Providing physics guidance in Bayesian neural networks from the input layer: The case of giant dipole resonance predictions

Xiaohang Wang, Long Zhu, and Jun Su ●*

Sino-French Institute of Nuclear Engineering and Technology, Sun Yat-sen University, Zhuhai 519082, China

Background: A Bayesian neural network (BNN) approach has been applied to evaluate and predict the nuclear data. The BNN is a numerical algorithm. When one incorporates this algorithm in nuclear physics analyses, how to maintain the scientific rigor is a key problem and presents new challenges.

Purpose: In this paper, a case study on giant dipole resonance (GDR) energy is presented to illustrate the effectiveness and maneuverability of the method to provide physics guidance in the BNN from the input layer.

Methods: Pearson's correlation coefficients are applied to assess the statistical dependence between nuclear properties in the ground state and the GDR energies. Then the optimal ground-state properties are employed as variables of the input layer in the BNN to evaluate and predict the GDR energies.

Results: Those selected ground-state properties actively contribute to reduce the predicted errors and avoid the overfitting.

Conclusions: This paper gives a demonstration to find effects of the GDR energy by using the BNN without the physics motivated model, which may be helpful to discover physics effects from the complex nuclear data.

I. INTRODUCTION

Bayesian machine learning is a powerful tool to make predictions with confidence intervals after training by data. In this respect, the Bayesian neural network (BNN) has rapidly drawn much attention in nuclear physics in the past few years. However, the BNN is merely a versatile and numerical algorithm. When one incorporates this algorithm in nuclear physics analyses, how to maintain the scientific rigor is a key problem and presents new challenges. The seminal ideas to use neural networks in nuclear physics date back to early works [1–4]. A clear illustration of the underlying philosophy behind the implementation of the BNN approach originally appears in works by Utama *et al.*, where theoretical predictions of nuclear masses [5] and nuclear charge radii [6] were presented. In those works, they claimed that one can include as much physics as possible in the initial prediction by using physics motivated models and use the BNN to fine tune these models by modeling the residuals. To facilitate the narrative, this approach is hereinafter called the residual approach.

The residual approach has been proved to be a very powerful tool to improve the physical model based predictions of nuclear observables, such as nuclear binding energies [7], fission yields [8], and isotopic cross sections in spallation [9]. As stated in Ref. [5], the basic requirement of the residual approach is the existence of a robust physics motivated model to provide physics information. The physics motivated model ensures efficient extrapolation in the region lacking experimental information. In contrast, the residual provided by the

BNN is tiny and hence will not cause a large deviation to prior physics knowledge, even though the BNN is merely a numerical algorithm and does not contain any physics information. However, a challenge arises if the physics motivated model is not available. In Ref. [9], it has been shown that the BNN method without initial prediction by physics motivated models provides good predictions for the cross sections of fragments in spallation. But when the predictions are extended to the region far from the measured data, one observes a numerical fluctuation that is contrary to prior physics knowledge, i.e., cross sections of neutron-rich fragments decrease exponentially with increasing neutron excess. Responding to this challenge, the physics motivated model should be developed to provide reliable initial predictions, otherwise physics guidance should be provided to avoid the overfitting of the BNN. In fact, beyond the residual approach, successful attempts to further consider physical information in the BNN approach can also be found in the literature. For example, it was found that better predictive performance for nuclear masses can be achieved by adding two quantities related to the well-known nuclear pairing and shell effects into the input layer of the BNN [10]. Inspired by this paper, we try to add ground-state properties into the input layer in the BNN so that it can predict the giant dipole resonance (GDR) energy without the initial prediction by the physics motivated model.

The GDR is the most well-known and oldest collective excited mode of nuclei in the excitation energy from single nucleon separation energy to dozens of MeV [11–13]. Over the last decades, hundreds of GDR data have been experimentally measured using photonuclear reactions or γ decay [14–16]. The γ spectrum displays a wide peak, which can be characterized by a Lorentzian function with

