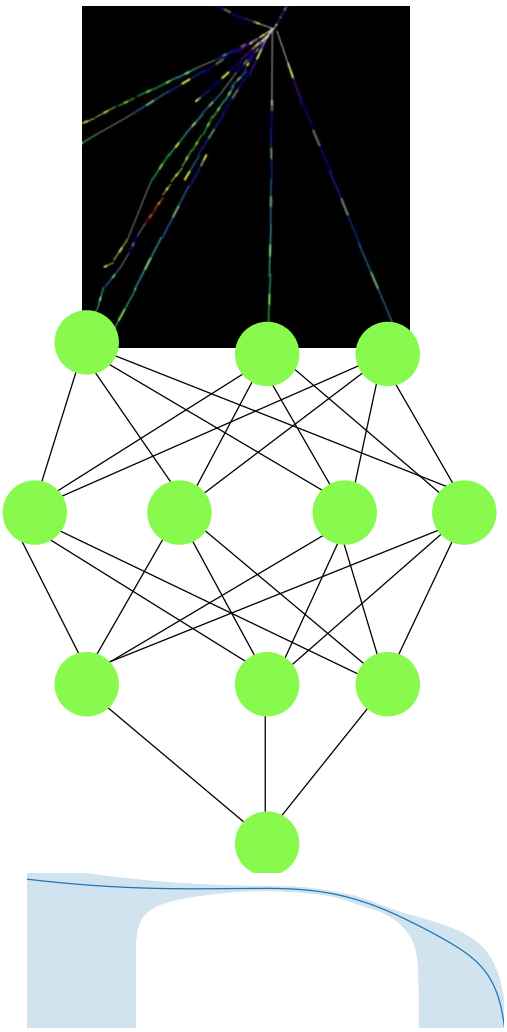
A Determination of Forward Neutrino Fluxes From the LHC with Machine Learning

Jukka John

Master Physics and
Astronomy
Theory track



Supervisor:
Prof. Dr. Juan Rojo
VU & Nikhef

Daily Supervisor:
Peter Krack
VU & Nikhef

Second Examiner:
Dr. Lydia Brenner
Nikhef

15 July 2025

Abstract

In the past two years incredible observations of TeV-energy neutrinos have been made by LHC far-forward detectors SND@LHC and FASER. These observations include the measurement of LHC forward neutrino fluxes as a function of flavour, energy and rapidity, which can be used to constrain theoretical predictions of forward hadron production in proton-proton(pp) collisions. In this work, the FASER neutrino event yield measurements from Run 3 are used to determine the LHC forward neutrino fluxes. This is done in a theory-agnostic manner by employing machine learning (ML) and by exploiting the equivalence with the NNPDF framework, where parton distribution functions (PDFs) are fitted using LHC data and ML. For the first time, the full $E\nu$ dependence on the muon neutrino flux is constrained. This flux is compared to theoretical predictions, showing a disfavour towards the `DPMJET 2019` model. The methodology used in this work is then used for several applications including comparing several forward hadron production models, constraining three BSM enhanced decays of neutral mesons into neutrinos and research the possibility of an intrinsic charm component in the proton.

