

# Machine learning based event classification for the energy-differential measurement of the ${}^{\text{nat}}\text{C}(n,p)$ and ${}^{\text{nat}}\text{C}(n,d)$ reactions

P. Žugec<sup>a,\*</sup>, M. Barbagallo<sup>b,c</sup>, J. Andrzejewski<sup>d</sup>, J. Perkowski<sup>d</sup>, N. Colonna<sup>b</sup>, D. Bosnar<sup>a</sup>, A. Gawlik<sup>d</sup>, M. Sabaté-Gilarte<sup>c,e</sup>,  
M. Bacak<sup>c,f</sup>, F. Mingrone<sup>c</sup>, E. Chiaveri<sup>c</sup>,  
The n\_TOF Collaboration<sup>1</sup>

<sup>a</sup>Department of Physics, Faculty of Science, University of Zagreb, Croatia

<sup>b</sup>Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy

<sup>c</sup>European Organization for Nuclear Research (CERN), Geneva, Switzerland

<sup>d</sup>Uniwersytet Łódzki, Łódź, Poland

<sup>e</sup>Universidad de Sevilla, Spain

<sup>f</sup>Technische Universität Wien, Vienna, Austria

arXiv:2204.04955v1 [physics.data-an] 11 Apr 2022

## Abstract

The paper explores the feasibility of using machine learning techniques, in particular neural networks, for classification of the experimental data from the joint  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  reaction cross section measurement from the neutron time of flight facility n\_TOF at CERN. Each relevant  $\Delta E$ - $E$  pair of strips from two segmented silicon telescopes is treated separately and afforded its own dedicated neural network. An important part of the procedure is a careful preparation of training datasets, based on the raw data from Geant4 simulations. Instead of using these raw data for the training of neural networks, we divide a relevant 3-parameter space into discrete voxels, classify each voxel according to a particle/reaction type and submit these voxels to a training procedure. The classification capabilities of the structurally optimized and trained neural networks are found to be superior to those of the manually selected cuts.

**Keywords:** Silicon telescope, Particle recognition, Machine learning, Neutron time of flight, n\_TOF facility

## 1. Introduction

Motivated by an earlier integral cross section measurement of the  ${}^{12}\text{C}(n,p)$  reaction [1, 2], an energy-differential measurement of the  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  reactions [3] was performed at the neutron time of flight facility n\_TOF at CERN. n\_TOF is a sophisticated neutron production facility providing a highly luminous neutron flux spanning 12 orders of magnitude in energy, from 10 meV to 10 GeV. Neutron production, based on the spallation of Pb nuclei from a massive lead spallation target, is induced by 20 GeV proton beam from the CERN Proton Synchrotron. The pulsed beam, with average repetition rate of 0.4 Hz, allows the time of flight technique to be employed for determination of the neutron energy dependence of the measured neutron-induced reactions. Operating since 2001, n\_TOF facility is currently in the third phase of its operation (n\_TOF-Phase3), characterized by parallel utilization of the two experimental areas, referred to as EAR1 and EAR2. Each experimental area is designed in response to a specific set of challenges in measuring the neutron-induced reactions. EAR1 addresses the requirements of the high-energy-resolution as well as the high-neutron-energy measurements, thanks to the long horizontal distance of 185 m from the spallation target, that ensures

a long time of flight and an excellent neutron energy resolution. EAR2, with a vertical flight path of 20 m above the same spallation target, is characterized by a significantly increased neutron flux, relative to EAR1, thus being ideal for measurements of reactions with low cross sections and measurements with small and/or highly radioactive samples. A general description of the n\_TOF facility and EAR1 in particular can be found in Ref. [4]. Detailed characteristics of EAR2 are well documented in Refs. [5, 6, 7]. We also refer the reader to the in-depth description of the neutron flux evaluation in both experimental areas [8, 9] and the concise overview of experimental activities at n\_TOF [10].

Experimental setup for the energy-differential measurement of the  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  reactions consists of two silicon telescopes placed outside the neutron beam, surrounding a 0.25 mm thick natural carbon sample. One telescope is parallel to the neutron beam, while the other is parallel to the carbon sample, which itself is tilted by  $45^\circ$  in respect to the beam. This configuration has been specifically optimized in order to maximize a solid angle coverage and minimize systematic effects in data analysis [3]. Figure 1 shows a schematic diagram of the experimental setup and a close-up of a single silicon telescope. The excellent charged particle response properties of such telescope configuration are reported in Ref. [11]. In short, each telescope consists of two ( $\Delta E$  and  $E$ ) silicon layers – 20  $\mu\text{m}$  and 300  $\mu\text{m}$  thick, respectively, and 7 mm apart. Every layer is

\*Corresponding author. Tel.: +385 1 4605552

Email address: pzugec@phy.hr (P. Žugec)

<sup>1</sup>www.cern.ch/ntof

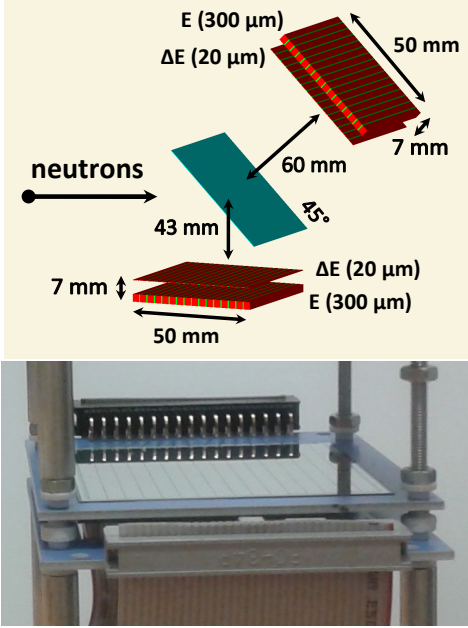


Figure 1: Top: schematic diagram of the experimental setup, consisting of two silicon telescopes surrounding the carbon sample (tilted by  $45^\circ$  relative to the neutron beam). Silicon layers are exaggerated in width. Bottom: close-up of a single telescope, showing a stripped structure of a  $\Delta E$ -layer.

5 cm  $\times$  5 cm in lateral dimensions, each comprising 16 silicon strips, 3 mm wide and separated by a thin layer of inactive silicon. The strips in two layers are oriented in the same direction.

Electronic data from two telescopes (64 channels in total) were digitally recorded at 125 MS/s sampling rate with a 14-bit resolution and analyzed offline by dedicated pulse shape fitting procedures described in Ref. [12]. The shape of the electronic signals is reported in Refs. [11, 13].

One telescope has already been successfully used in an extremely challenging measurement of the  ${}^7\text{Be}(n,p)$  reaction at n\_TOF [14]. Due to the central importance to as yet unresolved Cosmological Lithium Problem, this experiment was also accompanied by a measurement of the  ${}^7\text{Be}(n,\alpha)$  reaction [15], performed by using a similar type of a silicon sandwich detector [16].

Unlike the  ${}^7\text{Be}(n,p)$  measurement performed at EAR2, a joint  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  measurement was performed at EAR1, thus relying on a different data acquisition chain. In the latter case some time offsets were observed between separate data acquisition channels, each corresponding to a particular silicon strip. In order to correct for this in the offline analysis, a simple but effective synchronization method has been developed [13]. Furthermore, due to the challenging features of the  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  reactions – in particular the sensitivity to excited states in daughter nuclei and the expected anisotropy in angular distributions of the reaction products – a special method was developed for the analysis of the experimental data, aiming at the optimal extraction of energy-differential cross sections [17]. What remains is a classification of the experimental counts according to the reaction type (i.e. the reaction products), in order for the above method to be correctly applied to proper datasets.

Such classification, its feasibility as well as its optimal implementation, is the subject of this work. We aim to investigate the feasibility of applying machine learning techniques for this purpose and to put forward all the necessary steps for their implementation.

This work is a part of the ongoing efforts to introduce machine learning techniques into a widespread practice at n\_TOF [18, 19], as they naturally lend themselves to a wide variety of the classification and inference problems. Their amenability to physical sciences has long since been recognized, especially within the nuclear and particle physics [20]. As such, they are widely used in the experimental neutron physics [21, 22, 23], heavy ion collisions [24] and are famously adopted in various Higgs boson related analyses [25, 26, 27], to name just a few applications.

