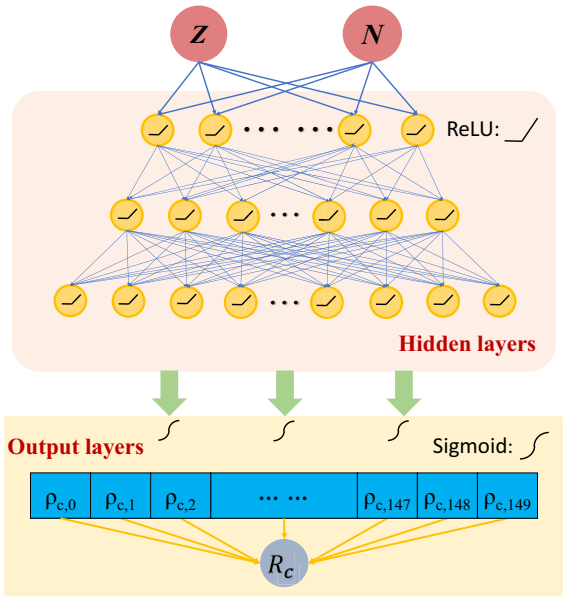

FIG. 1. Schematic diagram of the structure of the charge density generator neural networks.

embark on this research by incorporating density information into the prediction of charge radii. Based on the Hohenberg-Kohn maps [52] of DFT and the proven strong generalization ability of neural networks in describing density distributions [53], in this paper, we collectively constrain the charge density distributions by back-propagation with experimental data of charge radii, which makes the residual information flow back from radii to densities. We also construct the map from charge densities to binding energies to achieve further transfer of information from radii to binding energies.

II. THEORY

A. The charge density generators

Previously, a multilayer feed-forward neural network with a back-propagation algorithm of error has been elaborated to perform the maximum likelihood estimation in the process of generating density distributions [53] approximating the theoretical calculation. It is shown that a network trained by the density distributions of about 200–300 nuclei is sufficient to describe the density distributions of all the nuclei on the nuclear chart and has a powerful extrapolation capability [53].

Based on such a high computational efficiency and generalization ability, in this study, we introduce the correlation between the theoretical charge densities and the experimental root-mean-square charge radii (R_c) to a new hybrid neural network. The structure of this network is shown in Fig. 1, where $\rho_{c,i} = \rho_c(r_i)$ is the density on the mesh with $r_i = 0.1 \times i$ fm ($i = 0, 1, \ldots, 149$). The input is the proton and neutron numbers of a nuclide, $\mathbf{x} = \{Z, N\}$. The outputs are the charge density ρ_c and the root-mean-square charge radius R_c gained by the customized integration layer $\hat{l}_{\rho_c \to R_c}$,

$$\hat{l}_{\rho_c \to R_c} : R_{c,\mathrm{pre}} = \sqrt{\frac{\int \rho_{c,\mathrm{pre}}(r) r^4 dr}{\int \rho_{c,\mathrm{pre}}(r) r^2 dr}}, \tag{1}$$

where $\rho_{c,\mathrm{pre}}$ is the normalized predicted charge density distribution, i.e., $\int 4\pi \rho_{c,\mathrm{pre}}(r) r^2 dr = Z$. For the hidden module, three attempts are made: (i) a deep fully connected neural network, (ii) a convolutional neural network, suitable for learning the gradient information of features, and (iii) a neural network with a feature layer $\hat{l}_{i2 \to i8}$, containing the parity, pairing, and shell effects,

$$\hat{l}_{i2 \to i8} : \{Z, N\} \to \{Z, N, \mathcal{O}_p, \mathcal{O}_n, \delta, v_p, v_n, \mathcal{P}\}, \tag{2}$$

where

$$\begin{aligned} \mathcal{O}_p &= Z \mod 2, & \mathcal{O}_n &= N \mod 2, \\ \delta &= [(-1)^Z + (-1)^N]/2, & \mathcal{P} &= v_p v_n/(v_p + v_n), \end{aligned} \tag{3}$$

and v_p (v_n) is the difference between the actual proton (neutron) number Z (N) and the nearest magic number (8, 20, 28, 50, 82, or 126). Hereinafter, we refer to the three schemes of neural networks $\mathcal{M}$ as charge density generator-1 (CDG-1), CDG-2, and CDG-3, respectively. See the Supplemental Materials (SM) for details [54], where Refs. [53,55–58] are also referred. The map is denoted as

$$(\rho_c, R_c)_{\mathrm{pre}} = \mathcal{M}(\mathbf{x}, \mathbf{w}), \tag{4}$$

where $\mathbf{w}$ is the set of trainable parameters.