Summary Popular Science

Neutrinos are incredibly abundant particles, in fact 100 billion billion neutrino fly through your body every single second. Since the late physicist Wolfang Pauli proposed these incredibly light (neutrinos are a million times lighter than the electron) and weak-interacting (this is why you are not noticing anything of the 100 billion billion neutrinos) particles almost a century ago and its first detection in 1956, neutrinos have become to play an important role in several areas of physics: from astrophysics, where neutrinos have been measured coming from the sun and supernovae, to particle physics where they have become important assets in answering questions on the standard model and beyond the standard model physics. Considering the fact that there are so many neutrinos, one might think it is relatively straightforward to detect them, however, this is not the case at all. Actually, only in recent years high-energy neutrinos have been detected at the Large Hadron Collider (LHC) at CERN. The LHC is, as its acronym suggests, a large circular collider at CERN, Geneva. In the LHC, protons are accelerated to almost the speed of light after which they collide and, because they have such a high energy, produce a lot of other particles. Amongst these particles, neutrinos appear and, contrary to charged particles, they escape the detector and the ring because the magnets in the collider are not affecting the trajectory of these neutrinos. Only in the past 2 years a detector has been built outside the LHC where the neutrinos are detected. One may wonder: that is great, but how can these neutrinos be utilized to the physicists' advantage? Using the data from the detector which consists of the number of neutrinos measured as a function of the neutrino energy, also called the neutrino flux, one can employ machine learning(ML) to analyse this data. Machine learning has been a very powerful tool in particle physics to analyse data, and also in other areas of physics and beyond. This is because ML is theory agnostic i.e. it does not assume anything about the underlying theory and the laws of physics when analysing this data. In this work, ML techniques have been used to analyse this new LHC data to derive the underlying physical law governing this data. These new results are also compared to several different models predicting the neutrino flux based on the current physical laws. Using this comparison, some models can be excluded and improved predictions have been made for the neutrino flux. These models are describing the so-called strong force, which also governs the structure of the proton. Since the structure of the proton is still not completely unravelled,

putting constraints on these models describing the strong force can in turn teach us more about the structure of the proton. The ML extraction of the neutrino flux is also helping neutrino telescopes in detecting neutrinos, because they can help to improve predictions of the neutrino fluxes detected by these telescopes. These telescopes are trying to detect neutrinos coming from outer space. One example of such a neutrino telescope is called IceCube, which is trying to detect ultra-high-energy cosmic neutrinos and is a located near the south pole deep in the ice.
All in all, this work is using neutrino measurements from the LHC and machine learning to deepen our understanding of the strong force, refine models of proton structure, and improve the predictive capabilities of neutrino observatories like IceCube in their quest to detect cosmic neutrinos from the far reaches of the universe.

Contents

"Op dit moment gaat een tipje van de sluier omhoog die over het hele leven ligt: dat ik altijd en in alles weerloos, machteloos en vervangbaar als een atoom ben en dat alle bewustzijn, alle wil, hoop en vrees alleen maar manifestaties zijn van het mechanisme waarvolgens de menselijke moleculen zich bewegen in de peilloze kosmische materiedamp."

- *Willem Frederik Hermans, Nooit meer slapen.*

1 Introduction

100 billion billion. This mind-boggling number is how many neutrinos pass through your body every single second. One might conclude from this that neutrinos are all over the place and thus abundant, which is correct. However, one would not be correct in stating that neutrinos are therefore easy to detect in any sort of way. This is because of the properties of neutrinos, e.g. their mass is incredibly small compared to the lightest standard model particles and they only interact via the so-called weak force[1]. Because of these properties it is incredibly hard to detect neutrinos, in fact, only in the past 2 years observations at the Large Hadron Collider (LHC) by the FASER/FASERν[2, 3] and SND@LHC[4] experiments have been made, which heralded the beginning of collider neutrino physics and LHC neutrinos.

These high-energy neutrinos originate from proton-proton (pp) collisions at the ATLAS Interaction Point (IP) at the LHC. After the initial pp collision, the outgoing forward(small angle wrt beam axis) hadrons are not affected by the LHC magnets and eventually decay into neutrinos. However, these particles going in the forward direction are not be measured. This is why the promising Forward Physics Facility (FPF) has been proposed, including FASERν2, which is an upgrade to the current FASERν detector, to take full advantage of these LHC neutrinos, because it promises to measure the particles going in the forward direction[5]. This FPF is also promising, since all the particles and events that the several experiments will measure come for free from the pp collisions at the ATLAS IP. It is expected that around 100,000 electron neutrinos will be measured. These events, in combination with other measurements at the FPF will provide constraints on a variety of physical models [6]. In the meantime, the FASER and the SND@LHC detector continue the effort of detecting LHC neutrinos at TeV energies and its scattering rates.