A beautiful example (unrelated to n\_TOF) of the particle identification using a  $\Delta E$ - $E$  technique may be found in Ref. [28]. The general shape of  $\Delta E$ - $E$  patterns is well understood from a theoretical point of view [29]. Expectedly, one can find earlier applications of neural networks for the particle identification in the silicon-detector based measurements [30, 31]. A quite interesting example is a neural network based method that does not require learning [32]. Understandably, our analysis features its own specific requirements, warranting a careful documentation of issues and used methods. Of course, neural networks have also been used for the particle identification by means other than  $\Delta E$ - $E$  discrimination, using the detectors different from silicon-based ones. One recent example may be found in Ref. [33].

Section 2 lays out the motivation for adopting machine learning techniques. Section 3 describes the preparation of the training datasets. Section 4 reports the details of the neural network optimization and examines the quality of the obtained classification. Section 5 sums up the main conclusions of this work.

## 2. Motivation

In the analysis of the  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  data recorded by two silicon telescopes, each relevant  $\Delta E$ - $E$  pair of silicon strips must be separately taken into account. The relevance of a particular pair of strips is established on a basis of the rate of signals in coincidence (within a time window of  $\pm 100$  ns [13]) caused by the detection of charged particles passing through that particular pair. These pairs mostly include the closest and next-to-closest strips from the opposing  $\Delta E$  and  $E$  layers of a given silicon telescope. In a context of a joint  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  measurement these particles include protons, deuterons, tritons and  $\alpha$ -particles from different types of neutron-induced reactions on carbon. In that,  $\alpha$ -particles are of no importance to this work as they are not the subject of current experimental investigation and are well separated from other particle patterns in the coincidence spectra.

Among procedures to be separately applied to each relevant  $\Delta E$ - $E$  pair of strips is a particle recognition, i.e. an identification of a charged particle causing a coincidence of signals in the silicon strips. This is done by observing a distinct spectral signature of specific particles. Based on their mass and

charge, each particle type yields separate and recognizable correlations between energies ( $\Delta E$  and  $E$ ) deposited in two silicon layers. The spectral  $\Delta E$ - $E$  patterns also depend on the energy  $E_n$  of incident neutrons. A mechanism behind this dependence is clear: the neutron energy  $E_n$  determines the energy of released reaction products, thus affecting both the details of their interaction with the silicon detectors and the available energy to be deposited between the silicon layers. Each coincidence is therefore characterized by three parameters – energies  $\Delta E$  and  $E$  deposited in two silicon layers, and the neutron energy  $E_n$  – requiring a characterization of the particle-specific patterns in a 3-dimensional parameter space. In that, the neutron energy  $E_n$  is inferred from the neutron time of flight, extracted from the signal time relative to reference time provided by a so-called  $\gamma$ -flash [4, 13].

Although similar between different  $\Delta E$ - $E$  pairs of strips, these patterns are by no means identical. One of the reasons for these discrepancies is geometric in nature: particles incident on different pairs of strips leave the sample under different exit angles (affecting the path-length through the sample) and impinge on pairs of strips at angles dependent on the strip position (affecting the path-length through given strips). The other reason is kinematic: pairs of strips at different positions around the sample intercept the reaction products with different angle-energy correlations due to their kinematic boost from the center of mass frame of the incident neutron and the target nucleus. This affects the correlation between the incident neutron energy  $E_n$  and the energies  $\Delta E$  and  $E$  deposited in different pairs of strips.

Therefore, we are faced not only with the task of fencing the portions of a 3-parameter space ( $\Delta E, E, E_n$ ) corresponding to different particle types – i.e. of identifying the optimal 2-dimensional boundaries separating these parameter subspaces – but also of doing that separately for each relevant pair of silicon strips. Adjusting these boundaries manually with hope of achieving an *optimal* separation is a daunting challenge even for a single pair of silicon strips, let alone for a multitude of them. Fortunately, machine learning techniques are an appropriate toll for such task. They excel at complex classifications that would otherwise require a massive manual effort, such as having to identify a proper (analytical or otherwise) form for a multidimensional boundary within a multidimensional parameter space, where the separation of data into desired classes might not even be easily visualized. On the other hand, machine learning techniques usually require a careful preparation and advance investigation of how well they can be applied to a task at hand; it is by no means guaranteed that they can yield an optimal or even satisfactory solution to a particular problem. In this work we describe these preparatory considerations and analyze the applicability of neural networks to our particle classification problem. Many of these preliminary procedures – e.g. preparing a quality training dataset by means of detailed simulations and implementing the methods for assessing the quality of particle classification – need to be performed regardless of a selected approach to a problem. This will soon become clear, as we will also attempt a simple manual classification in order to judge it against the results obtained from the

trained neural networks and to justify a selection of one procedure over the other. We note that this work presents a self-contained *proof of concept* that machine learning techniques may be used for a high-fidelity classification of data obtained by silicon telescopes, with discrimination quality superior to that of the manual classification procedures. The experimental data from a joint measurement of the  $^{nat}\text{C}(n,p)$  and  $^{nat}\text{C}(n,d)$  reactions are pending the completion of several analysis procedures, required before the application of classification techniques described herein.

### 3. Training dataset preparation

#### 3.1. Geant4 simulations

We use Geant4 simulations of the experimental setup, comprising a realistic software replica of the natural carbon sample and of two silicon telescopes, in order to obtain the 3-dimensional coincidence spectra for all relevant  $\Delta E$ - $E$  pairs of strips, and for each relevant particle type ( $p, d, t$ ) from the separate neutron induced reactions. As the natural carbon consists of 98.9% of  $^{12}\text{C}$  and 1.1% of  $^{13}\text{C}$ , for each of these isotopes we consider four types of reactions – ( $n,p$ ), ( $n,np$ ), ( $n,d$ ) and ( $n,t$ ) – making a total of eight reactions.  $Q$ -values of these reactions and their energy thresholds in the laboratory frame, where the carbon target is at rest, are listed in Table 1.

Angular distributions of these reactions do not affect the basic shapes of the 3-dimensional coincidence patterns, as these shapes are determined by the kinematics of reaction products and their interaction with the materials in their path. In other words, a spectral place of a given coincidence event, i.e. of a triple ( $\Delta E, E, E_n$ ) corresponding to a given particle/reaction, is almost fully (apart from fluctuations related to the particle interaction with surrounding materials) determined by the initial energy and the emission angle of this particular reaction product. However, the angular distributions do affect the heights of these spectra, i.e. the occupancy of specific portions of a 3-parameter space. Since the specific particle/reaction spectra are expected to overlap, at least in part, these occupancies become relevant in determining which particle/reaction not only occupies, but dominates a given section of a parameter space – a consideration relevant for the later analysis. Thus, some care must be taken not only in implementing the kinematics of simulated reaction products, but also in implementing the statistical distributions governing their emission. The relevant distributions include (1) the angular distribution of reaction products, (2) the distribution of energy thresholds (related to the existence of excited states in daughter nuclei), and (3) the distribution of energy among reaction products when there are more than two products in the exit channel. Naturally, one should use accurate distributions, provided they were available. In case of the neutron induced reactions on carbon isotopes, many of the distributions are poorly known. For this reason we assume the simplest form of these distributions – uniform whenever possible – as the most unbiased course of action in the absence of reliable preexisting data. We describe the role of each relevant distribution below.

Table 1: Relevant neutron induced reactions on the  $^{12}\text{C}$  and  $^{13}\text{C}$  isotopes, characterized by their  $Q$ -value and the corresponding energy threshold in the laboratory frame, where the carbon target is at rest prior to reaction.