---

*Corresponding author: sujun3@mail.sysu.edu.cn

	A	Z	δ	Q_α	$Q_{p\alpha}$	$Q_{n\alpha}$	$Q_{\beta n}$	$Q_{d\alpha}$	$Q_{4\beta}$	$Q_{2\beta}$	E_d	Q_β	$Q_{\epsilon p}$	β_2	β_6	β_4	S_p	B	E_s	S_{2p}	S_n	S_{2n}	$\bar{E}_r$	ΔE_r
A	1																							
Z	1	1																						
δ	0.96	0.95	1																					
Q_α	0.93	0.94	0.84	1																				
$Q_{p\alpha}$	0.91	0.91	0.88	0.9	1																			
$Q_{n\alpha}$	0.92	0.92	0.83	0.94	0.93	1																		
$Q_{\beta n}$	0.77	0.75	0.84	0.69	0.83	0.67	1																	
$Q_{d\alpha}$	0.85	0.85	0.8	0.83	0.95	0.93	0.74	1																
$Q_{4\beta}$	0.74	0.72	0.84	0.64	0.71	0.62	0.88	0.62	1															
$Q_{2\beta}$	0.54	0.54	0.55	0.59	0.6	0.53	0.7	0.51	0.66	1														
E_d	0.64	0.62	0.77	0.55	0.65	0.54	0.88	0.56	0.97	0.65	1													
Q_β	0.59	0.57	0.68	0.51	0.7	0.54	0.91	0.67	0.76	0.61	0.8	1												
$Q_{\epsilon p}$	0.6	0.62	0.4	0.69	0.5	0.68	0.26	0.53	0.19	0.33	0.073	0.16	1											
β_2	0.32	0.32	0.32	0.39	0.38	0.33	0.42	0.31	0.39	0.76	0.36	0.35	0.2	1										
β_6	0.13	0.13	0.1	0.1	0.091	0.14	-0.0038	0.1	0.079	-0.06	0.065	-0.02	0.073	-0.22	1									
β_4	-0.014	-0.012	-0.0036	-0.011	-0.027	-0.065	0.038	-0.089	0.038	-0.017	0.031	0.042	-0.072	0.14	0.039	1								
S_p	-0.59	-0.6	-0.44	-0.63	-0.67	-0.62	-0.46	-0.68	-0.19	-0.34	-0.12	-0.44	-0.66	-0.27	0.099	0.067	1							
B	-0.69	-0.69	-0.68	-0.69	-0.66	-0.65	-0.46	-0.58	-0.39	-0.41	-0.3	-0.3	-0.3	-0.37	-0.075	-0.00096	0.3	1						
E_s	-0.78	-0.78	-0.79	-0.66	-0.67	-0.67	-0.53	-0.63	-0.51	-0.29	-0.4	-0.39	-0.38	-0.24	0.066	0.14	0.39	0.75	1					
S_{2p}	-0.76	-0.78	-0.58	-0.84	-0.68	-0.82	-0.4	-0.67	-0.33	-0.39	-0.21	-0.23	-0.92	-0.28	-0.053	0.012	0.73	0.48	0.55	1				
S_n	-0.82	-0.82	-0.85	-0.79	-0.8	-0.8	-0.74	-0.76	-0.71	-0.54	-0.67	-0.69	-0.4	-0.33	-0.13	-0.021	0.35	0.65	0.61	0.49	1			
S_{2n}	-0.89	-0.88	-0.94	-0.86	-0.87	-0.81	-0.87	-0.76	-0.86	-0.63	-0.82	-0.7	-0.35	-0.4	-0.11	-0.015	0.35	0.65	0.66	0.53	0.88	1		
$\bar{E}_r$	-0.92	-0.92	-0.89	-0.89	-0.88	-0.86	-0.8	-0.8	-0.78	-0.7	-0.7	-0.64	-0.54	-0.52	-0.085	-0.059	0.54	0.63	0.68	0.7	0.78	0.88	1	
ΔE_r	0.21	0.21	0.23	0.24	0.27	0.19	0.34	0.18	0.26	0.59	0.27	0.31	0.079	0.72	-0.079	0.11	-0.17	-0.28	-0.17	-0.13	-0.22	-0.31	-0.47	1

Legend:
- A: mass number
- Z: charge number
- $\delta=(A-2Z)/A$: isospin asymmetry
- Q_α: α decay energy
- $Q_{p\alpha}$: Q-value of (p,α) reaction
- $Q_{n\alpha}$: Q-value of (n,α) reaction
- $Q_{\beta n}$: Q-value of (β^-,n) reaction
- $Q_{d\alpha}$: Q-value of (d,α) reaction
- $Q_{4\beta}$: four β^- decay energy
- $Q_{2\beta}$: two β^- decay energy
- E_d: deformation energy
- Q_β: β^- decay energy
- $Q_{\epsilon p}$: Q-value of (β^+,p) reaction
- β_2: quadrupole deformation
- β_6: hexacontatetrapole deformation
- β_4: hexadecapole deformation
- S_p: proton separation energy
- B: binding energy per nucleon
- E_s: shell correction
- S_{2p}: two proton separation energy
- S_n: neutron separation energy
- S_{2n}: two neutron separation energy
- $\bar{E}_r=0.5(E_{r1}+E_{r2})$: Average of resonance energies
- $\Delta E_r=E_{r2}-E_{r1}$: Difference of resonance energies
- E_{r1}: low resonance energy
- E_{r2}: high resonance energy

FIG. 1. Pearson's correlation coefficient ρ between ground-state properties and GDR energies. The pairs of variables will be strongly positive correlated for high ρ value (or identical for a correlation $\rho = 1$), and strongly negative correlated for a low ρ value (or fully opposed for a correlation $\rho = -1$). A coefficient ρ near zero indicates that there is no correlation between pairs of variables. The ground-state properties are taken from Refs. [31–34]. The data of the GDR energy are taken from Ref. [18].

GDR parameters including the peak energy (i.e., GDR energy), the width, and the strength [17,18]. Various models have been developed to investigate the GDR, including Hartree-Fock-based random phase approximation (HF-based RPA) [19,20], time-dependent Hartree-Fock theory (TDHF) [21,22], the Boltzmann-Uehling-Uhlenbeck model [23,24], extended quantum molecular dynamics [25], and macroscopic Langevin equation coupling with the Skyrme Hartree-Fock-Bogoliubov model [26,27]. Several effects of the GDR, such as the mass dependence, the isospin dependence, and the shape effect, have been studied [23,26,28–30]. Those effects were also found in the nuclear properties in the ground state, the data of which are more abundant than those of GDR [31–33]. In fact, proposed by microscopic theory such as HF-based RPA and TDHF, the GDR energies depend strongly on the HF ground-state wave function [20,22]. In the macroscopic point of view, the GDR energy has been determined when the proton and neutron density distributions in the ground state are known [26]. Thus, there may exist dependence between the ground-state properties and the GDR energies.

