Letter

Learning correlations in nuclear masses using neural networks

Aman Sharma,* A. Gandhi, and Ajay Kumar†
Department of Physics, Banaras Hindu University, Varanasi 221005, India

(Received 29 January 2022; accepted 16 March 2022; published 29 March 2022)

There have been great improvements in the predictions of nuclear masses, yet it is difficult to exactly reproduce the measured nuclear mass. It has been suggested that the cause of such discrepancies is due to the negligence of many-body effects in the available theoretical models. The errors in the prediction of the nuclear mass show residual correlations due to the missing physics in the mass models. In the present Letter we have tried to learn such correlations by using the neural networks. We have used a neural network architecture which adaptively learns the linear and nonlinear correlations between the data of different fidelity. We have used the theoretical predictions of finite range droplet model and Hartree- Fock-Bogoliubov models in the input of the neural networks. The present approach show significant improvements in the accuracy of the predictions. It has been clearly presented that the difference between the predictions from the present approach and the experimental data behave more, such as white noise, showing that using the present approach the residual correlations arising due to the missing physics from the available mass models can be learned.

DOI: 10.1103/PhysRevC.105.L031306

Nuclear masses can be measured experimentally with the great precision because of the advances made in the field of nuclear mass spectroscopy. The accurate knowledge of the nuclear mass is very important in describing many nuclear processes. The accuracy in the nuclear mass has a direct impact on our understanding about the nuclear structure [1], nuclear effective interactions [2], and of nucleosynthesis [3]. Despite the great progress in measuring the nuclear mass [4,5], the theoretical models have to be used for predicting the nuclear masses in the region far from the stability [1]. Many theoretical models have been proposed over the years to predict nuclear masses. The mainly used nuclear mass models consist of macroscopic models (e.g., Bethe-Weizsäcker mass formula [6,7]), macroscopic-microscopic models (e.g., the finite-range droplet model (FRDM) [8] and the Weizsäcker-Skyrme model [9]) and microscopic models (e.g., Hartree-Fock-Bogoliubov (HFB) mass models [10–12]). Although accuracy of these models in known mass region vary slightly from each other, yet it is difficult, in general, to predict with accuracy better than $\approx$500 keV [1].

The deviation of FRDM predictions from measured nuclear mass show systematic dependence on the neutron and proton numbers. In different studies it has been confirmed that these deviations of model predictions from the measured data are correlated [13–19], and the strength of correlations decreases as we go from macroscopic models to macroscopic-microscopic models and microscopic model predictions [18]. Also the predictions of nuclear masses far from stability may differ by several MeVs for different mass models. The reason for such discrepancies and the residual correlations in the prediction error of the models have been attributed to the neglected many-body effects in the mass models and the chaotic motion in nuclei [14,18]. These discrepancies between mass models and measured mass can be reduced by incorporating more physical information in to the mass models by taking in to account the residual interactions or by using local information, such as Garvey-Kelson relations [14,18].

In the present Letter we have explored whether such correlations in the mass predictions can be reduced by using neural networks. Neural networks are a very powerful tool, and it has seen great advancements in recent years. It has been successfully used in a variety of applications. There have been many efforts in the direction to use the neural networks for nuclear mass predictions [20–24]. It is also observed in recent studies that by incorporating some physical features, such as nuclear pairing and shell effects in the input layer can improve the accuracy significantly [25]. Integrating physics with the machine learning methods can help in improving their performance and reliability, and recently many efforts have been put in this direction [26]. One such study uses multifidelity neural network strategy [27] to leverage on the low-fidelity (lf) data to produce better estimates for high-fidelity (hf) data. Although neural networks have emerged as powerful tools, yet to train it, a large amount of data of good accuracy is required. But, in practice, high-fidelity data are scarce and expensive to acquire as compared to the low-fidelity data. In recent studies the neural networks were trained on the difference between measured masses and predictions of different mass models. Then these learned differences were used with the existing model predictions to produce new estimates. In this Letter we have followed a different approach, instead of learning the residual between measured mass and model predictions, we have used multifidelity strategy as discussed in Ref. [27] to learn the correlations between the experi-