Reaction	$Q$ [MeV]	$E_{\text{thr}}$ [MeV]
$^{12}\text{C}(n, p)^{12}\text{B}$	-12.59	13.64
$^{12}\text{C}(n, np)^{11}\text{B}$	-15.96	17.30
$^{12}\text{C}(n, d)^{11}\text{B}$	-13.73	14.89
$^{12}\text{C}(n, t)^{10}\text{B}$	-18.93	20.52
$^{13}\text{C}(n, p)^{13}\text{B}$	-12.65	13.64
$^{13}\text{C}(n, np)^{12}\text{B}$	-17.53	18.89
$^{13}\text{C}(n, d)^{12}\text{B}$	-15.31	16.49
$^{13}\text{C}(n, t)^{11}\text{B}$	-12.42	13.38

We simulate each reaction from Table 1 separately. For the initial neutron energy  $E_n$  we generate the relevant reaction products: protons, deuterons or tritons. The neutron energies are sampled within a relevant energy range from the reaction threshold up to 30 MeV, several MeV above initially estimated upper limit for the reliable analysis of experimental  $^{\text{nat}}\text{C}(n, p)$  and  $^{\text{nat}}\text{C}(n, d)$  data [3]. The initial kinematic parameters of reaction products are calculated from the manually implemented (relativistic) kinematics, based on emission angle and emission energy sampled from relevant distributions. In that, we also account for their spatial sampling along the width and depth of the carbon sample, in accordance with the neutron beam profile.

All daughter nuclei from Table 1 –  $^{10}\text{B}$  [34],  $^{11}\text{B}$  [35],  $^{12}\text{B}$  [36] and  $^{13}\text{B}$  [37] – feature a rich spectrum of excited states. As a consequence, each listed reaction may leave a daughter nucleus either in the ground state or in any of its available excited states. For a nucleus left in an excited state, the  $Q$ -value is lower and the energy threshold is higher than those from Table 1, which correspond to a daughter nucleus left in the ground state. Furthermore, due to considerable overlap of many excited states because of their large energy widths, there is a continuum of energy thresholds, rather than just a few discrete values. In the absence of detailed information on branching ratios of these states, it would make little sense trying to implement other details of these states in the sampling of reaction products. Instead, after sampling the initial neutron energy  $E_n$ , we uniformly sample a reaction threshold between the lowest threshold from Table 1 and the available energy  $E_n$ , corresponding to a selection of a daughter nuclei state between the ground state and the highest one energetically available. In other words, for a given neutron energy we sample a continuous and uniform distribution of branching ratios.

For the  $(n, p)$ ,  $(n, d)$  and  $(n, t)$  reactions with two particles in the exit channel, the selected neutron energy and the selected reaction threshold uniquely determine the energy of reaction products in the center-of-mass frame. Thus, we have yet to sample only their emission angle. We select an isotropic emission in the center-of-mass frame. By means of manually implemented particle kinematics we boost the kinematic parameters of reaction products into the laboratory frame and generate a relevant particle  $(p, d, t)$  with thus obtained energy and emission angle.

For the  $(n, np)$  reaction with three particles in the exit channel

an additional sampling is required. This is due to the fact that for more than two particles in the exit channel, the energy distributions of reaction products are continuous, even in the center-of-mass frame. Again, in the absence of detailed information on these energy distributions, we uniformly sample a proton energy in the center-of-mass frame, between 0 and the maximum kinematically allowed value. As before, proton’s kinematic parameters are then boosted into the laboratory frame, and a proton is generated.

In order to investigate a possibility of background counts from other reactions besides  $(n, np)$  and  $(n, t)$ , we have run an independent batch of Geant4 simulations, generating incident neutrons and observing a detection of reaction products from three different inelastic cascade models: Binary cascade, Bertini cascade and INCL++/ABLA model (INCL intranuclear cascade coupled to the ABLA deexcitation model). Excluding the counts from the  $(n, np)$  and  $(n, t)$  reactions, the background from other reactions was found completely negligible within a relevant portion of a 3-parameter space, i.e. it is well separated from the patterns of the  $(n, p)$  and  $(n, d)$  reactions.

Figure 2 shows examples of simulated  $\Delta E$ - $E$  patterns for different particle/reaction types, corresponding to the neutron energy of  $E_n = 27$  MeV – an arbitrarily selected slice through a 3-parameter space, from an arbitrarily selected  $\Delta E$ - $E$  pair of silicon strips. All simulated data have been smeared by the energy resolution of silicon strips – relative resolution of 3% FWHM for  $\Delta E$ -strips and 0.5% FWHM for  $E$ -strips [11]. Proton spectra of the  $(n, p)$  and  $(n, np)$  reactions and a deuteron spectrum of the  $(n, d)$  reaction are shown for  $^{12}\text{C}$ . For visual reasons related to a high energy threshold of the  $^{12}\text{C}(n, t)$  reaction (see Table 1), a triton spectrum of the  $(n, t)$  reaction is shown for  $^{13}\text{C}$ . The patterns of remaining particle/reaction types are equal in shape between  $^{12}\text{C}$  and  $^{13}\text{C}$ . For display purposes all spectra are normalized such that the highest-content pixel is set to 1.

Solid lines from Fig. 2 show the manual cuts that will later be judged against trained neural networks. They also serve as a guide for the eye in a visual analysis of spectra from separate plots. A unique set of manual cuts was determined for all  $\Delta E$ - $E$  pairs of silicon strips, by observing the projected spectra of all counts, from all pairs of silicon strips and from all neutron energies. It is, of course, an entirely fortunate circumstance that such unique set of cuts can be identified at all (at least approximately), implying that the variation of these cuts across the neutron energy and among silicon pairs is not drastic. However, it is expected that this simplified set should not be optimal. It is precisely the goal of machine learning techniques to alleviate both the technical limitations of manual cuts – such as the necessity for simple analytic forms, preferably independent of  $E_n$  – and to obviate the manual repetition of finding the cuts for each pair of silicon strips separately.

Manual cuts from Fig. 2 are clearly labeled in the first plot, that of the  $^{12}\text{C}(n, p)$  reaction. Beyond cuts A and E the amount of data from relevant reactions is negligible or entirely nonexistent. One can mostly find the background counts below cut A. Above cut E is the spectrum of  $\alpha$ -particles from the  $(n, \alpha)$  reactions, which is well separated from proton, deuteron and triton spectra, hence not being of interest to this work. For this rea-

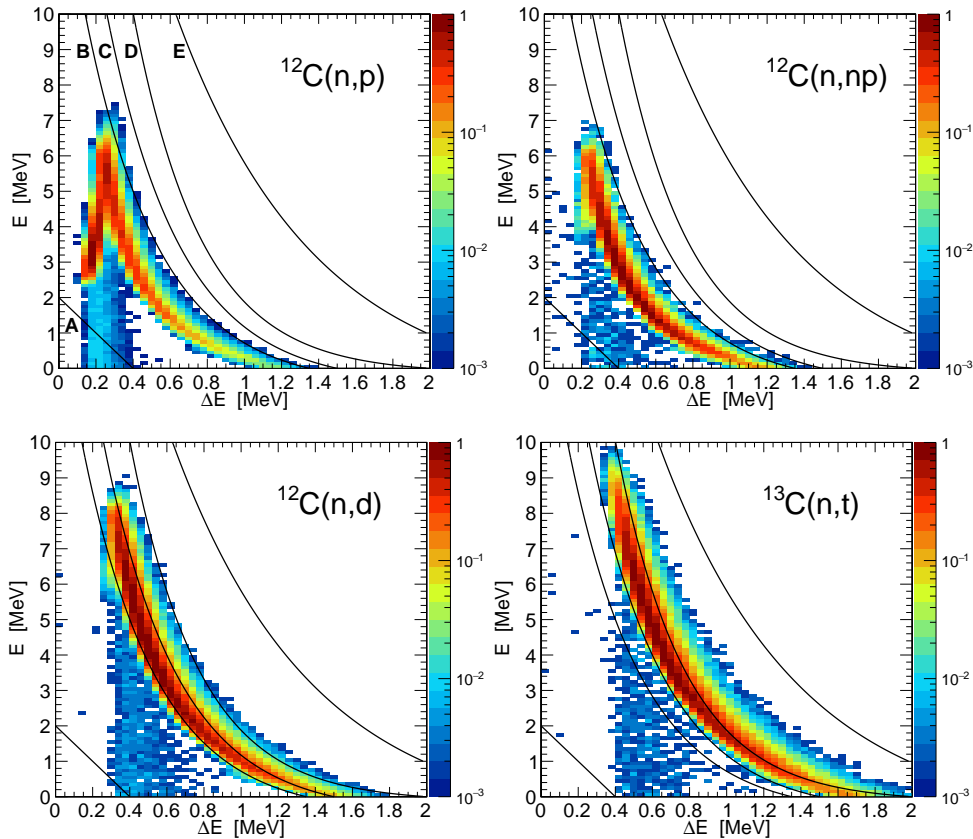


Figure 2: Simulated  $\Delta E$ - $E$  patterns for an arbitrarily selected pair of silicon strips, showing a detector response to protons from the  $^{12}\text{C}(n,p)$  and  $^{12}\text{C}(n,np)$  reactions, to deuterons from the  $^{12}\text{C}(n,d)$  and to tritons from the  $^{13}\text{C}(n,t)$  reaction, at a neutron energy of  $E_n = 27$  MeV. All spectra are smeared by the energy resolution of silicon strips and are normalized such that the highest-content pixel is set to 1. The meaning of displayed manual cuts is discussed in the main text.

