A Bayesian-neural-network prediction for fragment production in proton induced spallation reaction

Chun-Wang Ma(马春旺) 1,2,1) Dan Peng(彭丹) 2 Hui-Ling Wei(魏慧玲) 2 Yu-Ting Wang(王玉廷) 1,2 Jie Pu(普洁) 1,2

1 Institute of Particle and Nuclear Physics, Henan Normal University, Xinxiang 453007, China
2 School of Physics, Henan Normal University, Xinxiang 453007, China

Abstract: Fragment production in spallation reactions yields key infrastructure data for various applications. Based on the empirical SPACS parameterizations, a Bayesian-neural-network (BNN) approach is established to predict the fragment cross sections in proton-induced spallation reactions. A systematic investigation has been performed for the measured proton-induced spallation reactions of systems ranging from intermediate to heavy nuclei systems and incident energies ranging from 168 MeV/u to 1500 MeV/u. By learning the residuals between the experimental measurements and SPACS predictions, it is found that the BNN-predicted results are in good agreement with the measured results. The established method is suggested to benefit the related research on nuclear astrophysics, nuclear radioactive beam sources, accelerator driven systems, proton therapy, etc.

Keywords: Bayesian neural network (BNN), spallation reaction, cross sections

DOI: 10.1088/1674-1137/abb657

1 Introduction

A spallation reaction is one of the violent nuclear reactions, and it occurs when a high-energy light particle strikes a target nucleus. Spallation reactions can naturally occur in the cosmos where high-energy cosmic rays collide with nuclei [1], resulting in the variation of elements in the universe. They can also be artificially induced at a nuclear facility, such as accelerator driven systems (ADS) for nuclear waste disposal, radioactive nucleus production, or during the proton therapy process using accelerated protons. The incident energy in a spallation reaction is above tens of $A$ MeV, which covers the ranges of intermediate energy, relativistic energy, and even higher energies. As a result of the spallation reaction, various radioactive nuclei can be produced, the research regarding which has important applications in several disciplines, including nuclear physics, nuclear astrophysics, isotopic-separation-online (ISOL)-type radioactive-ion-beam facilities, upcoming third-generation radioactive nuclear beams facilities, ADS for nuclear energy [2] and nuclear waste transmutation [3, 4], radioactive nuclei synthesis (especially for extreme nuclei and nuclear isomers [5]), accelerator material radiation [6], and proton therapy [7, 8].

Its extensive applications have attracted both experimental and theoretical interests. In early times, the experiments were usually carried out using accelerated protons on the synchrocyclotron or using cosmic rays. After the reverse kinetic technique was proposed, massive experiments have been performed to measure the spallation fragments above $^{56}$Fe, covering the incident energy range of a few hundreds of $A$ MeV to higher than $A$ GeV. Prodigious data for fragments in spallation reactions have been assembled.

On the theoretical front, many models have been developed. The quantum molecular dynamics model has been improved for spallation reactions [9-11]. The statistical multi-fragmentation model [12-14] and the Liège intranuclear cascade (INCL++) model [15-17] (which has been implanted in the OpenMC, GEANT4, and FLUKA toolkits [18, 19]) can be used to simulate spallation reactions, which are usually followed by a secondary decaying simulation to reproduce the experimental results (a review is also recommended [20]). Some semi-empirical parameterizations, including the EPAX [21] and the SPACS [22], can globally predict the residue fragments in the spallation reaction. The nuclear energy agency (NEA) systematically compared the international codes and models for the intermediate energy activation yields.
to meet the needs of ADS designation, energy amplification, and medical therapy [23]. Difficulties still exist, as the spallation reaction involves a wide incident energy range, as well as a wide range of nuclei, from light to heavy ones. Since the important applications of spallation reactions and the resultant residues productions, it is important to improve the theoretical predictions for the proton induced spallation reactions. The existing models are limited with regard to predicting the light fragments, particularly for proton-induced reactions at low energy. It is necessary to propose a new method for the accurate prediction of the fragments in spallation reactions.

