Bayesian evaluation of energy dependent neutron induced fission yields*

Ming Xiang Xiao (肖名翔)† Xiao-Jun Bao (包小军)‡ Zheng Wei (韦峥)‡ Ze En Yao (姚泽恩)
1Department of Physics, Hunan Normal University, Changsha 410081, People's Republic of China
2School of Nuclear Science and Technology, Lanzhou University, Lanzhou 730000, China

Abstract: From both fundamental and applied points of view, fragment mass distributions of fission are important observables. We apply the Bayesian neural network (BNN) approach to learn existing neutron induced fission yields and predict unknowns with uncertainty quantification. By comparing the predicted results and the experimental data, the BNN evaluation results are quite satisfactory on distribution positions and energy dependencies of fission yields. Predictions have been made for the fragment mass distributions of some actinides, presuming that this might help for future experiments.

Keywords: Bayesian evaluation, Energy dependent, Neutron induced fission yields

I. INTRODUCTION

Fission fragment mass and charge yields has a wide range of applications in various fields, ranging from understanding of the $r$-process of nucleosynthesis in astrophysical physics to reactor operations. But fission is a very complicated dynamical process of quantum many body system, for such a large-scale collective motion and producing hundreds of isotopes characterized by different charge and mass yields and kinetic energies [1–5], many challenges remain to be fully understood in its mechanism [5–10].

The experimental distributions of fission yields are rather complex. It has been observed that the yield distribution of nuclear fission has the structure of single peak, double peak and triple peak [5,11–14]. Inspired with these abundant experimental data, many approaches have been developed to analyze the observed data and understand the fission mechanism [15–24]. There are mainly two methods to calculate the fission yield distribution: dynamical approach [25–29] and statistical approach [30–33]. Above two approaches require potential energy surface as input.

The macroscopic-microscopic model [3,4,24,34] and microscopic model [13,35] have been widely used to calculate potential energy surface of fission process. However, the calculation of the multidimensional potential energy surface for a heavy nuclear system is a complicated physical problem, which has not yet been fully solved. It is because the proper choice of common degrees of freedom is an important and difficult task. The number of degrees of freedom should not be too large in order to allow the numerical analysis of the corresponding of dynamical equations. This means that the calculated potential energy surface is the upper limit of the optimum potential that the nucleus can adopt [24].

It is important to note the strong dependence of microscopic effects on the temperature (or, excitation energy) of the nucleus. At higher energies, temperature effects should be included not only in the potential energy surface [36–40], but also in in the friction tensor and in the mass tensor in the dynamical equations. However, this is rarely done. In addition, to describe yields of compound nucleus at higher excitation energies, it is necessary to take multi-chance fission into account [41]. But the estimates of the percentages of multi-chance fission contributions are difficult and have considerable uncertainties [41,42]. Therefore, it is a challenge both experimentally and theoretically to obtain accurate and complete energy dependent the distributions of fission yields for any fissile nucleus ($^{235}X$).

An alternative class of approaches being actively explored is based on machine learning techniques. While applications of machine learning approaches to the many body problem in condensed matter, quantum chemistry, and quantum information have been proliferating in the past few years, there are also significantly more machine learning applications in low-energy nuclear physics, such as the neutron induced fission yields distribution [43–48], the nuclear masses [49–51] and various nuclear reactions

Received 16 June 2023; Accepted 8 September 2023

* Supported by the National Natural Science Foundation of China (Grants No.12175064, U2167203). Hunan Outstanding Youth Science Foundation (2022JJ10031).
† E-mail: baoxiaojun@hunnu.edu.cn
‡ E-mail: weizheng@lzu.edu.cn
©2023 Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd
and structures properties [52–57]. The main objective of this work is to evaluate energy dependent neutron induced fission yields based on the Bayesian neural networks (BNN) approach.