son all data beyond cuts A and E will be manually discarded, both from neural networks training procedure and from analysis of the experimental data. Cut B separates proton data from deuteron data. Above cut D the amount of deuteron counts is negligible. Cut C was visually estimated as an optimal separation boundary between deuteron and triton spectra. The ordering of cuts C and D indicates a certain overlap between deuteron and triton spectra, clearly demonstrating a necessity of finding an *optimal* separation boundary between them.

### 3.2. Training approach

After obtaining the sets of counts corresponding to specific particle/reaction types we are, at least in principle, ready to apply classification methods directly to these raw data, on a count-by-count basis. However, there are several technical issues to consider. The first one is the expected processing time required for training the neural networks, as the amount of counts to be taken for a reliable training should be quite large. At the same time, training needs to be repeated for each relevant pair of silicon strips separately. In that, we will consider 72 pairs of strips in this work, consisting of the relevant closest and next-to-closest  $\Delta E$ - $E$  neighbors. Though training time issues are not crucial to the quality of training results, it is certainly desirable to reduce this time as much as possible and even to make it independent of a considered amount of counts.

The second issue, of much greater impact on the quality of results, is related to a statistical significance of specific counts. The first, most basic measure of this significance is the amount of counts of a specific type. We wish that separation of  $\Delta E$ - $E$  patterns be *at all* neutron energies  $E_n$  in *equal measure* affected by training data from the corresponding neutron energy. However, despite a uniform  $E_n$  sampling the number of detected coincidences may differ significantly between separate  $E_n$ -slices. This may be due to low energy of reaction products at a given neutron energy or due to their angular distribution affecting the emission rate into a solid angle covered by a specific pairs of silicon strips. Since, at the end of training, an optimized boundary between specific particle/reaction counts will be continuous in a 3-parameter space, a separation of  $\Delta E$ - $E$  patterns from certain  $E_n$ -slices may be disproportionately affected by a larger number of counts from other slices. Nevertheless, we do not wish to perform a separate training for each  $E_n$ -slice, as the separation boundaries would discontinuously fluctuate between slices, while a well-behaved boundary should be continuous throughout the entire parameter space.

For these reasons we take different approach to constructing a training dataset, aiming to solve both issues at once. We divide a relevant portion of a 3-dimensional parameter space into discrete volumetric units – *voxels*. Based on a statistical significance of specific counts we determine to which particle/reaction

Table 2: Parameters defining a relevant portion of a 3-parameter space and its partitioning into voxels.

	Minimum	Maximum	Bins	Bin width
$\Delta E$	0 MeV	2 MeV	50	0.04 MeV
$E$	0 MeV	10 MeV	100	0.1 MeV
$E_n$	10.5 MeV	30.5 MeV	60	0.33 MeV

type each voxel (dominantly) belongs. Then we submit these voxels, instead of single counts, to a neural network training procedure. This approach evidently solves the training time issue, since (for statistically significant amount of counts) the number of filled voxels is considerably smaller than and approximately independent of the number of counts occupying the parameter space. On the other hand, a statistical significance of counts within and between  $E_n$ -slices is easily affected by manual manipulations during the voxel-type assignment procedure.

### 3.3. Training dataset construction

A relevant portion of a parameter space that we consider spans from 0 MeV to 2 MeV in  $\Delta E$ , from 0 MeV to 10 MeV in  $E$ , and from 10.5 MeV to 30.5 MeV in  $E_n$ . Other details of its partitioning are listed in Table 2. Since we expect to analyze the experimental data from joint  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  measurement within the neutron energy windows of 1 MeV [3], we adopt for  $E_n$  somewhat finer neutron energy resolution of 3 bins per MeV. In total, we divide a parameter space into  $3 \times 10^5$  voxels (the same partitioning as displayed in Fig. 2). Only a subset of all these voxels will be submitted to a neural network training procedure.

In order to determine which particle/reaction type (if any at all) dominates specific voxels, we need to manually estimate each type's *detection yield* within a given voxel, since each particle/reaction was simulated separately. For simplicity of notation, we enumerate voxels by a triple  $(\Delta E, E, E_n)$  of energies relevant to that voxel (e.g. energy coordinates of its center). In that, we are not interested in absolute detection yields but only in relative contributions between different particle/reaction types. Thus, for each particular type of reaction  $\mathcal{R} - (n,p)$ ,  $(n,np)$ ,  $(n,d)$  or  $(n,t)$  – on each particular carbon isotope  $C$  (either  ${}^{12}\text{C}$  or  ${}^{13}\text{C}$ ;  $C = 12, 13$ ), detected in coincidence by each particular  $\Delta E$ - $E$  pair of silicon strips  $\mathcal{P}$ , we define a simpler quantity  $x_{C,\mathcal{R}}^{(\mathcal{P})}$  that is proportional to a true detection yield<sup>2</sup>, up to the areal density of a sample as an omitted multiplicative factor. For simplicity and clarity we still refer to this quantity as

<sup>2</sup>For thin samples, a *reaction yield* (the amount of reaction products per incident neutron) equals  $Y_{C,\mathcal{R}}(E_n) = \eta_C \sigma_{C,\mathcal{R}}(E_n)$ , with  $\sigma_{C,\mathcal{R}}(E_n)$  as the reaction cross section and  $\eta_C = A_C \eta$  as an areal density of a particular carbon isotope (in number of atoms per unit area),  $A_C$  and  $\eta$  being an isotopic abundance and a total areal density of carbon sample. A *true detection yield*  $X_{C,\mathcal{R}}^{(\mathcal{P})}$  is then:

$$X_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) = \epsilon_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) Y_{C,\mathcal{R}}(E_n) = \eta \frac{N_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}{N_{C,\mathcal{R}}(E_n)} A_C \sigma_{C,\mathcal{R}}(E_n),$$

with  $\epsilon_{C,\mathcal{R}}^{(\mathcal{P})}$  as a detection efficiency, corresponding to a fractional term from Eq. (1). A simplified definition of detection yield from Eq. (1) differs only by a factor  $\eta$  from true detection yield.

*detection yield* throughout the paper. For each voxel we estimate these specific detection yields as:

$$x_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) \approx \frac{N_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}{N_{C,\mathcal{R}}(E_n)} A_C \bar{\sigma}_{C,\mathcal{R}}. \quad (1)$$

$N_{C,\mathcal{R}}(E_n)$  is a simulated number of reaction products, generated for the reaction  $\mathcal{R}$  on the carbon isotope  $C$  within the neutron energy window at  $E_n$ .  $N_{C,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)$  is a number of reaction products detected within a given voxel of the parameter space by the silicon pair  $\mathcal{P}$ .  $A_C$  are natural abundances of relevant carbon isotopes ( $A_{12} = 98.9\%$  and  $A_{13} = 1.1\%$ ).  $\bar{\sigma}_{C,\mathcal{R}}$  the roughly estimated cross section averages within neutron energy intervals spanning from a particular reaction threshold up to 30 MeV. Since evaluated data on these cross sections show large variations between available libraries – both in range and in values – the cross section estimates were obtained from Geant4 simulations, by averaging the cross sections predicted by three different inelastic cascade models: Binary cascade, Bertini cascade and INCL++/ABLA model (INCL intranuclear cascade coupled to the ABLA deexcitation model). Energy dependence of each reaction cross section was first extracted from Geant4 simulations using a method from Ref. [2], described in Appendix A. For each reaction an energy dependence was averaged between three cascade models and the averages  $\bar{\sigma}_{C,\mathcal{R}}$  within the energy intervals up to 30 MeV were estimated. We obtained the following values:  $\bar{\sigma}_{12,(n,p)} = 22$  mb,  $\bar{\sigma}_{12,(n,np)} = 15$  mb,  $\bar{\sigma}_{12,(n,d)} = 27$  mb,  $\bar{\sigma}_{12,(n,t)} = 8.5$  mb,  $\bar{\sigma}_{13,(n,p)} = 3.3$  mb,  $\bar{\sigma}_{13,(n,np)} = 15$  mb,  $\bar{\sigma}_{13,(n,d)} = 16$  mb,  $\bar{\sigma}_{13,(n,t)} = 14$  mb. Absolute values of these cross sections are of no importance to this work, only their relative magnitude. In addition, we have confirmed that these values do not play a crucial role in the upcoming analysis of our data, i.e. a voxel classification is rather insensitive to variations in thus obtained  $\bar{\sigma}_{C,\mathcal{R}}$ .