The neural network method, one of the machine learning technologies, is being increasingly employed in nuclear physics. In the standard neural network, it is difficult or even impossible to control the complexity of the model, which is likely to lead to overfitting and reduce the generalization ability of the network [24]. The Bayesian neural network (BNN) provides a good means for avoiding overfitting automatically by defining vague priors for the hyperparameters that determine the model complexity [25]. A prior distribution of the model parameters can be incorporated in Bayesian inference and combined with training data to control the complexity of different parts of the model [26]. Successful examples of BNN applications in nuclear physics can be found in the predictions of nuclear mass [27-29], nuclear charge radii [30], nuclear β-decay half-life [31], fissile fragments [32], and spallation reactions [33]. Based on the vast numbers of measured fragments in laboratories around the world, it is quite promising to predict spallation cross sections accurately and give reasonable uncertainty evaluations with the BNN approach.

In this article, a new method is proposed to predict the fragment cross sections in proton-induced spallation reactions. The BNN method is described in Sec. 2. The results and discussions are presented in Sec. 3, and a conclusion is given in Sec. 4.

2 BNN method

The key principle of Bayesian learning is to deduce the posterior probability distributions through the prior distribution. The process of Bayesian learning is started by introducing the prior knowledge for model parameters. Based on the given training data \( D = (x^{(i)}, y^{(i)}) \) and model assumptions, the prior distributions for all the model parameters are updated to the posterior distribution using Bayes’ rule,

\[
p(\theta | D) = \frac{p(D | \theta) p(\theta)}{p(D)} \propto L(\theta) p(\theta),
\]

where \( \theta \) denotes the model parameters. The posterior distribution \( p(\theta | D) \) combines the likelihood function \( p(D | \theta) \) with the prior distribution \( p(\theta) \), which contains the information about \( \theta \) derived from the observation and the background knowledge, respectively.

The introduction of the prior distribution is a crucial step that allows the prediction to go from a likelihood function to a probability distribution. The normalized quantity \( p(D) \) can be directly understood as the edge distribution of the data, which can be obtained from the integration of the selected model hypothesis and prior distribution \( p(\theta) \),

\[
p(D) = \int p(D | \theta) p(\theta) d\theta.
\]

In this work, the prior distributions \( p(\theta) \) are set as the Gaussian distributions. The precisions (inverse of variances) of these Gaussian distributions are set as gamma distributions [28], which automatically control the complexity of different parts of the model. The likelihood function \( p(D | \theta) \) and the objective function \( \chi^2(\theta) \) are given by

\[
p(D | \theta) = \exp(-\chi^2/2), \quad \chi^2 = \sum_{i=1}^{N} \frac{[y_i - f(x_i; \theta)]^2}{\Delta y_i^2},
\]

where \( \Delta y_i^2 \) is the associated noise scale. The function \( f(x, \theta) \) is a multilayer perceptron (MLP) network, which is also known as the "back-propagation" or "feed-forward." A typical MLP network consists of a set of input variables \( (x_i) \), certain hidden layers, and one or more outputs \( (f_k(x; \theta)) \). For an MLP network with one hidden layer and one output, the function is defined as

\[
f(x; \theta) = a + \sum_{j=1}^{H} b_j \tanh \left(c_j + \sum_{i=1}^{I} d_{ji} x_i \right),
\]

where \( H \) denotes the number of hidden units, and \( I \) is the number of input variables. \( \theta = (d_{ji}, c_j) \) and \( \theta = (b_j, a) \) are the weights and bias of the hidden layers and output layer, respectively.

Based on the theoretical principles and prior knowledge, the posterior distribution can be obtained from the data using Eq. (1). The predictive distribution of output \( y_{\text{new}} \) for a new input \( x_{\text{new}} \) is obtained by integrating the predictions of the model with respect to the posterior distribution of the model parameters,

\[
p(y_{\text{new}} | x_{\text{new}}, D) = \int p(y_{\text{new}} | x_{\text{new}}, \theta) p(\theta) d\theta.
\]

In the MLP network, we are interested in obtaining a reasonable prediction with the new input \( x_{\text{new}} \) rather than the posterior distribution for parameters. For a new input \( x_{\text{new}} \), the model prediction \( y_{\text{new}} \) can be obtained from the mathematical expectation of the posterior distribution,

\[
y_{\text{new}} = E[y_{\text{new}} | x_{\text{new}}, D] = \int f(x_{\text{new}}, \theta) p(\theta | D) d\theta.
\]