II. METHOD

To assess the statistical dependence between two variables of the ground-state properties and GDR energies, Pearson's correlation coefficients $\rho(x, y)$ are calculated:

$$\rho(x, y) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}, \qquad (1)$$

where x_i is the ith data of the variable x, $\bar{x}$ is the average of the variable x over the data, and $\sum_i$ expresses the summation over the data. A similar situation exists for the variable y. Figure 1 lists the correlation coefficients between

two variables of the ground-state properties as well as the GDR energies. The ground-state properties include the derived quantities of atomic masses taken from the Ame2012 evaluation [31–33], and the extracted quantities from the Weizsäcker-Skyrme mass formula [34]. The GDR energy includes the low and high peak energies (E_{r1} and E_{r2}), which are taken from Ref. [18]. Generally speaking, there are two components of the GDR energy for quadrupole deformation nuclei, but there is only one for spherical nuclei. In Ref. [18], the high resonance energies are missing for spherical nuclei. We use $E_{r2} = E_{r1}$ when the high resonance energy is missing in Ref. [18]. It means that two peaks of the GDR spectrum overlap for spherical nuclei.

Intuitively, the pairs of variables will be strongly positive correlated for high ρ value (or identical for a correlation $\rho = 1$), and strongly negative correlated for a low ρ value (or fully opposed for a correlation $\rho = -1$). The above two cases both mean that those two variables contain similar physical information, and hence one variable can be used to predict another. A coefficient ρ close to zero indicates that there is no correlation between the pair of variables. We refer to Ref. [35] for the quantitative interpretation of the correlation coefficients, where the degrees of correlation are divided into five levels. The correlation is negligible for $|\rho| \in [0.0, 0.1)$, weak for $|\rho| \in [0.1, 0.4)$, moderate for $|\rho| \in [0.4, 0.7)$, strong for $|\rho| \in [0.7, 0.9)$, and very strong for $|\rho| \in [0.9, 1.0]$.

$\rho(A, \bar{E}_r) = \rho(Z, \bar{E}_r) = -0.92$ from Fig. 1 reveals that A (or Z) is the ground-state property the correlation of which to the average resonance energies is the strongest. On the other hand, the strong correlation between the isospin and the GDR energies is also found, i.e., $\rho(\delta, \bar{E}_r) = -0.89$. The mass number A and isospin asymmetry δ are independent variables in physics, but for the data set involved in Fig. 1 their correlation is very strong [$\rho(A, \delta) = 0.96$]. To understand this statement,

one should remember that only the GDR data for the stable nuclei are available, and the isospin asymmetry δ increases with increasing mass for the stable nuclei. It is indicated that for the data set only involving stable nuclei the strong isospin dependence of the GDR energy does not give us any more information after strong mass dependence has been found. For the difference of resonance energies ΔE_r, the correlation coefficient with quadrupole deformation is $\rho(\Delta E_r, \beta_2) = 0.72$, which indicates a strong correlation. Other correlations to ΔE_r are moderate or even weaker. So, concerning the available data, the splitting of the GDR energies is a quadrupole deformation effect.

As mentioned above, the mass and quadrupole deformation effects of the GDR energies can be extracted from the two-dimensional correlation of the data. High-dimensional correlation should be used to find more effects of the GDR energies. In this respect, the BNN is a powerful tool. On the other hand, we do not need all the ground-state properties. In Fig. 1, one can find not only very strong cases, such as $\rho(Q_{p\alpha}, Q_{d\alpha}) = 0.94$, but also negligible cases, such as $\rho(Q_{\varepsilon p}, Q_{2\beta}) = 0.073$ for the correlations between the ground-state properties. We can say that the data of $Q_{p\alpha}$ and $Q_{d\alpha}$ provide similar physical information but those of $Q_{\varepsilon p}$ and $Q_{2\beta}$ provide independent physical information. Therefore, when the data sets of the ground-state properties are used to predict the GDR energies, useful information should be selected by the BNN approach.

Early works of the BNN approach include Refs. [36–38], which laid the foundation in this area. In the following we introduce the approach briefly. The mapping from ground-state properties X (input layer) to the GDR energies E_r (output layer) is established by the neural network with one hidden layer:

$$E_r(X, \theta) = a + \sum_{j=1}^{H} b_j \mathrm{logsig}\left(c_j + \sum_{i=1}^{l} d_{ji} X_i\right), \quad (2)$$

where $\theta = \{a, b_j, c_j, d_{ji}\}$ are the parameters in the neural network, $\{a, c_j\}$ are biases, and $\{b_j, d_{ji}\}$ are weights. The sigmoid function logsig is used as the activation function. In the following the reasons to choose one hidden layer and the sigmoid function will be discussed. H is the number of hidden neurons, and l is the number of input neurons. The outputs are denoted by two-dimensional variable $E_r = \{E_{r1}, E_{r2}\}$. The inputs $X = \{X_1, X_2, \ldots, X_l\}$ include one or several ground-state properties shown in Fig. 1.

Under the Bayesian theorem, the posterior distribution of parameters θ given data set D is expressed as

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{\int P(D|\theta)P(\theta)d\theta}, \quad (3)$$

where $P(\theta)$ is the prior distribution of θ, and $P(D|\theta)$ is the likelihood of D given θ. $D = \{X^{(n)}, E_r^{(n)}\}_{n=1}^{N_d}$ include both the ground-state properties $X^{(n)}$ and the GDR energy $E_r^{(n)}$, where N_d is the sample size of the available data.