In general, the charge density distribution calculated by physical models $\rho_{c,\mathrm{theo}}$, e.g., that calculated by the Skyrme Hartree-Fock (SHF) theory with the Bardeen-Cooper-Schrieffer (BCS) pairing, is expected to be quite accurate. The R_c residuals between theory and experiment can be eliminated by assuming a correction to the charge density $\delta\rho_c$, satisfying $\int_0^\infty \delta\rho_c(r) r^2 dr = 0$. We refer to this process as calibration. The calibrated charge density distribution $\rho_{c,\mathrm{cali}}$, whose R_c is expected to be close to $R_{c,\mathrm{exp}}$ [1], can be obtained by

$$\rho_{c,\mathrm{cali}}(r) = \rho_{c,\mathrm{theo}}(r) + \delta\rho_c(r). \tag{5}$$

Additionally, we aim at *making the smallest possible corrections $\delta\rho_c$ to the theory.*

To this end, we design a composite loss function. The normalized mean-square-error L_ρ [53] is employed as an assessment of the density distribution:

$$L_\rho = \frac{1}{N_g} \sum_{i=0}^{149} [\rho_{c,\mathrm{pre}}(r_i) - \rho_{c,\mathrm{theo}}(r_i)]^2 \times 1\ \mathrm{fm}^6, \tag{6}$$

where $N_g = 150$ indicates the number of grid points and the factor 1 fm^6 makes L_ρ dimensionless. The charge density distributions $\rho_{c,\mathrm{theo}}(r)$ are calculated by the SHF + BCS theory with the SkM* energy density functional (EDF) [59]. The BCS correlation is treated by the commonly accepted constant-gap approach [58], which simply adopts the relationship that the pairing gap is inversely proportional to the square root of the number of nucleons. The $\rho_{c,\mathrm{theo}}(r)$ are obtained from the charge form factor F_c by the inverse Fourier-Bessel transform,

$$\rho_{c,\mathrm{theo}}(r) = \frac{1}{2\pi^2} \int dk\, k^2 j_0(kr) F_c(k; \rho_n, \rho_p, \ldots). \tag{7}$$

where j_0 is the spherical Bessel function. The contributions of the matter density and the spin-orbit current are folded in F_c [58] (see the SM [54] for details). Here, R_c has a large

range of variation (about 1–6 fm) and, therefore, the Pearson χ^2 divergence is picked; i.e., its loss function $L_{\rm R}$ reads

$$L_{\rm R} = \frac{(R_{\rm c,pre} - R_{\rm c,exp})^2}{R_{\rm c,pre}} \times 1\,{\rm fm}^{-1}, \tag{8}$$

where the factor 1 fm^{-1} also makes $L_{\rm R}$ dimensionless. The two loss functions L_ρ and $L_{\rm R}$ are combined as

$$\mathrm{Loss}(\mathcal{W}, \mathbf{w}) \equiv \frac{1}{B_s}\sum_{{\rm n_u}=1}^{B_s}\left[(1-\mathcal{W})L_\rho^{({\rm n_u})}(\mathbf{w}) + \mathcal{W}L_{\rm R}^{({\rm n_u})}(\mathbf{w})\right], \tag{9}$$

where $\mathcal{W}$ is the weighting factor and $B_s = 64$ is the batch size, which means that 64 nuclei are constrained simultaneously for each training session. Actually, the corrections for different nuclei are derived from the same parameter updates $\delta\mathbf{w}$; i.e., *the correction of each nucleus is uniformly constrained by the other nuclei*.

In comparison to traditional methods, the corrections of the densities based on CDGs are no longer restricted by the level of an EDF or the Kohn-Sham scheme, but are driven by experimental data. For example, it has been shown that the nucleon correlation including the α-cluster formation on the surface affects charge radii when the numbers of protons and neutrons are similar [27,60]. In the present study, the CDG-based corrections aim at making the smallest possible corrections to the traditional DFT calculations and improving the capabilities in the description of charge radii. In other words, some missing but crucial terms in an existing EDF can be implemented by the CDG-based corrections in an implicit way. We randomly take 640 nuclei (10 batches) of about 900 nuclei with measured charge radii to date as the training set. The remaining nuclei are recorded as the validation set (see the SM [54] for details). It is recorded as an epoch when all nuclei on the training set have been trained once.

B. Machine-learning processes

The process of machine learning is divided into two stages: simulating the SHF + BCS results and correcting with experimental data. The evaluation of these processes is shown in Fig. 2. The first 3000 epochs are the stage of simulating the SHF + BCS results, during which $\mathrm{Loss}(0, \mathbf{w}) = L_\rho$ is being minimized. One can see that the loss functions of the training and validation sets almost overlap. This means neither ρ_c nor R_c is overfitted, which shows the generalization ability of the network. As a theoretical reference, the dashed magenta line is the Pearson χ^2 divergence $L_{\rm R,theo}$ on the validation set between the SHF + BCS and experimental values:

$$L_{\rm R,theo} = \frac{1}{N_{\rm v}}\sum_{{\rm n_u}\in{\rm val.}}\frac{\left(R_{\rm c,theo}^{({\rm n_u})} - R_{\rm c,exp}^{({\rm n_u})}\right)^2}{R_{\rm c,theo}^{({\rm n_u})}} \times 1\,{\rm fm}^{-1}, \tag{10}$$

where $N_{\rm v}$ is the number of nuclei on the validation set. After a short training, the loss values $L_{\rm R}$ overlap with $L_{\rm R,theo}$. This indicates that the network naturally captures the $R_{\rm c}$ information well in the process of learning the density distribution.