The integral of Eq. (5) is complex, and a numerical approximation algorithm will reduce the complexity. The
Markov chain Monte Carlo (MCMC) methods are applied to optimize the model control parameters and obtain the predictive distribution. As one of the MCMC methods, the hybrid Monte Carlo (HMC) algorithm was first introduced by Neal [34] to deal with the model parameters and Gibbs sampling for hyperparameters. The HMC is a form of the metropolis algorithm, where the candidate states are found via dynamical simulation. It makes effective use of gradient information to reduce the random walk behavior. In theory, the Gibbs sampler is the simplest Markov chain sampling method, which is also known as the heatbath algorithm. The hyperparameters are updated separately using Gibbs sampling, which allows their values to be used in obtaining good step-sizes for the discretized dynamics, and they help to minimize the amount of tuning needed for a good performance in HMC. The integral of Eq. (5) is approximately calculated as

\[ y_{\text{new}} = \frac{1}{K} \sum_{k=1}^{K} f(x_{\text{new}}, \theta^k), \]  

(6)

where \( K \) is the number of iteration samples. In a previous work, the BNN approach has been adopted to learn and predict the cross sections directly [33]. To provide some physical guidelines, a recent empirical parameterization for fragment prediction, which is named as the SPACS [22], has been adopted for spallation reactions to obtain the fragment cross sections. In this work, the BNN approach is employed to reconstruct the residues between the experimental data (\( \sigma^{\exp} \)) and the theoretical predictions (\( \sigma^{\text{th}} \)), i.e.,

\[ y_i = \lg(\sigma^{\exp}) - \lg(\sigma^{\text{th}}). \]  

(7)

The cross section predictions via the BNN approach are then given as

\[ \sigma^{\text{BNN+th}} = \sigma^{\text{th}} \times 10^{y_{\text{new}}}, \]  

(8)

where \( \sigma^{\text{th}} \) and \( y_{\text{new}} \) denote the SPACS results and the BNN predictions, respectively. In this work, \( \sigma^{\text{th}} \) refers to the predictions obtained using the SPACS parameterizations. As mentioned earlier, the SPACS was proposed recently and has gained success with regard to spallation reactions.

The inputs of the neural network are the mass numbers \( A_{pi} \) \((A_i)\), the charge numbers \( Z_{pi} \) \((Z_i)\) of the projectile (fragment) nucleus, and the bombarding energy (in MeV/u) \( E_i \), i.e., \( x_i = (A_{pi}, Z_{pi}, E_i, Z_i, A_i) \). Systematic experiments have been performed at the Lawrence Berkeley Laboratory (LBL), RI Beam Facility (RIBF) RIKEN, and FFragment Separator (FRS) GSI, which cover a broad range of spallation nuclei, from \(^{197}\text{Au} \) to \(^{238}\text{U} \). The range of the incident energy varied from 168 MeV/u to 1500 MeV/u, which is relevant for the ADS and proton therapy applications. As listed in Table 1, 3511 datasets from 20 different reactions will be used in this work. The complete data are divided into two different sets, which serve as the learning set and the validation set, respectively. The learning set is built by randomly selecting 3211 datasets, and the remaining 300 datasets are selected as the validation set.

<table>
<thead>
<tr>
<th>( X + p )</th>
<th>( E ) (MeV/u)</th>
<th>numbers</th>
<th>( Z_i )</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{197}\text{Au} + p)</td>
<td>800</td>
<td>352</td>
<td>60-80</td>
<td>[42]</td>
</tr>
<tr>
<td>(^{208}\text{pb} + p)</td>
<td>500</td>
<td>249</td>
<td>69-83</td>
<td>[43]</td>
</tr>
<tr>
<td>(^{208}\text{U} + p)</td>
<td>1000</td>
<td>364</td>
<td>74-92</td>
<td>[45]</td>
</tr>
</tbody>
</table>