One might wonder why one wants to measure these neutrinos and what insights it can offer. Because the LHC neutrinos originate from the pp collisions at the ATLAS IP, they provide us with valuable information on these collisions. Also, measuring these neutrinos not only provides insight into the interactions of neutrinos at TeV energies, but also into the still not completely unravelled proton structure and the strong force which is exerted inside the proton and described by Quantum Chromodynamics (QCD), or, in other words, the measurements of these neutrinos can put constraints on several QCD models. Because the neutrino fluxes can validate predictions for forward hadron D-meson production, they can indirectly also be used to improve predictions for the high-energy neutrino fluxes at neutrino telescopes such as IceCube[7] and KM3NET[8]. On top of this, new kinematic regions of the strong interaction can be explored. Figure 1.2 shows the connection between collider neutrino

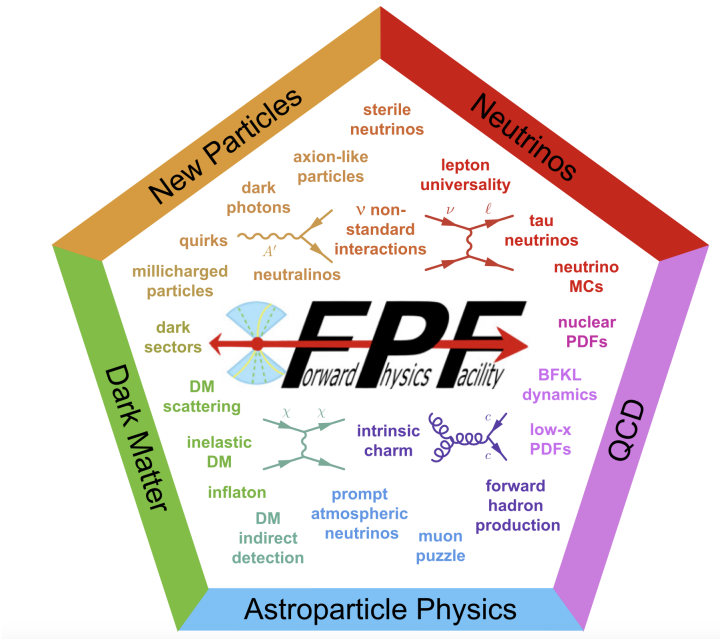


Figure 1.1. This image shows the full scope of the physical applications of the proposed Forward Physics Facility (FPF) at the LHC. The FPF would be located 620 m downstream of the ATLAS IP and is proposed to take advantage of particles produced from pp collisions and which fly in the forward detector [6]

.

flux measurements and QCD and astrophysics, showing an overview of all the possible applications of neutrino flux measurements.
[9]. Figure 1.1 shows the full potential of the FPF. In light of these applications as well as to improve predictions for the proposed FPF, this work will present how to use the recent event rate measurements from FASER and SND@LHC, as well as measurements from the proposed FPF, to perform a theory agnostic parametrization of the LHC forward neutrino fluxes by using feed-forward neural networks. Theory agnosticism is important, since one should be able to compare the parametrization to existing QCD calculations of forward particle production at the LHC, specifically from charm hadron decays. In this work the methodology developed will be applied to the differential measurement of muon neutrino event rates from FASER[3] to constrain for the first time the full E_ν dependence of the muon neutrino flux. The FASER collaboration performed an analysis of comparing the muon neutrino event rate measurements to different Monte Carlo (MC) event generators of forward hadron production [3]. In this work a similar analysis will be performed by comparing the parametrised neutrino flux to predictions of neutrino fluxes computed by several event generators. The parametrised neutrino flux is a suitable quantity to provide constraints on theoretical models of forward hadron production at the LHC, because it can provide unique constraints on these models using statistical compatibility. Next, the methodology will also be used to research several physics opportunities: First of all it will be used to discriminate between several event generators for forward hadron production. It will also be used to fit neutrino fluxes for the FASERν experiment at run-3 (which is the current setup) and for run 4+5, which is the high luminosity era at the LHC, which provides more performance and thus in the end more neutrino scattering event measurements. The proposed FASERν2 detector at the FPF will also be used in this work to fit neutrino fluxes. Furthermore, the method will be used to constrain BSM(beyond the standard model) scenarios predicting more neutrinos by so-called BSM decays of neutral hadrons into neutrinos for both current and future detector setups. Finally, the possibility of an intrinsic charm component in the proton(which is heavier than the proton itself) in the proton will

be researched for the current setup and the proposed FASERν2 detector at the FPF. This possibility would again lead to an excess of neutrinos, which would be visible in the neutrino flux. In this work, electron and muon fluxes are considered, with the note that the methodology can also easily be applied to tau neutrinos.