*aman.marley1314@gmail.com
†Corresponding author: ajaytyagi@bhu.ac.in

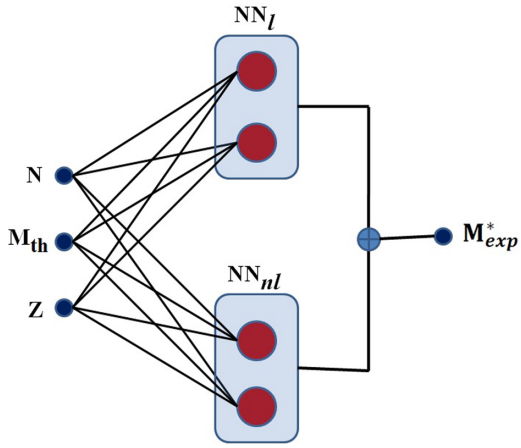

FIG. 1. Schematic of the neural network used in this Letter.

mentally measured nuclear masses and the theoretical model predictions. In this Letter, we have used HFB14 and FRDM95 mass excess predictions from Ref. [28] as the low-fidelity estimate and used the separate neural networks to learn the correlations between the model predictions and the experimental masses to produce the multifidelity approximations. For simplicity we will use only HFB and FRDM in place of HFB14 and FRDM95, respectively, in the rest of the Letter. The used neural network architecture adaptively learns the linear and nonlinear correlations between low- and high-fidelity data. The deviation of different mass model predictions from measured masses show residual correlations and the power spectrum of these fluctuations were found to follow the $1/f$ characteristic [18,29]. We have also calculated the power spectrum of the deviation of predictions from the present Letter with the experimental mass to check whether the present approach has reduced the correlations which are present in the existing model calculations.

In the present Letter, measured nuclear masses are labeled as high-fidelity data wherease mass predictions of theoretical models are labeled as low-fidelity data. In multifidelity modeling, it is of fundamental interest to construct the cross correlation between high- and low-fidelity data in order to produce better estimates for the data. Let the relation between high- and low-fidelity nuclear masses be given as

$$M_{\rm hf} = F_l(N, Z, M_{\rm lf}) + F_{\rm nl}(N, Z, M_{\rm lf}). \tag{1}$$

Here $M_{\rm hf}$ is the high-fidelity measured nuclear mass, N is the number of neutrons, Z is number of protons, and $M_{\rm lf}$ (predictions of theoretical models) is the low-fidelity data. In Eq. (1), F_l and $F_{\rm nl}$ represent linear and nonlinear functions, respectively, which are initially unknown. We have used the neural network architecture as presented in Fig. 1 to approximate both F_l and $F_{\rm nl}$. N, Z, and $M_{\rm th}$ (i.e., model predictions for nuclear mass excess) are used in the input layer. In Fig. 1 NN_l presents a neural network which is used to approximate the linear correlation function (F_l). In NN_l no hidden layer has been used because the linear correlation between high- and low-fidelity data is assumed easy to learn. In NN_l no activation function has been used, therefore, the output of the NN_l is a linear combination of the inputs. In order to approximate nonlinear correlation function ($F_{\rm nl}$), a feedforward neural network ($NN_{\rm nl}$) has been used. The outputs of $NN_{\rm nl}$ and NN_l are combined to give an estimate of $M_{\rm exp}$. Since feedforward neural networks are prone to the overfitting problem, therefore, to overcome the problem of overfitting, we have used L_2 regularization. It prevents the overfitting by adding a sum of the square of weights, which is a measure of the complexity of the model. The weights and biases of NN_l and $NN_{\rm nl}$ are learned by minimizing the following loss function:

$$\text{Loss} = \frac{1}{N_h}\sum_{i=1}^{N_h}(|M^*_{\rm exp} - M_{\rm exp}|)^2 + \lambda\sum_k w_k^2, \tag{2}$$

here N_h is total number of high-fidelity data points used in the training, $M^*_{\rm exp}$ is the estimate produced by the model, $M_{\rm exp}$ is the target value, λ is the regularization rate and w's represent the weights involved in $NN_{\rm nl}$. The regularization has been used only for $NN_{\rm nl}$ so that the present neural network can tend towards the linear correlations between the high- and low-fidelity nuclear masses. Weights were initialized by Xavier's initialization method [30], whereas the biases were initialized by zero. To minimize the loss, we have used the Adam optimization algorithm [31], and it has been used extensively in many fields of application. We have used two separate neural networks MF1 and MF2 to improve the predictions of FRDM and HFB model predictions respectively. $NN_{\rm nl}$ in both MF1 and MF2 consists two hidden layers with 50 neurons each with the hyperbolic tangent (tanh) function as the activation function. Initial learning rate of the optimizer was set to 0.001 and regularization rate (λ) was set to 0.0005 during the training of MF1 and MF2. The experimental mass excess values were taken from Audi *et al.* [32], although more than 3000 experimental masses are present in these compilations but we have used only those with experimental error less than or equal to 100 keV. Also mass excess values without any uncertainty estimate were also excluded, and the remaining 1919 experimental mass values were used in the training process. Out of these, randomly chosen 1534 nuclear mass values were used for training the neural network whereas rest 385 masses were used for testing the performance of the neural network. Corresponding FRDM and HFB estimates of mass excess of 1919 nuclei were used from the RIPLE-3 [28] library and were used as the low-fidelity estimate in the input of the neural networks.

The neural networks MF1 and MF2 were trained until the losses on testing and training set converged. The training was performed by using all the 1534 nuclei in a single batch. The root-mean-square (rms) deviation of mass excess from FRDM and HFB calculations with respect to the experimental mass were initially 573 and 679 keV, respectively, for the 1919 nuclei and after the training were 213 and 348 keV. Significant reduction in the rms values have been observed. Initial rms deviations for 1534 nuclei in the training set of FRDM and HFB predictions were 578 and 685 keV and for the testing set were 553 and 649 keV, respectively. After the completion of training the rms deviation corresponding to the training set were 201 and 336 keV for MF1 and MF2 predictions, respectively. Similarly, losses on the testing set were 259 and 382 keV, respectively, for MF1 and MF2. A reduction of rms deviation