Due to similar reaction thresholds for  ${}^{12}\text{C}$  and  ${}^{13}\text{C}$  (Table 1), reaction products of the same type from two carbon isotopes can not be experimentally distinguished, as they occupy the same portions of a parameter space. Therefore, for each reaction type  $\mathcal{R}$  we continue observing a joint contribution from both carbon isotopes, in accordance with their natural abundance. From this point on, each quantity characterized by  $\mathcal{R}$  and not by  $C$  refers to the natural carbon, as used in a joint measurement of the  ${}^{\text{nat}}\text{C}(n,p)$  and  ${}^{\text{nat}}\text{C}(n,d)$  reactions. Joint contributions to a detection yield follow simply as:

$$x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) = x_{12,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) + x_{13,\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n). \quad (2)$$

For each voxel from every silicon pair we now observe the relative contributions  $r_{\mathcal{R}}^{(\mathcal{P})}$  from specific reaction types:

$$r_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) = \frac{x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}{\sum_{\mathcal{R}'} x_{\mathcal{R}'}^{(\mathcal{P})}(\Delta E, E, E_n)}, \quad (3)$$

and classify the voxels, i.e. assign to each of them one of four labels using the following rules:

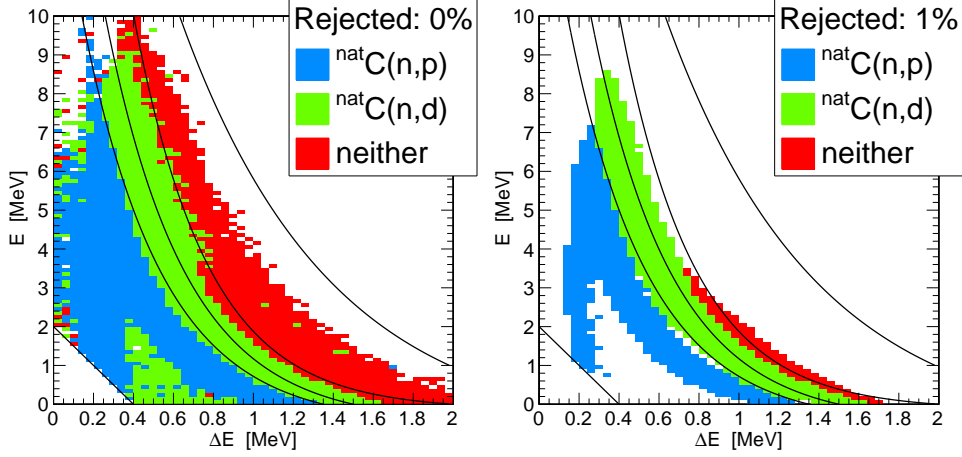


Figure 3: Assignment of the reaction types at  $E_n = 27$  MeV, for an arbitrarily selected pair of silicon strips from Fig. 2. The types are determined based on the reactions shown in Fig. 2 and four more reactions not shown there:  $^{13}\text{C}(n,p)$ ,  $^{13}\text{C}(n,np)$ ,  $^{13}\text{C}(n,d)$  and  $^{12}\text{C}(n,t)$ . Left panel: all voxels originally determined as  $(n,p)$ ,  $(n,d)$  or background type (none manually rejected). Right panel: voxels remaining after rejecting 1% of total detection yield from this particular  $E_n$ -slice. White voxels are of the undecided type (empty, of unclear classification or rejected).

- if the relative yield of the  $(n,np)$  and  $(n,t)$  reactions exceeds 50%:  $r_{(n,np)}^{(\mathcal{P})}(\Delta E, E, E_n) + r_{(n,t)}^{(\mathcal{P})}(\Delta E, E, E_n) > 50\%$ , label a voxel as the *background type*;
- if the relative yield of the  $(n,p)$  reactions exceeds 50%:  $r_{(n,p)}^{(\mathcal{P})}(\Delta E, E, E_n) > 50\%$ , label a voxel as the  $(n,p)$ -type;
- if the relative yield of the  $(n,d)$  reactions exceeds 50%:  $r_{(n,d)}^{(\mathcal{P})}(\Delta E, E, E_n) > 50\%$ , label a voxel as the  $(n,d)$ -type;
- otherwise, if empty of counts or not satisfying any of the previous conditions (i.e. if being of unclear classification), label a voxel as the *undecided type*.

Left panel from Fig. 3 shows an example of thus assigned types within the  $E_n$ -slice at 27 MeV, corresponding to the data from Fig. 2 (keep in mind that four other reactions have been taken into account, besides those shown in Fig. 2). Let us consider the effect on the quality of neural network training, caused by such arrangement of voxel types. There are some sporadic voxels far away from the main pattern corresponding to a specific voxel type, most prominently on a left side of the plot, where several isolated  $(n,d)$  and background voxels mix with the main  $(n,p)$  pattern. Furthermore, there is an entire island of  $(n,d)$  voxels closed off by the  $(n,p)$  pattern. From Fig. 2 we make two important observations. The first is that these voxels are statistically insignificant, relative to the main pattern of their type. The second is that the *primary* reason for these voxels not to have been assigned the  $(n,p)$  type is not so much the fact that non- $(n,p)$  counts from these voxels dominate over the  $(n,p)$  counts, but rather that there are no  $(n,p)$  counts in these voxels at all. If the neural networks were trained on such set of voxels, a 3-parameter boundary between the reaction types would be disproportionately affected by these voxels, considering their statistical (in)significance. For this reason we employ a simple method to exclude such insignificant voxels from interfering with training procedure, in order to ensure greater robustness of boundaries between the reaction types. A method

consists in rejecting from each  $E_n$ -slice the voxels with the lowest detection yield, until a predefined portion (specifically, 1%) of total yield from that particular  $E_n$ -slice has been removed. Formally stated, for each voxel let  $x^{(\mathcal{P})}(\Delta E, E, E_n)$  be a *joint detection yield* from all reactions:

$$x^{(\mathcal{P})}(\Delta E, E, E_n) = \sum_{\mathcal{R}} x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n). \quad (4)$$

Furthermore, for each  $E_n$ -slice let  $\hat{x}_i^{(\mathcal{P})}(E_n)$  be an *ordered*, increasing array of joint yields from within the same  $E_n$ -slice, where index  $i$  substitutes  $(\Delta E, E)$  pairs of parameters, ordered according to increasing voxel content. The number  $\kappa^{(\mathcal{P})}(E_n)$  of ordered voxels from a given  $E_n$ -slice, that we label as empty, i.e. as the undecided type, is the maximum  $\kappa^{(\mathcal{P})}(E_n)$  for which it holds:

$$\frac{\sum_{i=1}^{\kappa^{(\mathcal{P})}(E_n)} \hat{x}_i^{(\mathcal{P})}(E_n)}{\sum_{\Delta E, E} x^{(\mathcal{P})}(\Delta E, E, E_n)} \leq 1\%. \quad (5)$$

In other words, we reject at most 1% of total detection yield from each  $E_n$ -slice.

Right panel from Fig. 3 shows the result of such elimination of superfluous voxels from the left panel. Statistical insignificance of rejected voxels is to be appreciated, as they make up almost half of all initial voxels, while carrying only 1% of a statistical norm.