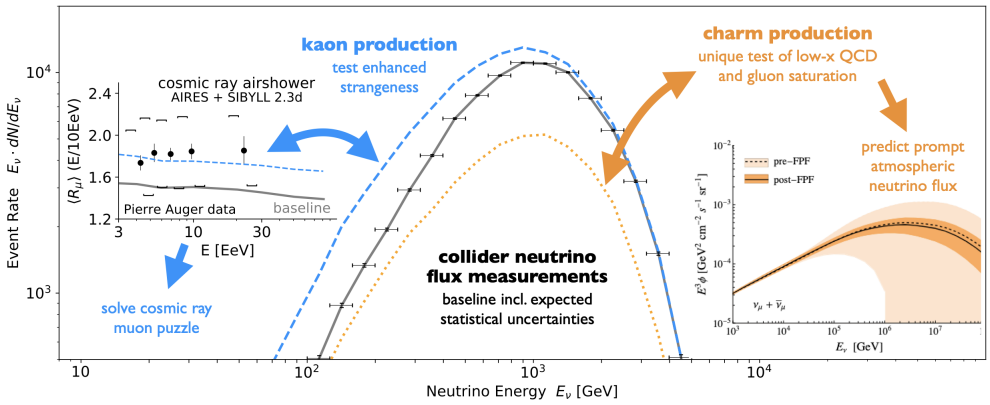


Figure 1.2. This image shows the connection between collider neutrino flux measurements and its applications to QCD, prompt atmospheric neutrino fluxes and astrophysics[9]

.

In order to achieve the goal of accurately parametrizing the neutrino flux using Machine Learning(ML), the equivalence between determining neutrino fluxes from neutrino scattering event rates and parton distribution functions (PDF) from Deep Inelastic Scattering (DIS) structure functions is exploited[10, 11]. These PDFs describe the individual constituents of a proton called partons and these functions describe the contribution of each parton species with a certain momentum fraction to the proton i.e. the probability for a parton species to carry a certain momentum fraction of the proton[12]. These PDFs can only be determined from experimental data, and not from a-priori theory. Figure 1.3 shows an example of the proton PDFs. It can be seen that around x = 03(the fraction of the proton's momentum) the valence quarks dominate, so each of them carry around a third of the proton's momentum.

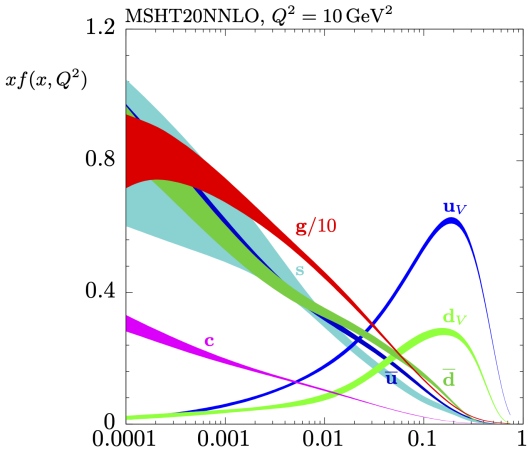


Figure 1.3. An example of proton PDFs. It can be seen that around x = 0.2 the valence quarks dominate, so each of them carry around a third of the proton's momentum. The Q^2 is the energy scale and the x is the fraction of the proton's momentum.