After 3000 epochs, the pretrained model is further tuned with an objective function, $\mathrm{Loss}(\mathcal{W} = 0.7, \mathbf{w})$, which allows the importance of experimental data to slightly exceed that of theoretical calculations. Thus, the charge density distributions under constraints are spontaneously corrected by the network. Since the correction makes the predicted charge radius close to the experimental value and the distribution naturally deviate from the model, the value of $L_{\rm R}$ falls and the value of L_ρ jumps as shown in Fig. 2. Remarkably, the training costs only 10 GPU min.

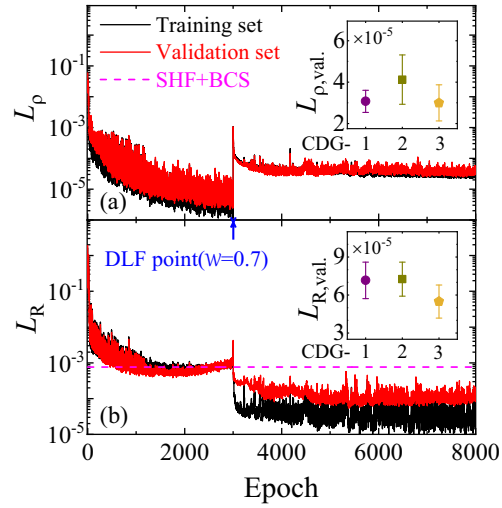

FIG. 2. (a) Normalized mean-square error for density distributions on the training and validation sets as a function of epochs with CDG-3. (b) Pearson χ^2 divergence for charge radii as a function of epochs with CDG-3. Inserts: The loss values on validation sets (a) $L_{\rho,\rm val}$ and (b) $L_{\rm R,val}$ for CDG-1, CDG-2, and CDG-3 with $\mathcal{W} = 0.7$.

It is clear that machine learning successfully eliminated a portion of residuals between SHF + BCS and experimental values. Such a process can be migrated to any other theoretical model. Certainly, we should be aware of the dependence of predictions on functionals and models, as theoretical calculations hold considerable weight in the target (loss) functions during the pretraining and calibration processes. Actually, such dependence is widely present in neural network-based nuclear structure studies and is one of the interesting research directions. An option for future research could be utilizing the BMA combined with multiple functionals for predictions, which has the potential to enhance the accuracy of predictions as model dependence is eliminated partially.

III. RESULTS

To explore the network performance, the errors of the densities and radii given by CDG-1, CDG-2, and CDG-3 on the validation set are plotted in the inserts of Fig. 2. It can be seen that the errors by CDG-3 are minimal for both densities and radii. This indicates the predictions of CDG-3 are closest to the experimental radii, while its corrections to the theory are the smallest. This agrees with our assumption.

Accuracy is more intuitive to show the CDGs' performance than error. Figure 3(a) shows the prediction accuracy of the three networks, where the accuracy indicates the percentages of the experimental data on the validation set falling within

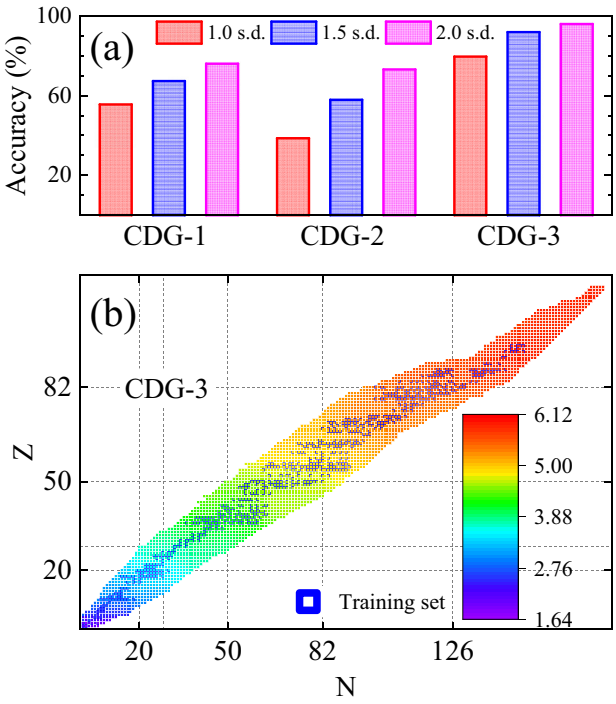

FIG. 3. (a) The accuracy of neural networks CDG-1, CDG-2, and CDG-3, expressed by the percentages of the experimental data on the validation set falling within different predicted standard deviations (red for 1 standard deviation (s.d.), blue for 1.5 s.d., and pink for 2 s.d.). (b) Charge radii predicted by CDG-3, where the nuclei in the training set are denoted with the blue squares. The results are shown in the unit of fm.