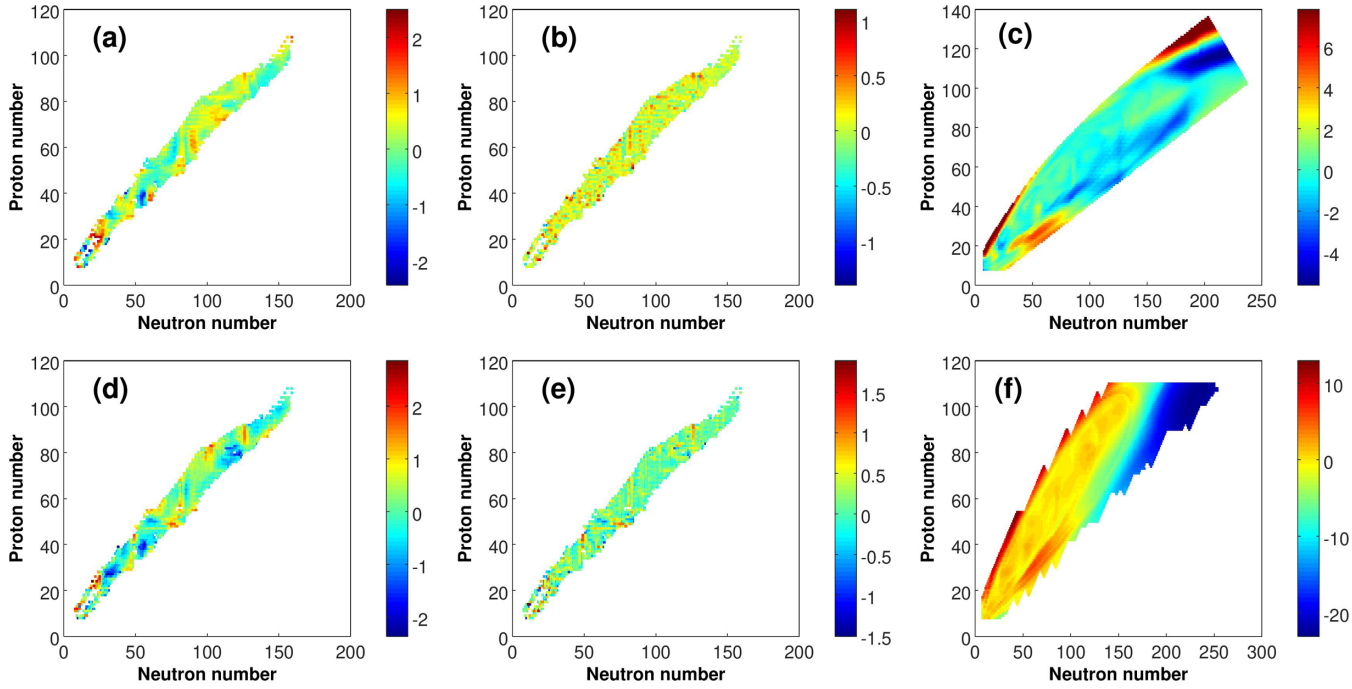

FIG. 2. Difference of (a) FRDM predictions with experimental mass excess, (b) MF1 predictions with the experimental mass excess, (c) FRDM predictions with the predictions of MF1, (d) HFB predictions with experimental mass excess, (e) MF2 predictions with the experimental mass excess, and (f) HFB predictions with the predictions of MF2.

of 65% and 53% has been observed for the training and testing sets, respectively in MF1. In MF2, a reduction of rms deviation of 51% and 41% has been observed for the training and testing sets, respectively, in MF2. These values indicate that the neural network has learned the correlations between model predictions and experimental mass excess with good generalization.

For more clarity, differences of FRDM and HFB predictions with experimental mass excess has been presented in Figs. 2(a) and 2(d), respectively. Regions with large differences can be observed in these figures, and such systematic discrepancies can be attributed to the missing physical information in the FRDM and HFB models. The motive behind the present Letter is that the neural network can complement the FRDM and HFB models by learning the correlations of FRDM and HFB predictions with experimental mass and by making them accountable for the contributions of the missing physics in these two models. In Figs. 2(b) and 2(e) differences of MF1 and MF2 predictions with the experimental mass excess are presented, respectively. It is observed from Figs. 2(b) and 2(e) that the regions of large deviation, such as in Figs. 2(a) and 2(d) are less pronounced. It is clear that MF1 and MF2 predicts with better accuracy than FRDM and HFB in known mass region. For completeness of the data, we have also presented the differences of FRDM with MF1 predictions, and HFB with MF2 predictions in Figs. 2(c) and 2(f), respectively. It is observed from these figures that MF1 and MF2 suggest large corrections to FRDM and HFM models for the regions far from the valley of stability. We have also presented the mass differences of experimental values with MF1, FRDM, MF2 and HFB predictions for the isotopes of Ba in Fig. 3. Mass excess for isotopes with neutron numbers 66 to 77 and 84 to 91 were used during the training process (i.e., were included in the training and testing set). It is clear from the Fig. 3 that the present approach is able to give reliable predictions in he neighborhood of measured mass region of the nuclear chart. In order to study the statistical characteristic of the deviation of different predictions from the experimental values, all the nuclei used in the training process were arranged in a one-dimensional boustrophedon single list [33]. All nuclei were arranged in