Both in this work, and in the determination of PDFs from data, the data is related to the PDFs via a convolution with a Fast Kernel(FK)-table, which encodes a theoretical description of observables[13, 14]. In the calculation of PDF fits from DIS, the FK-table contains information on partonic cross-

sections[15]. In this thesis the event rates measured at FASER are expressed as a convolution(matrix multiplication) between the neutrino flux and neutrino-nucleus interaction cross-section. The latter is assumed to be reliably computed by perturbative QCD and by the Standard Model (SM).
The fitting of PDFs from DIS structure functions and data has been extensively implemented in a framework called Neural Network Parton Distribution Function (NNPDF) which, as its acronym suggests, determines PDFs using Machine Learning[13, 16]. In this thesis close attention is paid to this framework and its implementation on how to fit these PDFs. This is because NNPDF provides a robust framework in employing statistical methods and validation of the results i.e. by employing closure tests of the results(this will be explained below).
In short, this work will present how to use event rate measurements from FASER to perform a model-independent parametrization of the LHC forward muon neutrino fluxes by using ML techniques. Using the framework presented in this work, several physics applications will be researched i.e. the framework will be used to extract electron and muon neutrino fluxes for different detectors and LHC runs, several event generators, three BSM decays and the possibility of intrinsic charm in the proton.
The thesis is structured as follows: first a theoretical framework is presented in Section 2 where DIS scattering is explained, alongside neutrino fluxes and event rates. Also, a formal explanation of FK-tables is presented and a more detailed explanation of the equivalence between parametrizing LHC neutrino fluxes and determining PDFs from DIS structure functions can be found in this section. Furthermore, the machine learning approach will be discussed here. Section 3 will be on the data used in this work, explaining both the rough workings of an event generator, which were used to generate data and a brief explanation of how the event rates were obtained from the FASERν experiment. Then the framework developed in this work will be explained in great detail in Section 4. Here, the process of building the framework will be described step-by-step and the methodology will be validated. The section will conclude with the muon neutrino flux extracted from the recent FASER event rate measurements. In Section 5 the framework will be used for the applications described above. The penultimate section is on transforming this framework from code written by physicists to a well-written python package. Finally, a summary and an outlook of this work are provided.

2 Theoretical framework

In this section the necessary background knowledge needed for this thesis will be explained in great detail. First of all the DIS scattering kinematics of neutrinos at the LHC far forward detectors, such as FASER, will be covered. The DIS variables will be related to event rate measurements, which can in turn be related to neutrino fluxes. Furthermore, the FK-table formalism will be discussed, which is necessary to express these relations in a numerically efficient manner, where the equivalence with determining PDFs from DIS structure functions will be highlighted. Finally, the neural net implementation to parametrize the neutrino flux in a theory agnostic manner will be presented.

2.1 Neutrino DIS Scattering Kinematics

At the LHC, and more specifically at the ATLAS IP, high-energy protons (E_{CoM} = 14TeV) collide and a high energy sea of quarks and gluons come flying out. The theory describing the structure of the proton is QCD, because the quarks and gluons inside the proton are bound together by the strong force. The strong force has a certain strength which can be quantified using the coupling constant $\alpha_s \sim \frac{1}{\ln(Q^2)}$. As can be seen from this equation, the coupling constant is not constant at all and is dependent on the momentum transfer Q^2. This is called the running of the coupling constant and it means that the strength decreases when the momentum transfer increases. As the strength of the coupling decreases, the partons inside the proton can be considered asymptotically free, which means they can be treated as free, point-like objects. At the interaction point at the LHC, the collisions happen at around 14TeV, which means the energy is high enough for the partons to be considered asymptotically free. In this region, perturbative QCD can be used to describe QCD interactions[17]. The strength of the coupling constant is high at lower energies, which means the partons are no longer asymptotically free and are in bound states i.e. they get confined inside hadrons. This is called hadronization which happens after the proton-proton collision has taken place. The forward hadrons eventually decay into electron, muon and tau neutrinos [18], where these neutrinos reach the FASER detector and react via a tungsten target. Figure 2.1 shows the journey of the forward travelling neutrinos.