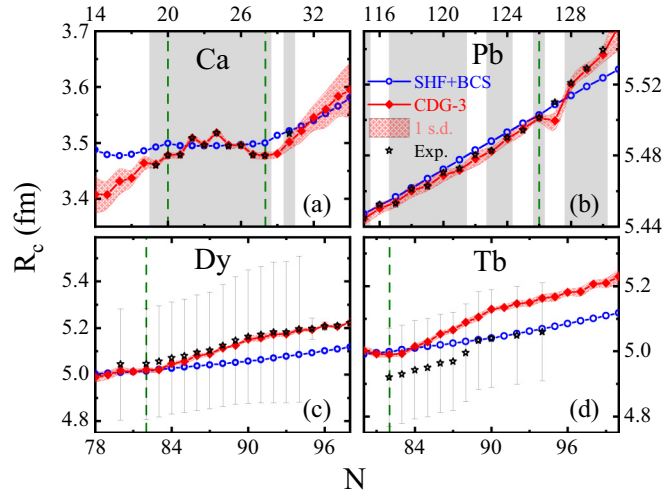

FIG. 4. Charge radii predicted by CDG-3 with 1 s.d. error for the Ca, Pb, Dy, and Tb isotopes, where the training regions are indicated by shadows and the magic numbers are indicated by the vertical dashed lines. The SHF + BCS results and experimental data are also shown for comparison.

different predicted standard deviations (red for 1 s.d., blue for 1.5 s.d., and pink for 2 s.d.). The values of accuracy are consistent with the errors on the validation set shown in Fig. 2 for different networks; i.e., the smaller the error, the higher the accuracy. In particular, the numbers of experimental data falling within 2 s.d. do not exceed 80% for CDG-1 and CDG-2. Meanwhile, the accuracy of CDG-3 (1 s.d.) is already close to 80%, which even reaches 96% for 2 s.d. Compared to CDG-1 and CDG-2, CDG-3 takes into account the odd-even staggering, pairing, and shell effects. Therefore, we conclude that the network structures considering more physical properties are more sufficient to improve the prediction accuracy, which is consistent with the mass research [31]. It can be speculated that CDG-3 can predict the radii of the remaining ∼2000 unmeasured nuclei with high precision. We present the predictions for nuclear charge radii in Fig. 3(b). In this regard, the risk of extrapolation still exists. Compared to the Kohn-Sham network proposed in Ref. [61], the current model lacks the modeling of single-particle information, which would lower its extrapolation performance. As such, this work draws on the experience of Bayesian neural networks to mitigate the risk of extrapolation. By constructing multiple networks with identical architectures through repeated training from different initializations, the results are then averaged to derive predictions that are relatively reliable.

By taking several isotopic chains as examples, we compare the charge radii calculated by SHF + BCS and the predictions of CDG-3 with 1 s.d. error with the available experimental data in Fig. 4. It is found that CDG-3, containing the parity, pairing, and shell effects, describes well the odd-even staggering for the Ca isotopes. Meanwhile, CDG-1 and CDG-2 failed in benchmarking the experimental data (see the SM [54]). For Pb isotopes, CDG-3 performs comparably with the theoretical model. The present predictions almost match the experimental data in the case of training several Pb isotopes. In particular, the predictive power remains strong even when we remove the Ca and Pb isotopes from the training set (see the SM [54]). Remarkably, it is believed that there are substantial beyond-mean-field effects in the charge radius evolution in the Ca isotopes, in particular, from ^{40}Ca to ^{48}Ca [62,63]. Given the excellent performance of CDG-3, we predicted the untrained Dy and Tb isotopes. For the Dy isotopes, the predictions are in agreement with experiment, but with much smaller uncertainty than the experimental error. The simultaneous precise description of the spherical (Pb and Ca) and deformed nuclei (Dy) is evidence of the validity of the CDG-based correction scheme. Originally, spherical symmetry is imposed in our theoretical calculations (SHF + BCS). However, after migrating the calculations to CDGs, the density can change more arbitrarily. Although the deformation is not fully considered physically (the densities before and after correction are still spherical, which in some sense correspond to the density distributions in the laboratory frame), the network also to some extent includes the information on the deformation under the influence of the features of input and the loss function. In detail, the shell and parity effects have been included in the input, with the experimental data integrated into the loss function. Those are closely tied to deformation, thus enabling the impact of deformation to be predicted from a macro perspective. From this point of view, we can compromise