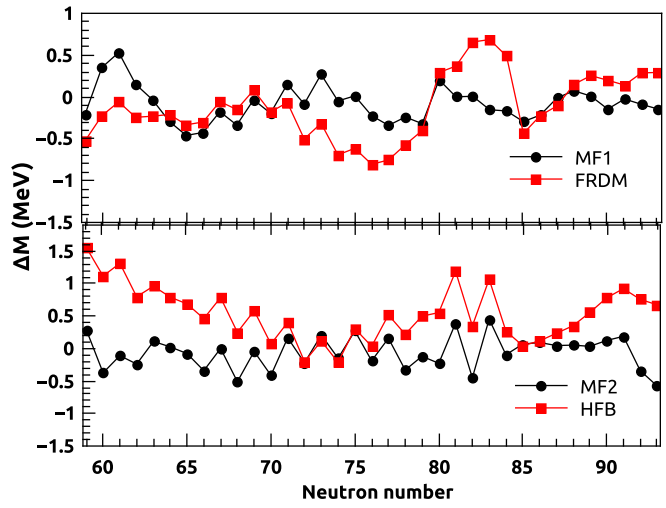

FIG. 3. Comparison of FRDM and MF1, HFB and MF2 predictions for the isotopes of Ba.

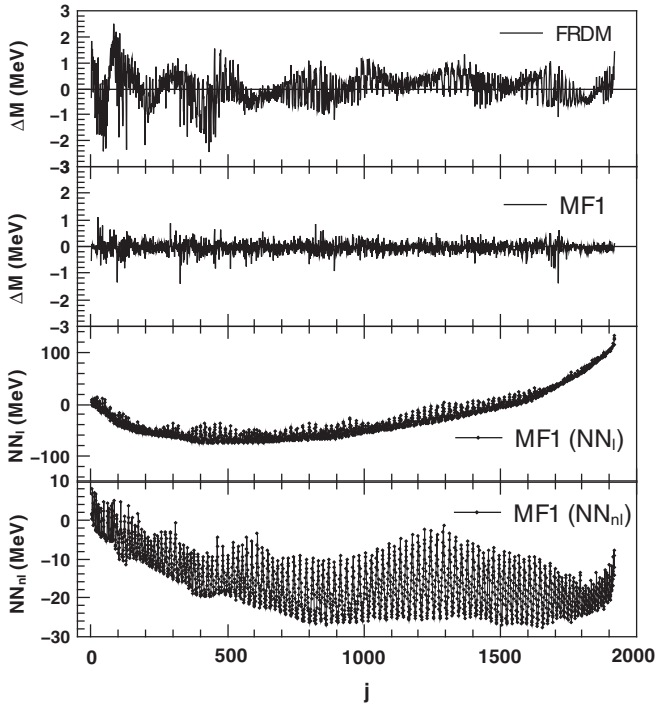

FIG. 4. Mass differences of experimental data with FRDM and MF1 predictions, MF1 (NN_l) and MF1 ($NN_{\rm nl}$) outputs with order number j along the x axis.