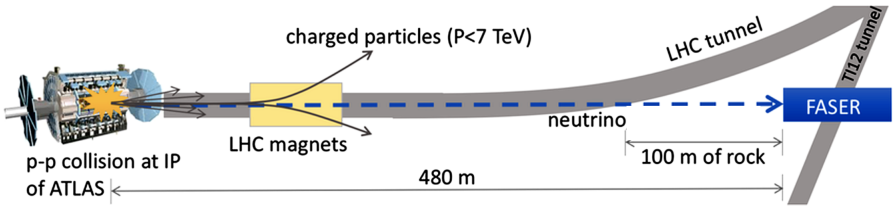


Figure 2.1. This figure shows the LHC and the ATLAS IP and FASER detector. At the ATLAS IP protons collide where the created neutrinos travel in the forward direction, eventually arriving at the FASER detector. Before the FASER detector there is 100 m of rock.[19]

2.1.1 DIS Neutrino Scattering

The neutrinos react with the tungsten target via a charged current DIS process, mediated by a $W^{\pm}$ boson.

$$\nu_i(\bar{\nu}_i) + N(p_N) \xrightarrow{W^{\pm}} \ell^-(\ell^+) + X_h(E_h), \quad i = e, \mu, \tau \quad \text{(charged current)}. \tag{2.1.1}$$

This process is called DIS, because the incoming neutrino probes deep into the tungsten target, and strikes a quark via an inelastic scattering[20]. The Feynman diagram of this process is shown in Figure 2.2.

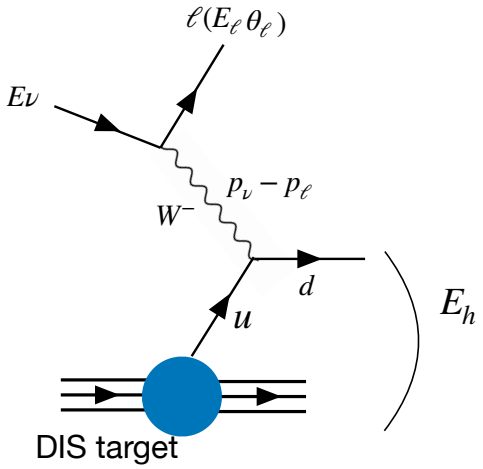


Figure 2.2. This Feynman diagram shows how the incoming neutrino with energy E_ν react with a tungsten target. This charged current DIS process is mediated by a $W^\pm$ boson. After the collision a generic final hadronic state with energy E_h and a lepton with energy E_ℓ and scattering angle θ_ℓ are produced. [19]

This DIS process can be described by three final-state variables: the lepton scattering angle θ_ℓ, the charged lepton energy E_ℓ and the total energy of the final hadronic state E_h. It is assumed that these variables can be measured with good resolution by the detectors such as FASERν. Using these kinematic variables, the DIS variables Q^2 (momentum transfer), Bjorken-x (Bjorken-x is the fraction of the proton's momentum carried by the struck quark) and the neutrino energy E_ν, can be reconstructed by the following equations:

$$\begin{aligned} E_\nu &= E_h + E_\ell \\ Q^2 &= 4(E_h + E_\ell)E_\ell \sin^2(\theta_\ell/2) \\ x &= \frac{4(E_h + E_\ell)E_\ell \sin^2(\theta_\ell/2)}{2m_N E_h}, \end{aligned} \tag{2.1.2}$$

where m_N is the nucleon mass. As one can see, the measurement of $(E_\ell, \theta_\ell, E_h)$ can reconstruct the DIS variables (E_ν, Q^2, x). Another common, alternative and equivalent, representation of the DIS variables are (Q^2, x, y), where y is the fractional energy loss of the incoming particle and is defined as

$$y = \frac{Q^2}{2xm_N E_\nu} \text{[21]}.$$

2.1.2 Dependence of DIS Variables on Neutrino Rapidity

For completeness, the dependence of the neutrino rapidity on the DIS kinematics is discussed in this section as well as the relation to the experimental setups of the FASERν and the SND@LHC detector. The rapidity is a measure which relates the energy of the neutrino and the neutrinos momentum in the z-direction. A difference in rapidity is Lorentz invariant, which is why it is a convenient measure[22].