II. THEORETICAL FRAMEWORK

The principle of Bayes’ rule is to establish a posterior distribution from all unknown parameters trained by a given data sample [58]. A detailed description of the origins and development of Bayesian neural networks (BNNs) goes beyond the scope of this paper, we limit ourselves to highlight the main features of the approach.

The BNN approach to statistical inference is based on Bayes’s theorem, which provides a connection between a given hypothesis (in terms of problem-specific beliefs for a set of parameters \( \omega \)) and a set of data \( (x, t) \) to a posterior probability \( p(\omega | x, t) \) that is used to make predictions on new inputs. In this context Bayes’s theorem may be written as [59]

\[
p(\omega | x, t) = \frac{p(x, t | \omega)p(\omega)}{p(x, t)} \tag{1}
\]

where \( p(\omega) \) is the prior distribution of the model parameters \( \omega \) and \( p(x, t | \omega) \) is the “likelihood” that a given model \( \omega \) describes the new evidence \( t(x) \). The product of the prior and the likelihood form the posterior distribution \( p(\omega | x, t) \) that encodes the probability that a given model describes the data \( t(x) \). The \( p(x, t) \) is a normalization factor which ensures the integral of posterior distribution is 1. In essence, the posterior represents the improvement to \( p(\omega) \) as a result of the new evidence \( p(x, t | \omega) \).

In the present work, the inputs of the network are given by \( x_i = [Z_i, N_i, A_i, E_i] \), which include \( Z_i \) (the number of charges of the fission nucleus), \( N_i \) (the number of neutrons in the fission nucleus), \( A_i \) (the mass number of fission fragments) and \( E_i = e_i + S_i \) (neutron incident energy plus neutron separation energy).

In our work, the neural network function \( f(x, \omega) \) adopted here has the following “sigmoid” form [60,61]:

\[
f(x, \omega) = a + \sum_{j=1}^{H} b_j \tanh \left( c_j + \sum_{i=1}^{l} d_{ij} x_i \right) \tag{2}
\]

where the model parameters are collectively given by \( \omega = [a, b_j, c_j, d_{ij}] \). \( H \) is the number of neurons in the hidden layer, and \( l \) is the number of inputs. \( a \) bias of output layers, \( b_j \) are the weights of output layers, \( c_j \) bias of hidden layers, and \( d_{ij} \) weights of hidden layers. The tanh is a common form of the sigmoid activation function that controls the firing of the artificial neurons. A schematic diagram of a neural network with a single hidden layer, three hidden neurons \( (H=3) \), and two input variables \( (l=2) \) is shown in Fig. 1.

In determining the optimal number of hidden layers and neurons in a neural network, we employ the technique of cross-validation. Firstly, we estimate a rough range for the model based on prior research papers. Subsequently, we randomly partition our dataset into two parts, one for training and the other for testing. We repeat this process, swapping the roles of the training and testing sets each time. We then compute the average prediction errors obtained from training the model twice with each set to estimate the performance of our Bayesian neural network model. Based on our analysis, we determine that the optimal configuration with double hidden layers, each consisting of 16 neurons.

Bayesian inference for neural networks calculates the posterior distribution of the weights given the training data \( p(\omega | x, t) \). This distribution answers predictive queries about unseen data by taking expectations. Each possible configuration of the weights, weighted according to the posterior distribution, makes a prediction about the unknown label given the test data item \( x \) [63].

In the present work, a variational approximation to the Bayesian posterior distribution on the weights is adopted. To obtain an estimate of the maximum a posteriori of \( \omega \), we use variational learning to find a probability \( q(\omega | \theta_\omega) \) to approximate the posterior probability of \( \omega \). To do this we use Kullback-Leibler divergence to minimize the distance between probabilities \( q(\omega | \theta_\omega) \) and \( p(\omega | x, t) \) [64–67]

\[
\theta^* = \arg \min_{\theta_\omega} KL [q(\omega | \theta_\omega) || p(\omega | x, t)]
\]

\[
= \arg \min_{\theta_\omega} \int q(\omega | \theta_\omega) \ln \frac{q(\omega | \theta_\omega)}{p(x, t | \omega)p(\omega)} d\omega
\]

\[
= \arg \min_{\theta_\omega} KL [q(\omega | \theta_\omega) || p(\omega)] - E_{q(\omega | \theta_\omega)} [\ln p(x, t | \omega)]
\]