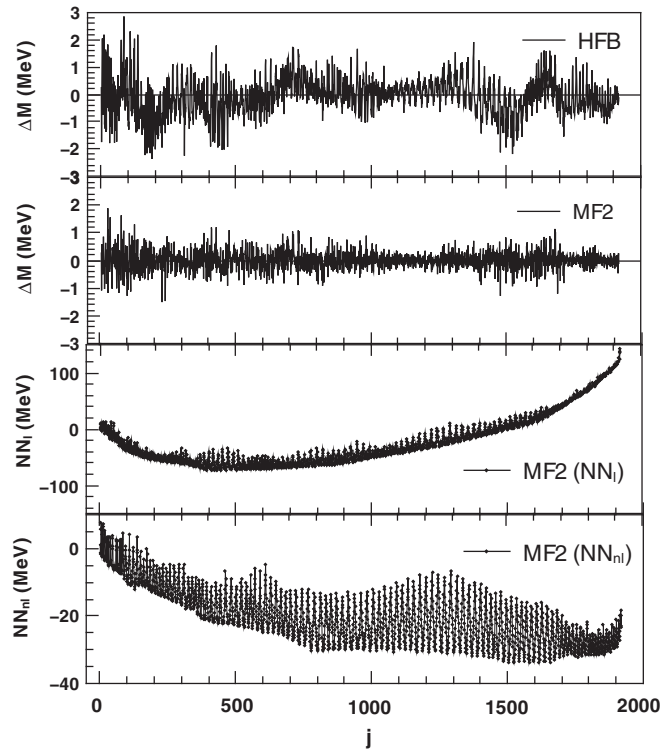

FIG. 5. Mass differences between experimental data and HFB and MF2 predictions, MF2 (NN_l) and MF2 ($NN_{\rm nl}$) outputs with order number j along the x axis.

increasing mass number order such that odd-A nuclei were arranged in decreasing N-Z value and even-A nuclei were arranged in increasing N-Z values. Based on this ordering scheme each nucleus is assigned a distinct order number (j). The deviation of FRDM, HFB, MF1, and MF2 predictions with experimental values is presented in Figs. 4 and 5. From Figs. 4 and 5 it is clear that significant correlations are present in the deviation of FRDM and HFB predictions from the experimental data, whereas the discrepancies in the MF1 and MF2 predictions with respect to experimental mass are highly fluctuating but no significant systematic can be observed. In Figs. 4 and 5, we have also presented the predictions of MF1 (NN_l), MF1 ($NN_{\rm nl}$), MF2 (NN_l). and MF2 ($NN_{\rm nl}$), which provide further insight in to the linear and nonlinear functions learned by the present approach. Using the optimized weight and bias values, the output of NN_l for the MF1 and MF2 can be written as $-0.781Z + 0.695N + 0.876M_{\rm th}^{\rm FRDM} - 0.616$, and $-0.772Z + 0.740N + 0.885M_{\rm th}^{\rm HFB} - 0.568$, respectively. It can be observed that, although MF1 and MF2 use different mass model predictions in the input layer, yet the outputs of the NN_l and $NN_{\rm nl}$ follow the similar systematic. The average ratio between the outputs of NN_l and $NN_{\rm nl}$ were observed to be 2.45 and 1.70, respectively. Now since we have arranged the nuclei in a one-dimensional list, we have calculated the discrete Fourier transform by using Eq. (3),

$$F_k = \frac{1}{\sqrt{N_j}} \sum_j \frac{M_j^{\rm exp} - M_j^{\rm th}}{\sigma_{\rm rms}} \exp\left(\frac{-2\pi i j k}{N_j}\right). \qquad (3)$$

The plot between $\ln(|F_k|^2)$ and $\ln(k/N_j)$ is presented in Figs. 6 and 7, here N_j represents the total number of nuclei. Then a linear fit has been performed, where the slope (m) of such a fit corresponds to the the m value in the power law of form $|F_k|^2 \approx \omega^m$ [18]. Systems exhibiting quantum chaotic motion, show $m \approx -1$ behavior which is present in the FRDM and HFB predictions as shown in Figs. 6 and 7. The value of slope in Figs. 6 and 7, corresponding to FRDM