For a nucleus with known ground-state properties X^*, the expected value of GDR energy $\hat{E}_r^*$ is expressed as the integration

$$\hat{E}_r^* = \int E_r(X^*, \theta)P(\theta|D)d\theta. \quad (4)$$

Monte Carlo techniques are applied to calculate the above integration:

$$\hat{E}_r^* \approx \frac{1}{N_s} \sum_{k}^{N_s} E_r(X^*, \theta^{(k)}), \quad (5)$$

where $\theta^{(k)}$ ($k = 1, 2, \ldots, N_s$) is the kth sample drawn from the posterior distributions $P(\theta|D)$, and N_s is the number of samples. A 95% confidence level is used to assess the uncertainty of the prediction. The confidence interval of the GDR energy E_r^* is

$$E_r^* = \hat{E}_r^* \pm 1.96 \frac{\sigma}{\sqrt{N_s}}, \quad (6)$$

where σ is the standard deviation of the samples $E_r(X^*, \theta^{(k)})$.

As the computation of the posterior distribution $P(\theta|D)$ is intractable due to the high dimension of parameters, the variation inference [39–41] is applied to find an approximation of $P(\theta|D)$. The variation inference tries to find κ so that $q(\theta|\kappa)$ is of minimum distance from $P(\theta|D)$ measured by Kullback-Leibler (KL) divergence:

$$\begin{aligned} \theta &= \arg\min \mathrm{KL}\left[q(\theta|\kappa)||P(\theta|D)\right] \\ &= \arg\min E_{q(\theta|\kappa)}\left[\ln \frac{q(\theta|\kappa)}{P(\theta|D)}\right] \\ &= \arg\min E_{q(\theta|\kappa)}\left[\ln \frac{q(\theta|\kappa)P(D)}{P(D|\theta)P(\theta)}\right] \\ &= \arg\min \sum_k [\ln q(\theta^{(k)}|\kappa) - \ln P(\theta^{(k)}) - \ln P(D|\theta^{(k)})]. \end{aligned}$$

$$(7)$$

There are 445 sets of data available. Tenfold cross validation is applied to determine the form of activation function, the number of hidden neurons, and the number of layers. We repeat the tenfold cross validation ten times, and for each time the validation data are used to calculate the root-mean-square error (RMSE) of predictions. The average and standard deviation of the RMSE of tenfold cross validation are calculated. We compare four activation functions: sigmoid, tanh, softplus, and ReLU (rectified linear unit). The curves of these four functions are shown in the upper panels in Fig. 2. Also, we take 10, 20, 30, 40, and 50 hidden neurons into consideration. One hidden layer and two hidden layers are compared as well. The results of the average and standard deviation of RMSE as a function of the hidden neuron are shown in Figs. 2(e) and 2(f). From the figure, we can see that the performance of the sigmoid function and one hidden layer are the best among four activation functions. We find that 10, 20, and 30 hidden neurons provide the similar RMSEs (about 1.3). It is the smallest value among of all cases. Therefore, we consider one hidden layer with ten hidden neurons and apply the sigmoid function as the activation function.

Figure 3 reports how the value of loss changes with the number of iterations. From the figure, we can see that the loss converges quickly, and the number of iterations 1000 is enough to reach a small and stable loss. In addition, the standard normal distribution is used as the prior of weights, and 10 000 samples are drawn for each predicted variable.
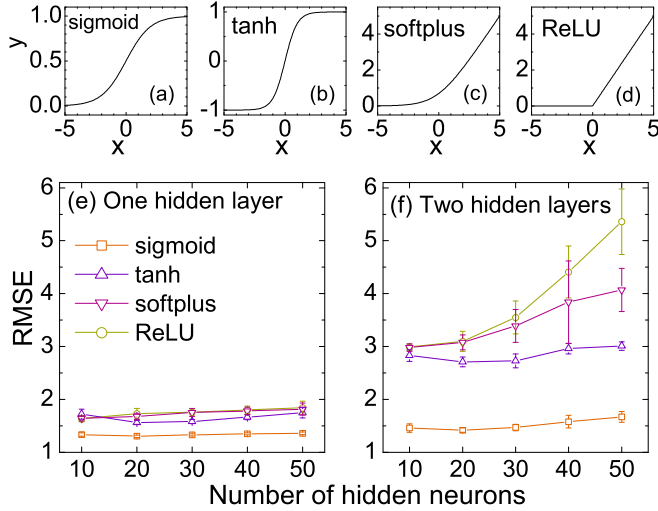
FIG. 2. Comparing the root-mean-square error of predictions using different activation functions and hidden layers.

III. RESULTS AND DISCUSSION

The BNN predictions are evaluated by the rms deviations,

$$
\mathrm{rms} = \sqrt{\frac{1}{2N_d}\left[\sum_{n=1}^{N_d}\left(\hat{E}_{r1}^{(n)} - E_{r1}^{(n)}\right)^2 + \sum_{n=1}^{N_d}\left(\hat{E}_{r2}^{(n)} - E_{r2}^{(n)}\right)^2\right]},
$$
(8)

where $E_{r1}^{(n)}$ and $E_{r2}^{(n)}$ are the nth data of low and high GDR energies, and $\hat{E}_{r1}^{(n)}$ and $\hat{E}_{r2}^{(n)}$ are their predicted values in the BNN.