And then the loss function of BNN can be written in the following form:

Fig. 1. (color online) A schematic diagram of a neural network with a single hidden layer, three hidden neurons \( (H = 3) \) and two input variables \( (l = 2) \).
\[ F(x, t, \theta_\omega) = KL \left[ q(\omega | \theta_\omega) \| p(\omega) \right] - E_{q(\omega | \theta_\omega)} \left[ \ln p(x, t | \omega) \right] \]

Finally, using the Monte Carlo method, we can get an approximate result.

**III. NUMERICAL RESULTS AND DISCUSSIONS**

In the process of running the program, we found that the function is often trapped in a local minimum point. And for this reason the results obtained by each program are always very different, especially when there is no experimental data for extrapolation, we have no way to judge which results are correct. Therefore, we use a method similar to annealing evolution algorithm. At each iteration, the program has a certain probability to accept a solution that is worse than the current one, so it is likely to jump out of the local optimal solution and reach the global optimal solution. After adding this mechanism, the effect is significant, and the difference between the results after each program execution is within an acceptable range.

In the absence of simulated annealing, iterative processes may lead to outcomes characterized by local optima, as shown in Fig. 2. This occurrence can be attributed to the iterative nature of the algorithm becoming trapped in suboptimal solutions. Conversely, the utilization of simulated annealing in neural networks mitigates this concern by effectively navigating the search space and avoiding convergence to suboptimal solutions.

To test the performance of the BNN approach, the independent mass distributions of the neutron induced fission of \( n + 2X \) with a neutron incident energy are studied. The training dataset includes 227Th(0.0253), 229Th(0.0253), 232Th(0.5), 232Th(14), 227Th(22.5), 231Pa(0.5), 232U(0.0253), 233U(0.0253), 231U(0.5), 233U(14), 234U(0.5), 234U(14), 235U(0.0253), 235U(0.5), 235U(5.04), 233U(14), 235U(15.5), 236U(0.5), 236U(14), 236U(20), 237U(0.5), 238U(0.5), 238U(1.6), 238U(2), 238U(3), 238U(4.5), 238U(5.5), 238U(5.8), 238U(10), 238U(10.05), 238U(13), 238U(14), 238U(16.5), 238U(20), 238U(22.5), 237Np(0.0253), 237Np(0.5), 237Np(1), 237Np(2), 237Np(4), 237Np(5.5), 237Np(14), 238Np(0.5), 238Pu(0.5), 240Pu(0.0253), 240Pu(0.5), 232Pu(0.5), 241Pu(0.0253), 241Pu(0.5), 242Pu(0.5), 242Pu(14), 241Am(0.0253), 241Am(0.5), 243Am(0.5), 242Cm(0.5), 243Cm(0.0253), 241Cm(0.5), 244Cm(0.5), 245Cm(0.0253), 246Cm(0.5), 248Cm(0.5), 251Cr(0.0253), 254Es(0.0253), 254Fm(0.0253), 255Fm(0.0253) is taken from Refs. [24, 68–70]. This is the neutron energy in parentheses.

We apply BNN to learn the existing evaluated distributions of mass yields, which includes 6996 data points of the neutron induced fissions of 66 nuclei. In order to examine the validity of the BNN approach, we separate the entire data set into the learning set and the validation set with a combination of 65+1. The learning set is built by randomly selecting 65 nuclei from the entire set, and the remaining 1 nucleus compose the validation set. The evaluation results about randomly selecting validation set (8 nuclei) are shown in Fig. 3–4. It can be seen from Fig. 3–4 that the evaluations of BNN are satisfactory, the position of distributions can be well described by the BNN approach. Note that the evaluations are less satisfactory around 254Fm, where neighboring nuclei in the learning set are not sufficient. This indicates the present BNN approach has a good ability for existing evaluated distributions of mass yields and there is no overfitting even only with the data in the learning set.