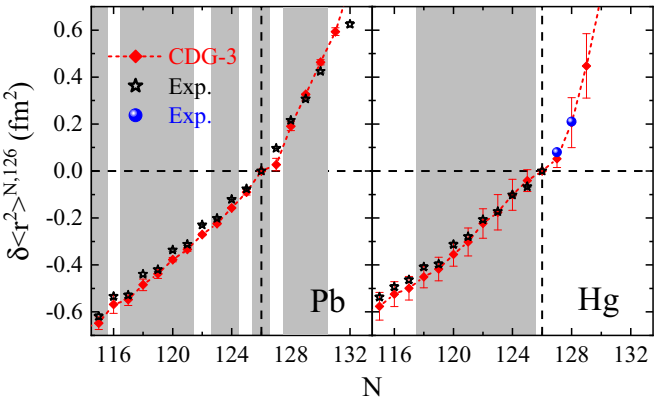

FIG. 5. The relative changes in the mean-square charge radii $\delta\langle r^2\rangle^{N,126}$ to ^{208}Pb (left panel) and ^{206}Hg (right panel). The experimental values in black stars are sourced from Ref. [1] and those in blue spherical symbols are sourced from Ref. [8]. The training regions are indicated by shadows and the magic number $N = 126$ is indicated by the vertical dashed lines.

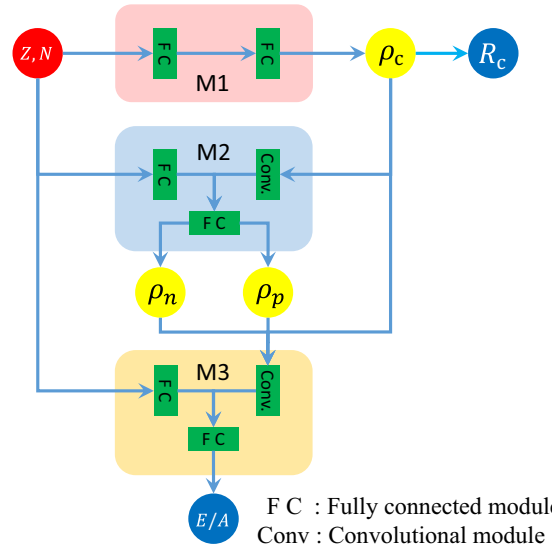

FIG. 6. Schematic diagram of the structure of a feed-forward neural network for the CDTBE map.

and assume that the density $\rho_c(r,\theta,\phi)$ has been angularly averaged out as $\rho_{c,pre}(r) = \int \rho_c(r,\theta,\phi)\, d\theta\, d\phi$. These conclusions well demonstrate the effectiveness of the present scheme beyond the existing forms of an EDF and the Kohn-Sham scheme. For the Tb isotopes, compared to the SHF + BCS results, the predicted isotope shifts are in good agreement with the experimental data, although the predicted charge radii are systematically larger than the present experimental central values by ∼0.1 fm. The above predictions can be testified to in the coming experiments.

As a further validation, we turn our attention to the kink in charge radii across the $N = 126$ shell closure, which is commonly believed to be present in relativistic structural models and to be absent in nonrelativistic frameworks. Regarding the kink, new experimental data for ^{207,208}Hg have been presented in the Ref. [8], along with theoretical explanations. The relative changes in the mean-square charge radii $\delta\langle r^2\rangle^{N,126}$ for lead and mercury are shown in Fig. 5, where $\delta\langle r^2\rangle^{N,126}$ is defined as

$$\delta\langle r^2\rangle^{N,126} = \langle r^2\rangle^{N} - \langle r^2\rangle^{126}. \tag{11}$$

In line with previous findings, there is generally good characterization of both the nuclei in the training regions and those internally inferred. A strong signal is manifested as the new experimental values for ^{207,208}Hg coincide accurately with the predicted average values of CDG-3, which again reinforces the reliability of our method. As such, this seems to imply that a neural network pretrained with a nonrelativistic model already contains the kink in the charge radii at the $N = 126$ neutron shell closure under experimental corrections. The majority of nonrelativistic EDFs, including the SkM* EDF used in this study, cannot reproduce such a kink, with the exception of a few EDFs such as SkI3 and SAMi. That is to say, CDG-3 has at least surpassed the level of the normal Skyrme interaction with the strength of the spin-orbit interaction $W_0 = W_0'$ [63]. The formation of a kink is the result of complex physical processes, involving tensor interaction, spin-orbit interaction, symmetry energy, and the pairing interaction [63]. Single-particle occupation probability decided by these complex correlations ultimately determines the existence of kink or anti-kink. In our method, the inclusion of experimental data from the corresponding nuclei (such as ^{210,211,212}Pb) is a natural way to incorporate such a kink at $N = 126$ shell closure into the neural network. In this sense, it is reasonable to assume that the majority of physical phenomena, including those yet to be discovered, have already been encapsulated to a large extent in the form of a black box.

IV. EXTENSION AND DISCUSSION

To understand what neural network-based calibration brings to DFT further, the impacts of the calibrated charge densities on binding energies are investigated by constructing and linking a series of neural networks.