One-dimensional input variable $X = \{X_i\}$ is considered, where X_i is one of the ground-state properties. The rms deviation for each input ground-state property is shown as a bar marked by $\{X_i\}$ in Fig. 4. It is found that the minimum of the rms deviations is obtained when using the input of mass number A or charge number Z (rms = 0.96 or 0.98 MeV, respectively). This is consistent with the fact in Fig. 1 that the strongest correlation is found between the average of GDR
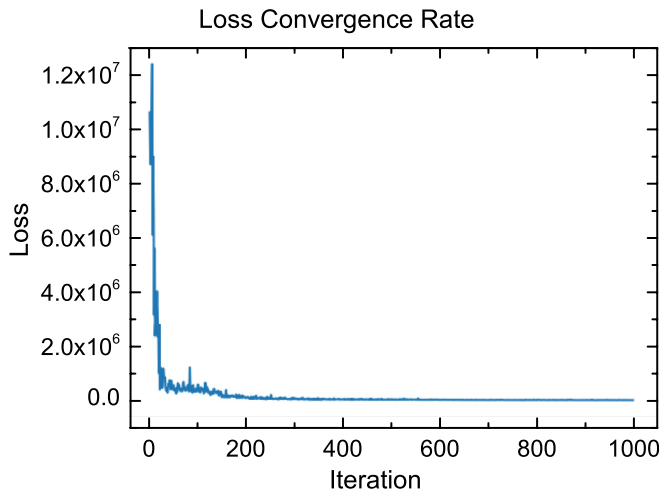
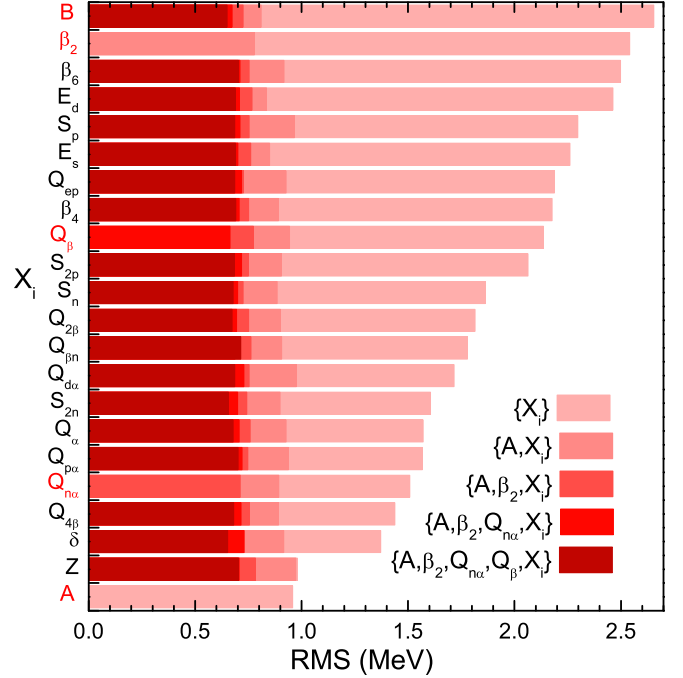FIG. 3. The value of loss changes with the number of iterations.

FIG. 4. Root-mean-square deviations between data and BNN predictions.

energies and mass or charge number. It is also consistent with the fact that the empirical formula $aA^{-1/3}$ is often used.

According to the smallest rms deviation, the mass number A is selected. Then the input ground-state properties are expanded to two-dimensional $X = \{A, X_i\}$. Their rms deviations are marked by $\{A, X_i\}$ in Fig. 4. It is found that adding one dimension in the input layer generally improves the BNN predictions compared to $X = \{A\}$. As is well known, the mass and charge numbers (or mass number and isospin asymmetry) are independent variables, which can be the identification of the nuclide. Thus, in the residual approach to study the nuclear charge radii [6], fission yields [8], and isotopic cross sections in spallation [9], the mass and charge numbers are applied as neurons of the input layer. This presupposition is not the best choice when using the BNN to predict the GDR energies. In Fig. 4, it is shown that the minimum rms deviation for the two-dimensional input is obtained by using $\{A, \beta_2\}$ as inputs, rather than $\{A, Z\}$ or $\{A, \delta\}$. The minimum rms deviation 0.78 MeV is 19% lower than that using one-dimensional input $\{A\}$. It is consistent with the quadrupole deformation effect found from Fig. 1. Based on $\{A, \beta_2\}$, the input data are added to three, four, and five dimensions in sequence. According to the minimum of the rms deviation, the input variables $Q_{n\alpha}$, Q_β, and B are selected. The optimal input variables to train the BNN for predicting GDR energies are $X = \{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$, and the rms deviation is 0.65 MeV.

The advantage of the BNN predictions using the optimal input $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$ compared to those using two-dimensional input $X = \{A, \delta\}$ can be found in Fig. 5. The 95% confidence intervals of predictions by the BNN approach with input $\{A, \delta\}$ are shown in Fig. 5(a) as a (blue) band. The global decrease of the data with increasing mass is
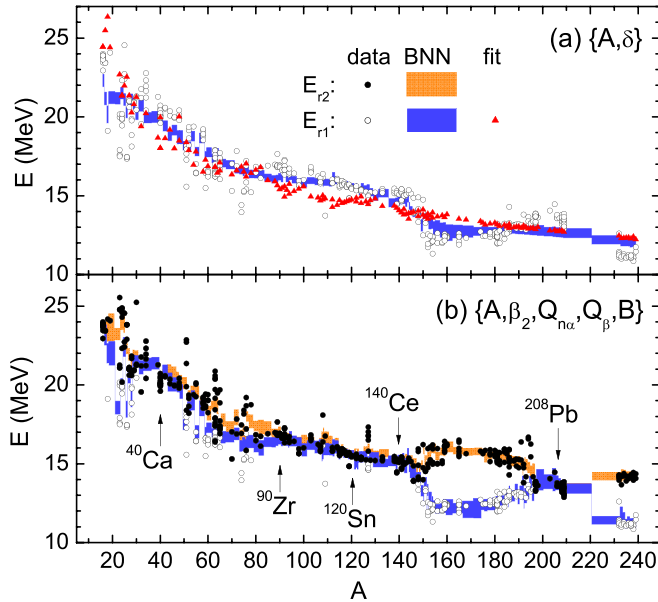
FIG. 5. Data of GDR energies compared to the predictions by BNN using (a) two-dimensional input layer $\{A, \delta\}$ and (b) five-dimensional input layer $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$. The fitting by the empirical formula $(a + b\delta)A^{-1/3}$ is also shown as (orange) triangles.