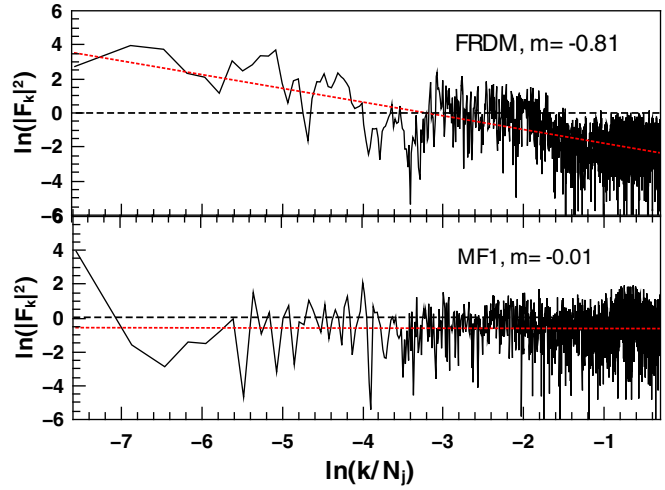

FIG. 6. Plot of the natural logarithm of the square of the amplitude of the Fourier transform of the prediction errors (as presented in Fig. 4) with the natural logarithm of the frequencies ($\omega = k/N_j$) along the x axis.

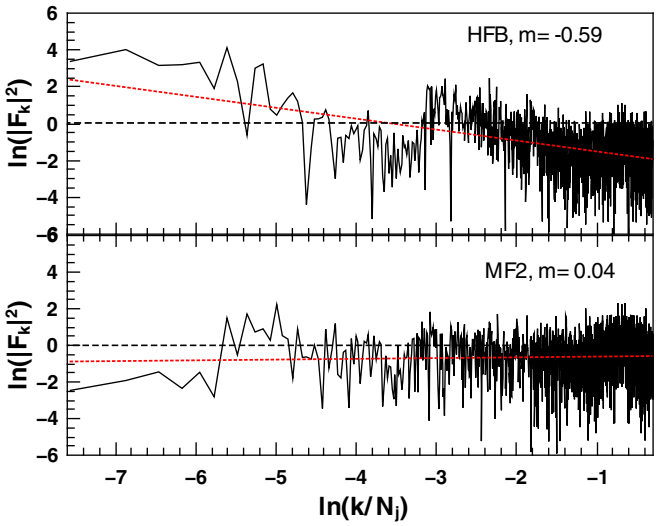

FIG. 7. Plot of the natural logarithm of the square of the amplitude of the Fourier transform of the prediction errors (as presented in Fig. 5) with the natural logarithm of the frequencies ($\omega = k/N_j$) along the x axis.

and HFB predictions were observed to be $m = -0.81 \pm 0.04$ and $m = -0.59 \pm 0.04$, respectively. For the MF1 and MF2 predictions the correlations are negligible as the power spectrum has $m = -0.01 \pm 0.03$ and $m = 0.04 \pm 0.03$, respectively. Hence, the fluctuations in MF1 and MF2 predictions show white-noise type of behavior. This indicates that the proposed neural networks are capable of learning the residual correlations in the mass model predictions.

It is observed in this Letter that multifidelity framework learns the cross correlations between FRDM, HFB predictions and experimental mass excess values. We have also observed that overall accuracy of the predictions are greatly improved, rms deviations for FRDM and HFB predictions have been reduced from 573 and 678 keV to 213 and 348 keV, respectively. The slope of the power spectrum of the deviations of the MF1 and MF2 is very small close to the white-noise-like behavior, which clearly indicates that the present approach is capable of learning the correlations arising due to the neglected many-body effects in the available mass models.

A.K. thanks the SERB-DST, Government of India (Sanction No. CRG/2019/000360), UGC-DAE Consortium for Scientific Research (Sanction No. UGC-DAE-CSR-KC/CRS/19/NP03/0913), and Institutions of Eminence (IoE) BHU [Grant No. 6031] for financial support of this work.