#### 4. Neural network training

We use the neural network training capabilities of `TMultiLayerPerceptron` class from ROOT. The sets of voxels that are submitted to a training procedure are represented by right panel from Fig. 3. Input data consist of the triples  $(\Delta E, E, E_n)$ , i.e. of parametric coordinates of each voxel. Output data consist of each voxel's type:  $(n,p)$ -type,  $(n,d)$ -type or background type. Voxels of the undecided type are not submitted to a training procedure at all. Considering the undecided

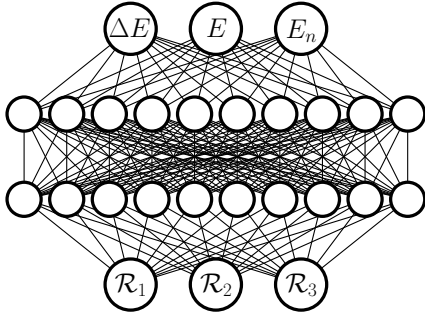


Figure 4: Structure of the adopted neural networks for each pair of  $\Delta E$ - $E$  strips. Inputs are the coordinates  $(\Delta E, E, E_n)$  of each voxel remaining after the rejection procedure from Eq. (5). Outputs are reaction labels  $\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3 \in \{0, 1\}$ .  $\mathcal{R}_1$  stands for the  $(n, p)$  reaction,  $\mathcal{R}_2$  for the  $(n, d)$  reaction and  $\mathcal{R}_3$  for the background reactions. Since the undecided voxels are excluded from training procedure, each submitted voxel has exactly one of the labels equal to 1 and the remaining ones equal to 0.

voxels as separate relevant type would overly restrict the boundary between the  $(n, p)/(n, d)$  patterns and the empty portions of parameter space, effectively causing the undecided type to behave as the background type. Background type indicates that certain portions of parameter space are decidedly not dominated either by the  $(n, p)$  or  $(n, d)$  reaction. On the other hand, the purpose of the undecided type is to allow the generalization and extension of specific type-patterns into the portions of parameter space that are not covered by training data.

Though the default Broyden-Fletcher-Goldfarb-Shanno (BFGS) training method from `TMultiLayerPerceptron` class performs reasonably well, we have found that the stochastic minimization with manually optimized internal hyperparameters  $\tau = 5$  and  $\eta = 0.01$ , and default values of  $\eta_{\text{decay}} = 1$ ,  $\delta = 0$  and  $\epsilon = 0$  converges faster (in fewer training iterations) and requires only about half of processing time per iteration, in comparison with the BFGS method with default settings. Therefore, in this work we use a stochastic minimization with the listed hyperparameter values. For hidden neurons a sigmoid function was used as an activation function.

We have found that the optimal network structure consists of 2 layers of neurons, each composed of 10 neurons. An important requirement in identifying this optimized structure was that both the quality and the consistency of trained boundaries remain stable even when introducing variations in the input data, e.g. by varying a rejection level from Eq. (5). In conclusion, we employ a 2-layer neural network with 10 neurons in each layer. The adopted network structure is displayed in Fig. 4.

Though of secondary importance, a voxel rejection procedure from Eq. (5) further reduces the amount of voxels to be submitted to a training procedure (thus further reducing a training time) since the undecided voxels are not submitted at all. For 72 pairs of silicon strips that we consider in this work, consisting of the relevant closest and next-to-closest  $\Delta E$ - $E$  neighbors, a number of specific-type voxels per silicon pair varies between approximately 47 000 and 87 600 in case of no voxel rejection (represented by left panel from Fig. 3). After rejecting 1% of the lowest voxel content (right panel from Fig. 3),

between 29 000 and 47 500 voxels remain per pair. These numbers should also be compared against the initial  $3 \times 10^5$  voxels composing an entire parameter space.

In order to recover a voxel classification from trained networks, we determine the voxel's type simply by checking which of the output neurons (see Fig. 4) provides a maximum result. We have, of course, tested more stringent conditions such as insisting that  $\mathcal{R}_i > 0.5 > \mathcal{R}_j + \mathcal{R}_k$  for a given voxel to be classified as  $i$ -type (otherwise it is left undecided). However, these stricter conditions do not significantly affect the overall voxel-type recovery rate (reported below). Within a relevant portion of a parameter space these conditions only result in a few sporadic undecided voxels between the specific reaction patterns (similar to the empty voxels between  $(n, p)$  and  $(n, d)$  patterns from the right plot of Fig. 3). In light of no notable improvement in the classification rate, we find this behavior to work against the desired classification outcome. For this reason we adopt the simplest classification recovery scheme (by finding the maximum  $\mathcal{R}_i$ ) as the most efficient one.

We now demonstrate the quality of trained neural networks and compare it with the quality of manual cuts from Figs. 2 and 3. While the convergence of network weights is satisfactory already after a couple of hundred iterations, we use in this work  $10^4$  iterations<sup>3</sup>. For this demonstration we consider the portions of a detection yield that are correctly (true positive) or incorrectly (false positive) classified as  $\mathcal{R} = (n, p)$  or  $\mathcal{R} = (n, d)$  type. We first define the true-positive contributions  $T_{\mathcal{R}}^{(\mathcal{P})}$  to a particular reaction type:

$$T_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) = \begin{cases} x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) & \text{if classified as } \mathcal{R} \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

together with the false-positive contributions  $F_{\mathcal{R}}^{(\mathcal{P})}$ :

$$F_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n) = \begin{cases} \sum_{\mathcal{R}' \neq \mathcal{R}} x_{\mathcal{R}'}^{(\mathcal{P})}(\Delta E, E, E_n) & \text{if classified as } \mathcal{R} \\ 0 & \text{otherwise} \end{cases}. \quad (7)$$

We further define the correctly and incorrectly classified portions  $P_{\mathcal{T}}^{(\mathcal{R})}$  and  $P_{\mathcal{F}}^{(\mathcal{R})}$  of a reaction-specific detection yield:

$$P_{\mathcal{T}}^{(\mathcal{R})}(E_n) = \frac{\sum_{\mathcal{P}} \sum_{\Delta E, E} T_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}{\sum_{\mathcal{P}} \sum_{\Delta E, E} x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}, \quad (8)$$

$$P_{\mathcal{F}}^{(\mathcal{R})}(E_n) = \frac{\sum_{\mathcal{P}} \sum_{\Delta E, E} F_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}{\sum_{\mathcal{P}} \sum_{\Delta E, E} x_{\mathcal{R}}^{(\mathcal{P})}(\Delta E, E, E_n)}. \quad (9)$$

Though these portions may be defined for each  $\Delta E$ - $E$  pair  $\mathcal{P}$  of silicon strips separately, for demonstration purposes we have defined them as averaged over 72 relevant pairs of strips.

<sup>3</sup>We used the parallel processing capabilities of `TProof` class from `ROOT` in order to distribute the training jobs for separate  $\Delta E$ - $E$  pairs of strips between available CPU cores. For training we have used a table computer based on Intel Xeon E5-1603 v3(2.80 GHz) processor, which (for the reported network structure of  $10 \times 10$  neurons) takes approximately  $2 \times 10^{-7}$  minutes per voxel and per training iteration. For the reported number of submitted voxels (between 29 000 and 47 500 per  $\Delta E$ - $E$  pair) it takes between 55 and 90 minutes for  $10^4$  training iterations to be completed for a given pair of  $\Delta E$ - $E$  strips.