FIG. 6. Data of GDR energies in nuclei $^{124-160}$Nd compared to the predictions by BNN using (a) two-dimensional input layer $\{A, \delta\}$ and (b) five-dimensional input layer $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$. In order to distinguish high GDR energies E_{r2} from low ones E_{r1}, the data and predictions of high GDR energies are moved upward by 3 MeV. The predictions by the BNN are shown as shadows expressing the 95% confidence intervals.

reproduced. However, the data in the regions near $A = 150$, 200, and 238 deviate from global mass dependence. From $A = 175$ to 200, the value of E_{r1} even increases with increasing mass. Those features of the data cannot be reproduced by the BNN approach using input $\{A, \delta\}$.

The BNN predictions using the input $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$ are shown as (blue and red) bands in Fig. 5(b). In the figure, the arrows point out the cases of nuclei with magic number 20, 50, 82, and 126. It is found that, for the nuclei near the shell, the two GDR energies are the same, which means there is only a peak in the GDR spectrum. This feature is reproduced by the BNN predictions. In the region out of the shell, there are differences between the low and high GDR energies. Especially in the region $150 < A < 196$ ($60 < Z < 78$ and $90 < N < 118$), the differences between two GDR energies are nearly 4 MeV. After training by the data including the ground-state properties, the BNN predictions display the deformation effect.

It is worth comparing the predictions by the BNN and empirical formula with the same input variables, i.e., $X = \{A, \delta\}$. The empirical formula with parameters is often proposed by the physics motivated model or phenomenology. For example, based on the mass dependence, the empirical formula $aA^{-1/3}$ is used to fit the data of the GDR energies. In order to include higher order effects, such as the isospin effect, the empirical formula is extended to $(a + b\delta)A^{-1/3}$. By fitting the data of the low GDR energy E_{r1}, the parameters $a = 61.6 \pm 0.5$ and $b = 66.9 \pm 4.0$ are obtained. As shown in Fig. 5(a), the formula $(a + b\delta)A^{-1/3}$ reproduces the global decrease of data with increasing mass. But the detailed features of the data in the regions near $A = 150$, 200, and 238 cannot be reproduced.

The description and prediction of the empirical formula may be further improved by adding terms for the higher
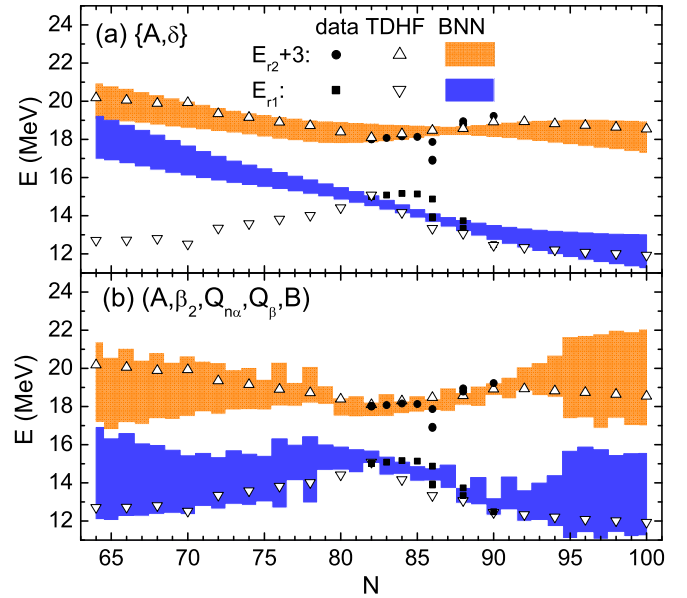
order effects, such as the shell effect or deformation effect. Nevertheless, it is not always straightforward as several prerequisites are needed. What is the next important effect? How can the effect be expressed? Is there any coupling with the known effects? In fact, GDR energies depend on the ground-state properties, the data of which are abundant and credible. The BNN approach provides a method for mining useful information from the data of ground-state properties without the above prerequisites and sorting the importance of the effects, as shown in Fig. 4. With this importance order, the effects may be further studied by the physics motivated model. That is to say, the physics motivated model is significant both for the empirical formula and the present BNN approach. It is used beforehand for the empirical formula but afterward for the BNN approach.