According to the Hohenberg-Kohn theorem [52,64], there exists a bijective map between the local matter density and the one-body potential. This means that the corrections on the density distributions can be mapped to other observables, such as binding energies [65]. Thus, we reconstruct a map from the charge density to the binding energy (CDTBE). The schematic structure of such a feed-forward neural network is shown in Fig. 6. There are three parts in the CDTBE map—M1, M2, and M3: M1 is the previously trained charge density generator, and we take the CDG-3; M2 is a map from the charge density to the matter density as the inverse of Eq. (7), where N and Z are also inputs; M3 is a map from the neutron, proton, and charge densities to the binding energy per nucleon. It is found that the networks adequately capture the relationship between densities and binding energy with different DFT effective interactions. As the charge density updates in CDG-3, the residual information flows to other observables, i.e., $\delta R_c \to \delta\rho_c \to \{\delta\rho_n, \delta\rho_p\} \to \delta(E/A)$. It is worth mentioning that, while M1 is driven by the experimental data, M2

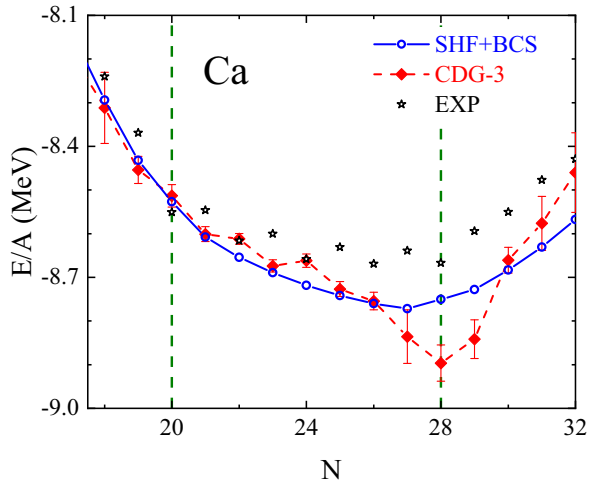

FIG. 7. Binding energy per nucleon of Ca isotopes with the statistical error (1 s.d.) obtained by CDG-3 residual information flow. The SHF + BCS results and experimental data are also shown.

and M3 adopted here are trained by the relations among the charge, proton, and neutron densities and the binding energy generated by the SHF + BCS calculations with given EDFs: SkM, SkM*, SkIII, SLy4, SkT, and SkT3 [66].

Taking the Ca isotopes as examples, the binding energies per nucleon obtained by the CDG-3 residual information flow are shown in Fig. 7. We note that the propagated corrections provide a better description, except ^{47,48,49}Ca. Especially for ^{40,42,44}Ca, both radii [see Fig. 4(a)] and binding energies coincide well with the experimental values. This illustrates that the physical features contained in the CDTBE map based on the density functional theory are adequate for these nuclei. In contrast, for ^{48}Ca, the radius is corrected to a value smaller than that of SHF + BCS, leading to an increase of nuclear densities in the central region and a further decrease of binding energy, which is consistent with the equation of state. However, it results in further deviations from the experimental value, which can be attributed to the incompatibility of the corrected densities with the used functional form. In Refs. [62,63], it is discussed that an indispensable beyond-mean-field effect appears near ^{48}Ca. As a future study, one can train M3 using EDFs with different structures to assess the effect of incorporating different terms. This may provide additional insights into the overbinding observed in ^{48}Ca.

On the other hand, we believe the present studies provide a systematically improvable method for the global description of various properties. For example, more corrections can be implemented step by step, such as incorporating experimental corrections for neutron-skin thicknesses in M2 and using the binding-energy residuals between theory and experiment as output for M3. Following this strategy, the description of each kind of observable can be improved by making segmental calibrations.

V. SUMMARY

In this study, a novel supervised learning on the combination of the theoretical charge density distributions and the experimental charge radii has been performed. In such a way, the physics features embedded in nuclear density functional theory can be preserved to a large extent and the feedback from the experimental data can be considered quantitatively. It is found that the description of charge radii can be improved globally on the nuclear chart. In particular, the specific charge radii evolution in the Ca isotopes can be well reproduced by taking the parity, pairing, and shell effects into account. This property remains valid even when all the Ca isotopes are excluded in the learning set. The predictive power is also shown with the charge radii in the Dy and Tb isotopes, where the present experimental uncertainties are much larger than the prediction uncertainties. Moreover, the charge-radius kink phenomenon observed across the $N = 126$ shell closure has been successfully reproduced, and it matches closely with the latest experimental data on ^{207,208}Hg.

Inspired by the Hohenberg-Kohn theorem, an information flow from the charge density to the binding energy has also been constructed and investigated. It is found that for the Ca isotopes the improvement in the description of charge radii can also be propagated to the improvement in the description of binding energies, except ^{47,48,49}Ca. The corresponding analysis implies the existence of the incompatibility of the corrected densities with the used functional form near ^{48}Ca, which merits further discussions on functional forms.