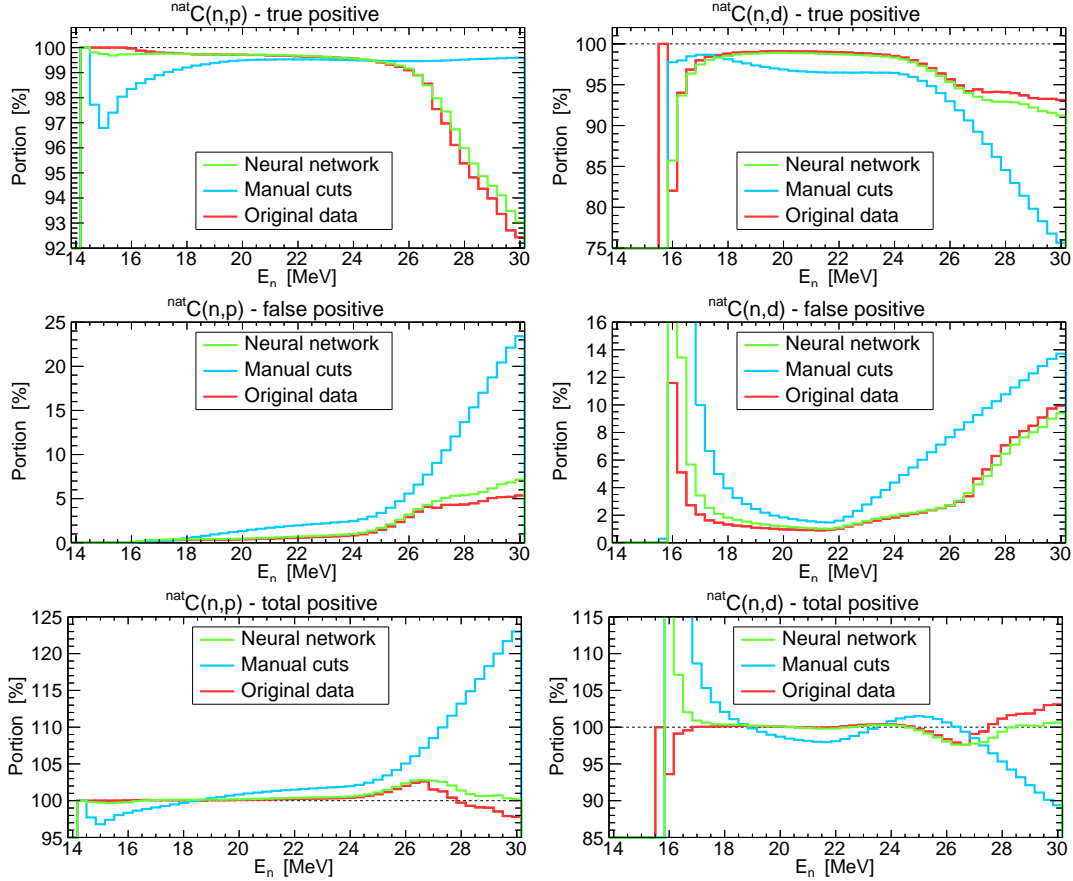


Figure 5: Portions of a reaction-specific detection yield classified correctly or incorrectly, either by the trained neural networks, or by applying the manual cuts, or by the original voxel classification based on the raw detection yield data. Top row: true-positive portions  $P_T^{(R)}(E_n)$ ; middle row: false-positive portions  $P_F^{(R)}(E_n)$ ; bottom row: sum of both contributions to the positive reaction-specific classification. Left column:  $(n,p)$  reaction; right column:  $(n,d)$  reaction.

Figure 5 shows thus defined portions of true-positive and false-positive classifications. Left column and right column of plots correspond to the  $(n,p)$  and  $(n,d)$  reactions, respectively. Top row shows the true-positive portions  $P_T^{(n,p)}$  and  $P_T^{(n,d)}$ . Middle row shows the false-positive portions  $P_F^{(n,p)}$  and  $P_F^{(n,d)}$ . Bottom row shows the total amount of positive classifications, i.e. the sum of true-positive and false-positive classifications for both types of reactions. The most important comparison is that between the trained neural networks and the laboriously selected manual cuts. A quality of these two classification schemes is also compared with the initial classification based on the raw detection yield data (an example from the left plot of Fig. 3), which we refer to as the *original classification*. The original classification shows how much of a reaction-specific detection yield could be successfully recovered if every single voxel could be correctly recognized (e.g. by overfitted neural networks), thus serving as a reference point for evaluating the quality of any other type of classification. It should be noted that one does not wish to use the original classification for the analysis of experimental data, because it does not lend itself to any kind of generalization outside the initially identified patterns. In other words, it cannot resolve the type of any (experimentally obtained) count which falls within originally undecided voxels.

It can be readily appreciated that the classification quality of trained neural networks barely deviates from the quality of original classification. A deviation of the original classification from 100% in true positives and from 0% in false positives is due to the fact that within some voxels there might have been counts from multiple reactions – e.g. a mixture of  $(n,p)$  and  $(n,d)$  counts – while each voxel is assigned only a single reaction type at the end of the classification procedure. It should be noted that the level of a reaction type recovery from original classification depends on the adopted voxel size, since the overlap of counts decreases by decreasing the voxel size. Therefore, one needs to ensure that the adopted voxel density is not too dissimilar from the one used in the analysis of experimental data.

In general, one aims for as high as possible portion of true-positive classifications, and as low as possible portion of false-positive classifications. The fact that at least up to 25 MeV both the original classification and the neural network classification keep close to 100% in true positives (top row) and to 0% in false positives (middle row), suggests that the experimental data from a joint  ${}^{\text{nat}}C(n,p)$  and  ${}^{\text{nat}}C(n,d)$  measurement can be reliably analyzed (at least) up to 25 MeV. Based on these results, one can expect an uncertainty in the identification within a few percent when using the optimized neural networks.

The results from trained neural networks are clearly superior to those from manual cuts, at least up to the relevant limit of 25 MeV. Even when manual cuts seem to outperform trained networks, which appears to be the case for true-positive ( $n,p$ ) classifications, they do so at the price of performing poorly at the complementary task, that of minimizing false-positive classifications. This is clearly seen from the total portion of positive classifications (bottom row) for both types of reactions. The desired outcome in these plots is to be as close to 100% as possible. Trained networks clearly outperform manual cuts in the overall positive classification.

Total portions of positive classification close to 100% also suggest an accidental – not to be relied upon – but possibly favorable type of outcome to be expected in the analysis of experimental data. When one starts losing portions of the correctly classified yield due to a decrease in true-positive classifications, a *simultaneous* increase in false-positive classifications may lead to a partial cancellation of these opposing effects and at least a partial restoration of a reaction-specific detection yield. For example, a portion of deuterons misclassified as protons may partially recover the portion of protons misclassified as deuterons, and vice versa. While each departure of true-positive classifications from 100% and of false-positive classifications from 0% is a negative effect in itself, their combined effect opens a possibility for a more optimistic outcome, conditional on the identification of the optimal type-separation boundaries.

## 5. Conclusions

We have developed a machine learning based procedure for classification of the proton and deuteron counts from a joint  $^{\text{nat}}\text{C}(n,p)$  and  $^{\text{nat}}\text{C}(n,d)$  measurement from n\_TOF. An important part of the procedure is a careful preparation of training datasets. Training data, consisting of triples  $(\Delta E, E, E_n)$  of relevant parameters identifying each count, were obtained by Geant4 simulations of eight relevant neutron induced reactions on natural carbon:  $(n,p)$ ,  $(n,np)$ ,  $(n,d)$  and  $(n,t)$  reactions on both  $^{12}\text{C}$  and  $^{13}\text{C}$  isotopes. In order to solve practical difficulties in applying a neural network training procedure to a set of raw counts, we have first constructed a 3-parameter space and divided it into discrete voxels. Each voxel was assigned a specific type, discriminating spectral patterns of the  $(n,p)$  and  $(n,d)$  reactions from competing contributions of the  $(n,np)$  and  $(n,t)$  reactions. These spectral patterns were further refined based on careful considerations and thus obtained sets of voxels were submitted to a neural network training procedure. Spectral patterns for each relevant  $\Delta E$ - $E$  pair of silicon strips, 72 in total, were treated separately, each being afforded its own neural network. To this end, a `TMultiLayerPerceptron` class from ROOT was used. Both the input data and the network parameters were varied in order to identify the optimal network configuration, as well as the optimal training procedure. A stochastic minimization with manually adjusted hyperparameters  $\tau = 5$  and  $\eta = 0.01$  was found to provide the best performance, regarding both the convergence rate and the processing time per training iteration. A 2-layer neural network with 10 neurons

in each layer was adopted as the optimal network structure. A performance of trained networks was compared against carefully determined manual cuts between reaction types, by examining the portions of true-positive and false-positive classifications for the relevant  $(n,p)$  and  $(n,d)$  reactions. Trained neural networks were found to be clearly superior in quality of classification and to be a basis for a reliable analysis of experimental data up to at least 25 MeV of neutron energy.