Figure 6 shows the low and high GDR energies in $^{124-160}$Nd as a function of neutron number. The Nd isotopes display the typical evolution of the quadrupole deformation [34]. Because of the magic number $N = 82$, the nucleus ^{142}Nd has a spherical shape. The quadrupole deformation parameter β_2 as a function of neutron number N shows that the ^{142}Nd nucleus has a minimum value 0.07, with gradual raises in two sides. The correlation between the GDR energy splitting $E_{r2} - E_{r1}$ and the deformation parameter β_2 has been found by the TDHF model in Ref. [22]. The TDHF calculations (up and down triangles in Fig. 6) show that two GDR energies are the same for ^{142}Nd, but gradually split in both sides of $N = 82$. This is the prior physics knowledge about the GDR energy splitting. Figure 6(a) shows the BNN predictions using input $\{A, \delta\}$ only. In the region from $N = 82$ to 90 where

the experimental data are available, the BNN provides good predictions for both E_{r1} and E_{r2}. The predictions by the BNN agree with the TDHF calculations for $N \geqslant 82$. However, one sees the overfitting in the region of $N < 82$. The GDR energy splitting $E_{r2} - E_{r1}$ by the BNN decreases with decreasing N, which is contrary to prior physics knowledge provided by the TDHF model.

The predictions by the BNN using the optimal ground-state properties $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$ as inputs are shown in Fig. 6(b). The selected input layer provides physics guidance in the BNN and hence actively contributes to reduce the risk of the numerical fluctuations. In the region with experimental data ($82 \leqslant N \leqslant 90$), the predictions are constrained and hence the 95% confidence intervals are small. Out of this region, the confidence intervals are quite large but cover the TDHF calculations in general. Concerning the prior physics knowledge about the GDR provided by the TDHF calculations, the overfitting by the BNN disappears after adding the ground-state properties in the input layer.

The data (circles and squares in Fig. 6) show that both the ^{142}Nd nucleus and the ^{144,145,146}Nd nuclei have the same low and high GDR energies. This is different from the TDHF calculations. In fact, the GDR energies are extracted from the ^{144,145,146}Nd(γ,sn) spectra. The ^{144,145,146}Nd(γ,sn) spectra (graphs 102–105 in Ref. [18]) display asymmetrical shapes, but were fitted by a single peak function. Both the TDHF calculations and BNN predictions indicate that two peaks fitted to the GDR spectra are necessary for ^{144,145,146}Nd nuclei. It should be interesting to train the BNN by the GDR spectra and not the GDR energies, which will be our future effort.

It is worth comparing the BNN to multitask neural networks (MNNs), which are used to describe GDR key parameters [42]. The advantage of the MNN is that it can exploit common information in the shared layer of different output variables. Also, it can save the number of parameters to be estimated compared to building multiple neural networks. Therefore, the MNN is suitable when the sample size is small. But, like the other variants of neural networks, it can only provide the predicted values without confidence intervals, while the BNN can measure the uncertainty of the predictions and provide confidence intervals. The confidence intervals are significant when the predictions are extended to the region far

from the measured data. Another difference exists in the input layer: the inputs are $\{Z, N, A, \beta_2\}$ in Ref. [42], while those in our paper are $\{A, \beta_2, Q_{n\alpha}, Q_\beta, B\}$. The same inputs A and β_2 are applied. What makes our paper novel is that we add the ground-state properties in the input layer, which can avoid overfitting. Considering the Bayesian statistics in the MNN is worthy and should be tried in future work.

IV. CONCLUSION

In conclusion, it has been proved in previous works [5–9] that one can include as much physics as possible in the initial prediction by using physics motivated models and use the BNN to fine tune these models by modeling the residuals. In this paper, a case study on the prediction of GDR energy is presented to illustrate the new method to provide physics guidance in the BNN from the input layer without the initial prediction by the physics motivated model. Microscopic theories such as Hartree-Fock-based random phase approximation and time-dependent Hartree-Fock theory have proved that the GDR energies depend strongly on the ground-state wave function [20,22]. Based on this dependence, it is proposed to predict the GDR energies by using the data of the nuclear properties in the ground state, which are abundant and credible. Pearson's correlation coefficients are applied to assess the statistical dependence between the ground-state properties and the GDR energies. Then the optimal ground-state properties are selected as neurons of the input layer in the BNN for predicting the GDR energies. It is shown that those selected ground-state properties provide physics guidance in the BNN and hence actively contribute to avoid overfitting. This paper gives a demonstration to find effects of the GDR energy by using the BNN without the physics motivated model, which may be helpful to discover physics effects from complex nuclear data.

[1] K. Gernoth, J. Clark, J. Prater, and H. Bohr, Phys. Lett. B **300**, 1 (1993).

[2] S. Athanassopoulos, E. Mavrommatis, K. Gernoth, and J. Clark, Nucl. Phys. A **743**, 222 (2004).

[3] N. J. Costiris, E. Mavrommatis, K. A. Gernoth, and J. W. Clark, Phys. Rev. C **80**, 044332 (2009).

[4] T. Bayram, S. Akkoyun, and S. O. Kara, Ann. Nucl. Energy **63**, 172 (2014).

[5] R. Utama, J. Piekarewicz, and H. B. Prosper, Phys. Rev. C **93**, 014311 (2016).

[6] R. Utama, W.-C. Chen, and J. Piekarewicz, J. Phys. G **43**, 114002 (2016).

[7] L. Neufcourt, Y. Cao, W. Nazarewicz, and F. Viens, Phys. Rev. C **98**, 034318 (2018).

[8] Z.-A. Wang, J. Pei, Y. Liu, and Y. Qiang, Phys. Rev. Lett. **123**, 122501 (2019).

[9] C.-W. Ma, D. Peng, H.-L. Wei, Z.-M. Niu, Y.-T. Wang, and R. Wada, Chin. Phys. C **44**, 014104 (2020).