Moreover, constructing the experimental-based observable networks will facilitate the realization of realistic nonparametric Hohenberg-Kohn maps for nuclear complex systems. Along this direction, we will not only improve the description of different nuclear observables consistently but also strengthen the interpretability of the supervised learning.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under Grants No. 12005175 and No. 11875070, the Fundamental Research Funds for the Central Universities under Grant No. SWU119076, the JSPS Grant-in-Aid for Early-Career Scientists under Grant No. 18K13549, the JSPS Grant-in-Aid for Scientific Research (S) under Grant No. 20H05648, the JSPS Grant-in-Aid for Research Activity Start-up under Grant No. 22K20372, the JSPS Grant-in-Aid for Transformative Research Areas (A) under Grant No. 23H04526, the JSPS Grant-in-Aid for Scientific Research (B) under Grant No. 23H01845, the JSPS Grant-in-Aid for Scientific Research (C) under Grant No. 23K03426, and the Anhui project (Grant No. Z010118169). This work is also partially supported by the RIKEN Pioneering Project: Evolution of Matter in the Universe, the RIKEN Special Postdoctoral Researchers Program, and the Science and Technology Hub Collaborative Research Program from the RIKEN Cluster for Science, Technology and Innovation Hub (RCSTI).

[1] I. Angeli and K. Marinova, At. Data Nucl. Data Tables **99**, 69 (2013).
[2] R. H. Helm, Phys. Rev. **104**, 1466 (1956).
[3] R. Hofstadter, Rev. Mod. Phys. **28**, 214 (1956).
[4] P. Campbell, I. Moore, and M. Pearson, Prog. Part. Nucl. Phys. **86**, 127 (2016).
[5] J. Wood, K. Heyde, W. Nazarewicz, M. Huyse, and P. van Duppen, Phys. Rep. **215**, 101 (1992).
[6] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Rev. Mod. Phys. **75**, 121 (2003).
[7] X.-F. Li, D.-Q. Fang, and Y.-G. Ma, Nucl. Sci. Tech. **27**, 71 (2016).
[8] T. Day Goodacre, A. V. Afanasjev, A. E. Barzakh, B. A. Marsh, S. Sels, P. Ring, H. Nakada, A. N. Andreyev, P. Van Duppen, N. A. Althubiti *et al.* (CLAS Collaboration), Phys. Rev. Lett. **126**, 032502 (2019).
[9] B. Dreher, J. Friedrich, K. Merle, H. Rothhaas, and G. Lührs, Nucl. Phys. A **235**, 219 (1974).
[10] J. Friar and J. Negele, Adv. Nucl. Phys. **8**, 219 (1975).
[11] A. Mutschler, A. Lemasson, O. Sorlin, D. Bazin, C. Borcea *et al.*, Nat. Phys. **13**, 152 (2017).
[12] X.-H. Fan, G.-C. Yong, and W. Zuo, Phys. Rev. C **99**, 041601(R) (2019).
[13] D. R. Yennie, D. G. Ravenhall, and R. N. Wilson, Phys. Rev. **95**, 500 (1954).
[14] N. Ullah, Pramana **43**, 165 (1994).
[15] H. Euteneuer, J. Friedrich, and N. Voegler, Nucl. Phys. A **298**, 452 (1978).
[16] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **106**, 162 (1957).
[17] N. N. Bogoljubov, V. V. Tolmachov, and D. V. Širkov, Fortschr. Phys. **6**, 605 (1958).
[18] T. Dytrych, K. D. Sviratcheva, C. Bahri, J. P. Draayer, and J. P. Vary, Phys. Rev. Lett. **98**, 162503 (2007).
[19] D. J. Dean, G. Hagen, M. Hjorth-Jensen, T. Papenbrock, and A. Schwenk, Phys. Rev. Lett. **101**, 119201 (2008).
[20] R. Roth, Phys. Rev. C **79**, 064324 (2009).
[21] T. Dytrych, K. D. Launey, J. P. Draayer, P. Maris, J. P. Vary, E. Saule, U. Catalyurek, M. Sosonkina, D. Langr, and M. A. Caprio, Phys. Rev. Lett. **111**, 252501 (2013).
[22] T. Neff, H. Feldmeier, and W. Horiuchi, Phys. Rev. C **92**, 024003 (2015).
[23] Y. Zhang, Y. Chen, J. Meng, and P. Ring, Phys. Rev. C **95**, 014316 (2017).
[24] T. Naito, X. Roca-Maza, G. Colò, and H. Liang, Phys. Rev. C **99**, 024309 (2019).
[25] C. A. Bertulani and J. Valencia, Phys. Rev. C **100**, 015802 (2019).
[26] Z. X. Yang, X. L. Shang, G. C. Yong, W. Zuo, and Y. Gao, Phys. Rev. C **100**, 054325 (2019).
[27] J. Tanaka, Z. Yang, S. Typel, S. Adachi, S. Bai *et al.*, Science **371**, 260 (2021).