## Acknowledgements

This work was supported by the Croatian Science Foundation under Project No. 8570. Geant4 simulations were run at the Laboratory for Advanced Computing, Faculty of Science, University of Zagreb.

## Appendix A. Extracting cross sections from Geant4

We describe a procedure for extracting any type of cross section from Geant4 simulations. We stress that the simulations described here are completely independent of and in all aspects different from simulations described in Section 3.1; they are in no manner related to or constrained by any particular experimental setup. We run these simulations separately for pure  $^{12}\text{C}$  sample and for pure  $^{13}\text{C}$  sample, which is of arbitrary but preferably regular shape and dimensions. For a given reaction type  $\mathcal{R}$  on a pure isotope  $C$ , a first-chance reaction yield  $Y_{C,\mathcal{R}}$  (without the multiple scattering effects) may be expressed as:

$$Y_{C,\mathcal{R}}(E_n) = \left(1 - e^{-\eta_C \sigma_C(E_n)}\right) \frac{\sigma_{C,\mathcal{R}}(E_n)}{\sigma_C(E_n)} = \frac{N_{C,\mathcal{R}}(E_n)}{N_C(E_n)}. \quad (\text{A.1})$$

Here  $\eta_C$  is the areal density of a specific sample (in number of atoms per unit area),  $\sigma_{C,\mathcal{R}}(E_n)$  is the cross section for a particular reaction of interest, and  $\sigma_C(E_n)$  is the total cross section for any reaction at all. The sample is irradiated by neutrons (or any other relevant incident particles) following some predefined energy distribution. Reaction yield (for a given sample) may be reconstructed by counting the total number  $N_C(E_n)$  of neutrons generated with the energy  $E_n$ , and the number  $N_{C,\mathcal{R}}(E_n)$  of first-chance reactions – first occurrences of a specific reaction, which are counted only if that reaction was the first one to take place among all possible neutron interactions, including the elastic scattering. By first counting the number  $N_C(E_n)$  all possible first-chance reactions, governed by the total cross section  $\sigma_C(E_n)$ , one obtains upon inverting Eq. (A.1):

$$\sigma_C(E_n) = -\frac{1}{\eta_C} \ln \left(1 - \frac{N_{C,\mathcal{R}}(E_n)}{N_C(E_n)}\right). \quad (\text{A.2})$$

From thus reconstructed total cross section one only needs to make the scaling between specific reactions of interest and all induced reactions:

$$\sigma_{C,\mathcal{R}}(E_n) = \frac{N_{C,\mathcal{R}}(E_n)}{N_C(E_n)} \sigma_C(E_n) \quad (\text{A.3})$$

in order to reconstruct the cross section of a particular reaction.

## References

- [1] P. Žugec, N. Colonna, D. Bosnar, et al., *Phys. Rev. C* 90 (2014) 021601(R).
- [2] P. Žugec, N. Colonna, D. Bosnar, et al., *Eur. Phys. J. A* 52 (2016) 101.
- [3] P. Žugec, N. Colonna, D. Bosnar, et al., CERN-INTC-2017-081 / INTC-P-522 (2017).
- [4] C. Guerrero, A. Tsinganis, E. Berthoumieux, et al., *Eur. Phys. J. A* 49 (2013) 27.
- [5] C. Weiß, E. Chiaveri, S. Girod, et al., *Nucl. Instrum. Methods Phys. Res. A* 799 (2015) 90.
- [6] N. Colonna, E. Chiaveri, F. Gunsing, *Nucl. Phys. News* 25:4 (2015) 19.
- [7] S. Barros, I. Bergström, V. Vlachoudis, C. Weiß, *J. Instrum.* 10 (2015) P09003.
- [8] M. Barbagallo, C. Guerrero, A. Tsinganis, et al., *Eur. Phys. J. A* 49 (2013) 156.
- [9] M. Sabaté-Gilarte, M. Barbagallo, N. Colonna, et al., *Eur. Phys. J. A* 53 (2017) 210.
- [10] F. Gunsing, O. Aberle, J. Andrzejewski, et al., *Eur. Phys. J. Plus* 131 (2016) 371.
- [11] M. Barbagallo, J. Andrzejewski, M. Mastromarco, et al., *Nucl. Instrum. Methods Phys. Res. A* 887 (2018) 27.
- [12] P. Žugec, C. Weiß, C. Guerrero, et al., *Nucl. Instrum. Methods Phys. Res. A* 812 (2016) 134.
- [13] P. Žugec, M. Barbagallo, J. Andrzejewski, et al., *Nucl. Instrum. Methods Phys. Res. A* 983 (2020) 164606.
- [14] L. Damone, M. Barbagallo, M. Mastromarco, et al., *Phys. Rev. Lett.* 121 (2018) 042701.
- [15] M. Barbagallo, A. Musumarra, L. Cosentino, et al., *Phys. Rev. Lett.* 117 (2016) 152701.
- [16] L. Cosentino, A. Musumarra, M. Barbagallo, et al., *Nucl. Instrum. Methods Phys. Res. A* 830 (2016) 197.
- [17] P. Žugec, M. Barbagallo, J. Andrzejewski, et al., *J. Instrum.* 15 (2020) P02011.
- [18] J. Balibrea-Correa, J. Leredegui-Marco, V. Babiano-Suárez, et al., *Nucl. Instrum. Methods Phys. Res. A* 1001 (2021) 165249.
- [19] V. Babiano-Suárez, J. Leredegui-Marco, J. Balibrea-Correa, et al., *Eur. Phys. J. A* 57 (2021) 197.
- [20] O. Behnke (ed.), K. Kröninger (ed.), G. Schott (ed.), T. Schörner-Sadenius (ed.), *Data Analysis in High Energy Physics*, Wiley-VCH, Weinheim, Germany, 2013.
- [21] S. Avdic, S. A. Pozzi, V. Protopopescu, *Nucl. Instrum. Methods Phys. Res. A* 565 (2006) 742.
- [22] A. Sharghi Ido, M. R. Bonyadi, G. R. Etaati, M. Shahriari, *Appl. Radiat. Isot.* 67 (2009) 1912.
- [23] S. A. Hosseini, *Radiat. Phys. Chem.* 126 (2016) 75.
- [24] A. Adare, C. Aidala, N. N. Ajitanand, et al., *Phys. Rev. C* 93 (2016) 014904.
- [25] C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, et al., *J. Phys. Conf. Ser.* 664 (2015) 072015.
- [26] M. Azhari, A. Abarda, B. Ettaki, et al., *Procedia Comput. Sci.* 170 (2020) 1141.
- [27] K. Lasocha, E. Richter-Was, M. Sadowski, Z. Was, *Phys. Rev. D* 103 (2021) 036003.
- [28] S. Carboni, S. Barlini, L. Bardelli, et al., *Nucl. Instrum. Methods Phys. Res. A* 664 (2012) 251.
- [29] L. Tassan-Got, *Nucl. Instrum. Methods Phys. Res. B* 194 (2002) 503.
- [30] M. Ambriola, R. Bellotti, F. Cafagna, et al., *Nucl. Instrum. Methods Phys. Res. A* 440 (2000) 438.
- [31] J. L. Flores, I. Martel, R. Jiménez, et al., *Nucl. Instrum. Methods Phys. Res. A* 830 (2016) 287.
- [32] M. Alderighi, A. Anzalone, M. Bartolucci, et al., *Comput. Phys. Commun.* 140 (2001) 13.
- [33] P. Abratenko, M. Alrashed, R. An, et al., *Phys. Rev. D* 103 (2021) 092003.
- [34] D. R. Tilley, J. H. Kelley, J. L. Godwin, et al., *Nucl. Phys. A* 745 (2004) 155.
- [35] J. H. Kelley, E. Kwan, J. E. Purcell, et al., *Nucl. Phys. A* 880 (2012) 88.
- [36] J. H. Kelley, J. E. Purcell, C. G. Sheu, *Nucl. Phys. A* 968 (2017) 71.
- [37] F. Ajzenberg-Selove, *Nucl. Phys. A* 523 (1991) 1.