[10] Z. Niu and H. Liang, Phys. Lett. B **778**, 48 (2018).

[11] G. C. Baldwin and G. S. Klaiber, Phys. Rev. **71**, 3 (1947).

[12] B. S. Ishkhanov and V. N. Orlin, Phys. Part. Nuclei **38**, 232 (2007).

[13] A. Bracco, E. Lanza, and A. Tamii, Prog. Part. Nucl. Phys. **106**, 360 (2019).

[14] S. Ceruti, F. Camera, A. Bracco, A. Mentana, R. Avigo, G. Benzoni, N. Blasi, G. Bocchi, S. Bottoni, S. Brambilla, F. C. L. Crespi, A. Giaz, S. Leoni, B. Million, A. I. Morales, R. Nicolini, L. Pellegri, S. Riboldi, O. Wieland, D. Bazzacco, M. Ciemala,

E. Farnea, A. Gottardo, M. Kmiecik, A. Maj, D. Mengoni, C. Michelagnoli, V. Modamio, D. Montanari, D. Napoli, F. Recchia, E. Sahin, C. Ur, J. J. Valiente-Dobón, B. Wasilewska, and M. Zieblinski, Phys. Rev. C **95**, 014312 (2017).

[15] D. Mondal, D. Pandit, S. Mukhopadhyay, S. Pal, S. Bhattacharya, A. De, N. D. Dang, N. Q. Hung, S. Bhattacharya, S. Bhattacharyya, B. Dey, P. Roy, K. Banerjee, and S. Banerjee, Phys. Lett. B **784**, 423 (2018).

[16] N. Kobayashi, K. Miki, T. Hashimoto, C. Iwamoto, A. Tamii, N. Aoi, M. P. Carpenter, K. Hatanaka, J. Isaak, E. Ideguchi, S. Morinobu, S. Nakamura, and S. Noji, Eur. Phys. J. A **55**, 231 (2019).

[17] Y. Tian, X. Tao, J. Wang, X. Ke, R. Xu, and Z. Ge, Chin. Phys. C **43**, 114102 (2019).

[18] V. Plujko, O. Gorbachenko, R. Capote, and P. Dimitriou, At. Data Nucl. Data Tables **123-124**, 1 (2018).

[19] N. Lyutorovich, V. Tselyaev, J. Speth, S. Krewald, F. Grümmer, and P.-G. Reinhard, Phys. Lett. B **749**, 292 (2015).

[20] G. Bonasera, M. R. Anders, and S. Shlomo, Phys. Rev. C **98**, 054316 (2018).

[21] S. Fracasso, E. B. Suckling, and P. D. Stevenson, Phys. Rev. C **86**, 044303 (2012).

[22] A. A. B. Mennana, Y. E. Bassem, and M. Oulne, Phys. Scr. **95**, 065301 (2020).

[23] R. Wang, Z. Zhang, L.-W. Chen, C. M. Ko, and Y.-G. Ma, Phys. Lett. B **807**, 135532 (2020).

[24] H.-Y. Kong, J. Xu, L.-W. Chen, B.-A. Li, and Y.-G. Ma, Phys. Rev. C **95**, 034324 (2017).

[25] W. B. He, Y. G. Ma, X. G. Cao, X. Z. Cai, and G. Q. Zhang, Phys. Rev. Lett. **113**, 032506 (2014).

[26] J. Su, Chin. Phys. C **43**, 064109 (2019).

[27] J. Su, L. Zhu, and C. Guo, Phys. Rev. C **101**, 044606 (2020).

[28] G. Colò, U. Garg, and H. Sagawa, Eur. Phys. J. A **50**, 26 (2014).

[29] D. R. Chakrabarty, N. D. Dang, and V. M. Datar, Eur. Phys. J. A **52**, 143 (2016).

[30] D. Pandit, S. Bhattacharya, D. Mondal, B. Dey, S. Mukhopadhyay, S. Pal, A. De, and S. R. Banerjee, Phys. Rev. C **99**, 024315 (2019).

[31] W. Huang, G. Audi, M. Wang, F. G. Kondev, S. Naimi, and X. Xu, Chin. Phys. C **41**, 030002 (2017).

[32] G. Audi, F. G. Kondev, M. Wang, W. Huang, and S. Naimi, Chin. Phys. C **41**, 030001 (2017).

[33] M. Wang, G. Audi, F. G. Kondev, W. Huang, S. Naimi, and X. Xu, Chin. Phys. C **41**, 030003 (2017).

[34] N. Wang, M. Liu, X. Wu, and J. Meng, Phys. Lett. B **734**, 215 (2014).

[35] P. Schober, C. Boer, and L. A. Schwarte, Anesthesia Analgesia **126**, 1763 (2018).

[36] D. J. C. MacKay, Neural Comput. **4**, 448 (1992).

[37] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University, New York, 1995).

[38] R. M. Neal, *Bayesian Learning for Neural Networks*, Vol. 118 (Springer Science & Business Media, New York, 2012).

[39] V. Mullachery, A. Khera, and A. Husain, arXiv:1801.07710v2.

[40] T. Salimans, D. Kingma, and M. Welling, in *Proceedings of the International Conference on Machine Learning*, 2015 (unpublished), pp. 1218–1226.

[41] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra, arXiv:1505.05424v2.

[42] J. Bai, Z. Niu, B. Sun, and Y. Niu, Phys. Lett. B **815**, 136147 (2021).