3 Results and discussions

The network is trained with different model structures, and 2000 iteration samples are taken in each training. Because the fragment cross sections may differ by several orders of magnitude, an A-factors method [46, 47] is introduced to indicate the validation results of different models, as shown in Fig. 1. The A-factor is defined as

\[ A_f = \frac{1}{N} \sum_{i=1}^{N} \frac{\sigma^{\exp} - \sigma^{\text{pre}}}{\sigma^{\exp} + \sigma^{\text{pre}}}, \]  

(9)

where \( \sigma^{\exp} \) and \( \sigma^{\text{pre}} \) denote the measured data and the predicted data, respectively. In the following section, we discuss the predictions obtained using the BNN + SPACS
method, and compare them with the measured data and the SPACS predictions.

In Fig. 1, the A-factors for the SPACS predictions and BNN method with different numbers of hidden neurons are compared. It is observed that even the BNN with five hidden neurons can significantly improve the predictions. The A-factor decreased with an increase in $H$. When $H$ increased to 23, the A-factor could not be minimized further. A 5-23-1 structure is taken as the optimal network structure, which means that 5 inputs $x_i = (A_{pi}, Z_{pi}, E_i, Z_i, A_i)$, 1 output $y_i = \log(\sigma^{\text{exp}}) - \log(\sigma^{\text{th}})$, and a single hidden layer with 23 hidden neurons are included.

The BNN + SPACS predictions for fragment cross sections in the 1 A GeV $^{136}$Xe + $p$, 168 A MeV $^{136}$Xe + $p$, 356 A MeV $^{40}$Ca + $p$, and 1 A GeV $^{238}$U + $p$ reactions are shown in Fig. 2 to Fig. 5, and compared with the experimental data as well as the SPACS predictions.

In Fig. 2, the predicted and measured fragment cross sections in the 1 A GeV $^{136}$Xe + $p$ reaction are compared. It is observed that the BNN + SPACS predictions agree quite well with the measured data for fragments from $Z = 3$ to 54. For the fragments with $Z < 25$, the underestimation of experimental data by SPACS has been improved significantly.

Figure 3 shows the BNN + SPACS predictions for fragment cross sections in the 168 A MeV $^{136}$Xe + $p$ reaction, which have been measured at RIBF, RIKEN recently [38]. In [38], only the cross sections for $Z > 48$ fragments were reported. Compared to the measured fragments, the BNN + SPACS predictions are very close to
Fig. 3. (color online) Same as Fig. 2 but for the 168 $A$ MeV $^{136}$Xe + $p$ reaction (experimental data taken from Ref. [38]).

Fig. 4. (color online) Same as Fig. 2 but for the 356 $A$ MeV $^{40}$Ca + $p$ reaction (experimental data taken from Ref. [36]).
the SPACS ones. For the light and medium fragments ($Z \leq 25$), the BNN + SPACS predictions are substantially higher than the SPACS ones, which are similar to the results shown in Fig. 2. In addition, the uncertainties are relatively large for the $Z = 11$ isotopes. This may be caused by the insufficient data in the training set for this incident energy.

The spallation of intermediate nuclei is of interest in proton therapy and nuclear astrophysics. The composition of interstellar matter is influenced by cosmic ray (mainly high-energy proton)-induced spallation reactions. The $^{40}$Ca + $p$ reaction at 356 $A$ MeV, for which the predicted and measured results are shown in Fig. 4, has been studied. Compared to the measured results, both the BNN + SPACS and SPACS predictions could reproduce the experimental data quite well. The fragment cross sections predicted by the BNN + SPACS method are in line with those predicted by SPACS except for the $Z = 3$ isotopes.

The predictions for the fragment cross sections in the $1$ $A$ GeV $^{238}$U + $p$ reaction are compared in Fig. 5. The measured data cover the fragments from $Z = 74$ to 92 [45]. It can be observed that the BNN + SPACS method can predict the results well, while the SPACS highly overestimated the measured results for $Z = 91$. The BNN + SPACS method shows the sign of larger than the SPACS method for the fragments of smaller $I$. It seems that for the $Z = 20$ isotopes, for fragments from the spallation of a heavy system such as $^{238}$U, the predictions by BNN + SPACS become more inaccurate, which indicates that the BNN should be further improved by incorporating more data for small $Z$ fragments produced in the heavy systems.