[28] M. Duer *et al.* (CLAS Collaboration), Nature (London) **560**, 617 (2018).
[29] G. A. Miller, Phys. Rev. Lett. **123**, 232003 (2019).
[30] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature (London) **323**, 533 (1986).
[31] Z. Niu and H. Liang, Phys. Lett. B **778**, 48 (2018).
[32] C. Ma, M. Bao, Z. M. Niu, Y. M. Zhao, and A. Arima, Phys. Rev. C **101**, 045204 (2020).
[33] S. Athanassopoulos, E. Mavrommatis, K. Gernoth, and J. Clark, Nucl. Phys. A **743**, 222 (2004).
[34] K. Gernoth, J. Clark, J. Prater, and H. Bohr, Phys. Lett. B **300**, 1 (1993).
[35] R. Utama, W.-C. Chen, and J. Piekarewicz, J. Phys. G: Nucl. Part. Phys. **43**, 114002 (2016).
[36] D. Wu, C. L. Bai, H. Sagawa, and H. Q. Zhang, Phys. Rev. C **102**, 054323 (2020).
[37] X.-X. Dong, R. An, J.-X. Lu, and L.-S. Geng, Phys. Rev. C **105**, 014308 (2022).
[38] G. Co', M. Anguiano, and A. M. Lallena, Phys. Rev. C **105**, 034320 (2022).
[39] R.-D. Lasseri, D. Regnier, J.-P. Ebran, and A. Penon, Phys. Rev. Lett. **124**, 162502 (2020).
[40] Y. Wang, X. Zhang, Z. Niu, and Z. Li, Phys. Lett. B **830**, 137154 (2022).
[41] G. A. Negoita, J. P. Vary, G. R. Luecke, P. Maris, A. M. Shirokov *et al.*, Phys. Rev. C **99**, 054308 (2019).
[42] G. Saxena, P. K. Sharma, and P. Saxena, J. Phys. G: Nucl. Part. Phys. **48**, 055103 (2021).
[43] Z. M. Niu, H. Z. Liang, B. H. Sun, W. H. Long, and Y. F. Niu, Phys. Rev. C **99**, 064307 (2019).
[44] Z.-A. Wang, J. Pei, Y. Liu, and Y. Qiang, Phys. Rev. Lett. **123**, 122501 (2019).
[45] C. Y. Qiao, J. C. Pei, Z. A. Wang, Y. Qiang, Y. J. Chen, N. C. Shu, and Z. G. Ge, Phys. Rev. C **103**, 034621 (2021).
[46] J. W. Clark and H. Li, Int. J. Mod. Phys. B **20**, 5015 (2006).
[47] S. A. Bass, A. Bischoff, J. A. Maruhn, H. Stöcker, and W. Greiner, Phys. Rev. C **53**, 2358 (1996).
[48] R. Utama, J. Piekarewicz, and H. B. Prosper, Phys. Rev. C **93**, 014311 (2016).
[49] L. Neufcourt, Y. Cao, W. Nazarewicz, E. Olsen, and F. Viens, Phys. Rev. Lett. **122**, 062502 (2019).
[50] L. Neufcourt, Y. Cao, S. A. Giuliani, W. Nazarewicz, E. Olsen, and O. B. Tarasov, Phys. Rev. C **101**, 044307 (2020).
[51] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz *et al.*, arXiv:2303.12712.
[52] J. R. Moreno, G. Carleo, and A. Georges, Phys. Rev. Lett. **125**, 076402 (2020).
[53] Z.-X. Yang, X.-H. Fan, P. Yin, and W. Zuo, Phys. Lett. B **823**, 136650 (2021).
[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevC.108.034315 for the detailed parameters not listed in the main text as well as the performance of the model.
[55] K. Simonyan and A. Zisserman, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, May 7-9, 2015, Conference Track Proceedings*, edited by Y. Bengio and Y. LeCun (ICLR, 2015).
[56] N. Ketkar, in *Deep Learning with Python* (Apress, New York, 2017), pp. 195–208.
[57] D. P. Kingma and J. Ba, in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, May 7–9, 2015, Conference Track Proceedings*, edited by Y. Bengio and Y. LeCun (ICLR, 2015).
[58] P.-G. Reinhard, in *Computational Nuclear Physics 1* (Springer, Berlin, 1991), pp. 28–50.
[59] J. Bartel, P. Quentin, M. Brack, C. Guet, and H.-B. Håkansson, Nucl. Phys. A **386**, 79 (1982).
[60] S. Typel, Phys. Rev. C **89**, 064321 (2014).

[61] Z.-X. Yang, X.-H. Fan, Z.-P. Li, and H. Liang, Phys. Lett. B **840**, 137870 (2023).
[62] U. C. Perera, A. V. Afanasjev, and P. Ring, Phys. Rev. C **104**, 064313 (2021).
[63] T. Naito, T. Oishi, H. Sagawa, and Z. Wang, Phys. Rev. C **107**, 054307 (2023).
[64] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[65] K. Ryczko, D. A. Strubbe, and I. Tamblyn, Phys. Rev. A **100**, 022512 (2019).
[66] J. Friedrich and P.-G. Reinhard, Phys. Rev. C **33**, 335 (1986).