Next, we analyzed the results based on the data of 255Fm(0.0253) have been excluded in the learning set. In this case, with 6890 points, the total \( \chi^2 = \sum (I_o - f(x))^2 / N \) is 2.24 \times 10^{-5}. Fig. 5-6 shows the calculated results of mass distributions of 238U(n,f) at different neutron energies, which are compared with experimental data. One can see that the training results are in agreement with the experimental data.
good agreement with experimental data at the neutron energy range from 0.50 MeV to 22.50 MeV; the peak-to-valley ratio decreases with increasing energy. The BNN can remarkably reproduce the overall evaluated fission yields.

To study the influence of excitation energy on the shape of the mass distribution, we calculate $^{238}\text{U}$ at incident neutron energies of 1.6, 5.8, 16.5 and 22.5 MeV as shown in Fig. 7. There are obvious changes such as the increase in valley height as well as the decrease in peak height in the mass distributions with increasing incident neutron energy. It is known that the symmetric fission mode will play a role as excitation energies increase. Our training results demonstrated that the features of energy-dependent fission yields can be successfully described by the BNN evaluation.

The main motivation of our BNN approach is to estimate incomplete experimental the distribution of fission yields based on the information learned from completed evaluations of other nuclei. It is well known that only a few experimental data points are available for some nuclei. Fig. 8(a-c) shows that the calculated results using BNN for fission yields of $n+^{238}\text{U}$ at energies of 3.60, 4.49 and 14.8 MeV. The calculated mass distributions of the fragments are denoted by solid lines. A comparison of the experimental data with the results evaluated by using BNN, taking into account the experimental data, it can be seen that the BNN predictions without learning the experimental data are satisfactory. It is demonstrated that our BNN method shows reliable capacities in predicting the fission fragment mass distributions for interpolation.

As described above, the neutron energy ranges from 0.053 to 22.5 MeV in the training and test sets. To explore the estimation ability of our BNN at higher excitation energies, the evaluation of $^{238}\text{U}$ (n,f) at energies of 32.8 and 59.9 MeV are also shown in Fig. 8(d-e). The calculated mass distributions of the fragments are denoted by solid lines. A comparison of the experimental data...
data with the results evaluated by using BNN, we find that the BNN can give a rather reasonable evaluation of fission yields at energy of 32.8 MeV. However, the predictions of fission yields are less satisfactory at energy of 59.9 MeV.

One noticed that the evaluation of the fission fragment mass distributions for $^{238}$U(n,f) reaction at different incident-neutron-energies. The shadow region corresponds to the CI estimated at 95%. The experimental data are taken from Refs.[24,68−70].

It would be desirable to show some predictions of the dependence of the shape of the mass distribution on the neutron energy. Predictions have been made for the dependence of the shape of the mass distribution for the Th(n,f) and Pu(n,f) reactions at different incident-neutron-energies, the results are shown in Fig. 9. It should be clearly indicated that the asymmetry of distributions were predicted even at very high excitation energies (about 40 MeV) of fissioning nuclei, presuming that this might help for future experiments.

**IV. CONCLUSIONS**

The BNN method was applied to estimate the fission fragment mass distributions in neutron induced fission with incident energies from 0.0253 MeV to 32.8 MeV. The input for BNN includes the incident neutron energy, the neutron and charge numbers of the fission nucleus, the mass numbers of the fragment. The accumulation of more experimental data in the future will help test the predictive power of the present BNN method. In addition, one noticed that both experimental measurements and results of BNN, it should be clearly indicated that the asymmetry of distributions persists even at very high excitation energies (about 50-60 MeV) of fissioning nuclei.

References