The BNN + SPACS predictions are further verified by using the correlation between the cross section and the average binding energy, which has been elucidated in Ref. [33]. It is generally believed that the isotopic cross section depends on the average binding energy in the form of $\sigma = C \epsilon_0 A^{3/4}$, where $C$ and $\tau$ are free parameters, and $B' = (B - \epsilon_p)/A$ (in which $\epsilon_p = 0.5((-1)^N + (-1)^Z)\epsilon_0 A^{-3/4}$ is the pairing energy for the fragment, and $\epsilon_0 = 30$ MeV). It is evident that the isotopic cross sections predicted by the BNN + SPACS model for $Z = 17$ and 41 obey the correlation very well, which improves both the previous BNN method and also the SPACS method Fig. 6.

From the above results, which cover the fragment cross section predictions for proton-induced reactions ranging from intermediate to heavy nuclei and for the incident energy range of 168 $A$ MeV to 1 $A$ GeV, it is observed that the BNN approach improves the quality of the empirical SPACS parameterizations through the reconstruction of the residual cross sections between the SPACS predictions and measured data. Thus, the BNN + SPACS method can be a new tool to predict the fragment cross section in spallation reactions since it can work in-

![Fig. 5. (color online) Same as Fig. 2 but for the 1 $A$ GeV $^{238}$U + $p$ reaction (experimental data taken from Ref. [45]).](image-url)
dependently after the network is formed.

If we revisit the foundation of the BNN approach, it is natural that the BNN + SPACS model should provide a better prediction than the SPACS parameterizations since the difference between the SPACS and measured data has been minimized. Therefore, the BNN + SPACS model makes improved predictions and also avoids the nonphysical phenomenon by forming a direct BNN learning network from the measured data, as shown in Ref. [33]. The physical implementations of SPACS play important roles to make the BNN + SPACS method reasonable in physics, and the learning and predicting abilities of the BNN also improve the predictions where the SPACS parameterizations do not perform well.

The results indicate that the SPACS tends to underestimate the cross sections for fragments with relatively small \( Z \), whereas it overestimates the cross sections for fragments with \( Z \) close to those for heavy spallation nuclei. These shortcomings have been overcome by the BNN + SPACS model. Limitations still exist for the BNN + SPACS model constructed in this work due to the absence of experimental data for reactions of incident energy below 100 \( A \) MeV and for small spallation systems. For applications in proton therapy, the incident energy can be lower than 100 \( A \) MeV, and the masses of the spallation nuclei are smaller than 30. The smallest spallation reaction adopted in this work is for \( ^{36}\text{Ar} \). If we consider interstellar matter, most of the nuclei have mass numbers smaller than 56. Regarding the spallation reactions induced by high-energy cosmic rays and the proton therapy process, we should improve the prediction model to account for small spallation systems, for which the SPACS parameterizations do not work well. It is important to improve the BNN + SPACS predictions by introducing new data for spallation reactions of intermediate energy (for example, below 100 \( A \) MeV) and intermediate/small systems (\( A < 20 \)), which requires new experiments.

4 Conclusion

In this article, the BNN approach combined with the SPACS parameterizations is proposed to predict the fragment cross sections in proton-induced spallation reactions. Based on 3511 measured fragment cross sections corresponding to 20 spallation reaction systems, the optimal network structure has been established to be 5-23-1, which includes 5 inputs, 1 output, and a single hidden layer with 23 hidden neurons. By reconstructing the residuals between the measured data and the BNN predictions, the BNN + SPACS predictions by introducing new data for spallation reactions of intermediate energy (for example, below 100 \( A \) MeV) and intermediate/small systems (\( A < 20 \)), which requires new experiments.

References

5. S.A. Karamian et al., NIMA, 600: 488 (2009)
8 H. A. Shih et al., Inter. J. Radi. Oncology Biology Physics, 72: S642 (2008)
18 J. Lampinen and A. Vehtari, Neural Networks, 14: 257 (2001)