[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] E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle, Rev. Mod. Phys. **29**, 547 (1957).
[4] R. S. Van Dyck, S. L. Zafonte, S. Van Liew, D. B. Pinegar, and P. B. Schwinberg, Phys. Rev. Lett. **92**, 220802 (2004).
[5] M. Mukherjee, A. Kellerbauer, D. Beck, K. Blaum, G. Bollen, F. Carrel, P. Delahaye, J. Dilling, S. George, C. Guénaut, F. Herfurth, A. Herlert, H.-J. Kluge, U. Köster, D. Lunney, S. Schwarz, L. Schweikhard, and C. Yazidjian, Phys. Rev. Lett. **93**, 150801 (2004).
[6] H. A. Bethe and R. F. Bacher, Rev. Mod. Phys. **8**, 82 (1936).
[7] C. F. Von Weizsäcker, Z. Phys. **96**, 431 (1935).
[8] P. Möller, W. D. Myers, H. Sagawa, and S. Yoshida, Phys. Rev. Lett. **108**, 052501 (2012).
[9] N. Wang, M. Liu, X. Z. Wu, and J. Meng, Phys. Lett. B **734**, 215 (2014).
[10] S. Goriely, S. Hilaire, M. Girod, and S. Péru, Phys. Rev. Lett. **102**, 242501 (2009).
[11] S. Goriely, N. Chamel, and J. M. Pearson, Phys. Rev. C **93**, 034337 (2016).
[12] S. Goriely, N. Chamel, and J. M. Pearson, Phys. Rev. Lett. **102**, 152503 (2009).
[13] O. Bohigas and P. Leboeuf, Phys. Rev. Lett. **88**, 092502 (2002).
[14] A. Molinari and H. A. Weidenmuller, Phys. Lett. B **637**, 48 (2006).
[15] J. G. Hirsch, V. Velázquez, and A. Frank, Phys. Lett. B **595**, 231 (2004).
[16] J. G. Hirsch, A. Frank, and V. Velázquez, Phys. Rev. C **69**, 037304 (2004).
[17] H. Olofsson, S. Åberg, O. Bohigas, and P. Leboeuf, Phys. Rev. Lett. **96**, 042502 (2006).
[18] J. Barea, A. Frank, J. G. Hirsch, and P. Van Isacker, Phys. Rev. Lett. **94**, 102501 (2005).
[19] J. Barea, A. Frank, J. G. Hirsch, P. Van Isacker, S. Pittel, and V. Velázquez, Phys. Rev. C **77**, 041304(R) (2008).
[20] S. Gazula, J. W. Clark, and H. Bohr, Nucl. Phys. A **540**, 1 (1992).
[21] S. Athanassopoulos, E. Mavrommatis, K. A. Gernoth, and J. W. Clark, Nucl. Phys. A **743**, 222 (2004).
[22] K. A. Gernoth, J. W. Clark, J. S. Prater, and H. Bohr, Phys. Lett. B **300**, 1 (1993).
[23] H. F. Zhang, L. H. Wang, J. P. Yin, P. H. Chen, and H. F. Zhang, J. Phys. G: Nucl. Part. Phys. **44**, 045110 (2017).
[24] R. Utama, J. Piekarewicz, and H. B. Prosper, Phys. Rev. C **93**, 014311 (2016).
[25] Z. M. Niu and H. Z. Liang, Phys. Lett. B **778**, 48 (2018).
[26] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Nat. Rev. Phys. **3**, 422 (2021).
[27] X. Meng and G. Em Karniadakis, J. Comput. Phys. **401**, 109020 (2020).
[28] R. Capote *et al.*, Nucl. Data Sheets **110**, 3107 (2009).
[29] E. Faleiro, J. M. G. Gómez, R. A. Molina, L. Muñoz, A. Relaño, and J. Retamosa, Phys. Rev. Lett. **93**, 244101 (2004).
[30] X. Glorot and Y. Bengio, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010* (MLR, Phoenix, AZ, 2010), pp. 249–256.
[31] D. P. Kingma and J. Ba, arXiv:1412.6980.
[32] G. Audi, A. H. Wapstra, and C. Thibault, Nucl. Phys. A **729**, 337 (2003).
[33] J. G. Hirsch, A. Frank, J. Barea, P. Van Isacher, and V. Velázquez, Eur. Phys. J. A **25**, 75 (2005).