It is defined as:

$$y_\nu = \frac{1}{2}\ln\left(\frac{E_\nu + p_{\nu,z}}{E_\nu - p_{\nu,z}}\right), \tag{2.1.3}$$

where the neutrino four-momentum is defined as:

$$p^\mu_\nu = (E_\nu, \vec{p}_{\nu,\mathrm{T}}, p_{\nu,z}), \tag{2.1.4}$$

where the momentum in the z-component, which is aligned with the beam direction, and the transverse component, which is perpendicular to the beam direction, are separated. In the case of massless particles, like neutrinos, the rapidity is equal to the pseudo rapidity. The pseudo rapidity is a measure of the angle of the neutrino with respect to the beam axis and is defined as

$$y_\nu = \eta_\nu = -\ln\left(\tan\frac{\theta_\nu}{2}\right), \tag{2.1.5}$$

where θ_ν signifies the angle of the neutrino momentum relative to the beam axis[23].
The attentive reader might have noticed the four-momentum in equation 2.1.4 is not the one assumed at the LHC where the neutrino is aligned with the line-of-sight, given by

$$p^\mu_\nu = \left(E_\nu, \vec{0}, p_{\nu,z} = E_\nu\right). \tag{2.1.6}$$

Because of this, reconstructing the neutrino kinematics is not as easy and for example Q^2 is dependent on the neutrino rapidity. To show this, the four momenta of the initial state are (considering the CC DIS process)

$$p^\mu_\nu = \left(E_\nu, \vec{p}_{\nu,\mathrm{T}}, p_{\nu,z}\right) = (E_\nu, E_\nu\sin\theta_\nu, E_\nu\cos\theta_\nu), \qquad p^\mu_N = (m_N, \vec{0}, 0), \tag{2.1.7}$$

and the momenta of the final state variables are

$$p^\mu_\ell = (E_\ell, E_\ell\sin\theta_\ell\cos\phi_\ell, E_\ell\sin\theta_\ell\sin\phi_\ell, E_\ell\cos\theta_\ell), \qquad p^\mu_h = (E_h, \vec{p}_{h,T}, p_{h,z}). \tag{2.1.8}$$

The momentum transfer is now given by

$$Q^2 = -q^2 = -(p^\mu_\nu - p^\mu_\ell)^2 = 2E_\nu E_\ell\left(1 - \sin\theta_\nu\sin\theta_\ell - \cos\theta_\nu\cos\theta_\ell\right), \tag{2.1.9}$$

which obviously depends on neutrino (pseudo) rapidity.
For line-of-sight (LoS) detectors, such as FASERν accounting for the neutrino rapidity is a small effect. This is because it can be assumed that the transverse momentum carried by the neutrino is much smaller than the transverse momentum carried away by the final state particles. For off-axis detectors, such as SND@LHC this assumption cannot be made.
For the FASERν detector, the rapidity coverage is

$$\eta^{(\max)}_\nu = -\ln\left(\tan\frac{\theta^{(\max)}_\nu}{2}\right) \simeq 8.5, \tag{2.1.10}$$

which means that this detector can detect neutrinos with rapidities $\eta_\nu \gtrsim \eta^{(\max)}_\nu$. This corresponds to scattering angles of $\theta_\nu \lesssim \theta^{(\max)}_\nu \simeq 4\times10^{-4}$. The longitudinal and transverse components of the massless

neutrinos can be computed as follows:

$$p_z = E\cos\theta\,, \qquad p_T = E\sin\theta \tag{2.1.11}$$

The transverse momentum of the neutrinos that fall within the FASERν detector acceptance is

$$p_T^{\nu} \lesssim \left(E_{\nu}^{\max} \times \sin\theta_{\nu}^{(\max)}\right) \sim 0.8 \text{ GeV}, \tag{2.1.12}$$

which shows that we can assume the transverse momentum of the neutrino is negligible compared to the longitudinal momentum when evaluating the DIS scattering cross-section.
In conclusion, we assume that the neutrino detector can measure the neutrino rapidity η_ν, E_ℓ, θ_ℓ and E_h. In this work we also assume LoS detectors, which can be used to reconstruct the events by the DIS variables $\left(Q^2, x, E_\nu\right)$.The neutrino fluxes are parametrized using a single variable $(E\nu, E_\ell, \theta_\ell, E_h)$, and the dependence on multi-differential parametrizations i.e. including rapidity dependence in the flux is left for future work.

2.2 Event Rate Calculations

After having discussed the DIS kinematics and variables, the following section is on the event rate calculations and its relation to neutrino fluxes, because the goal of this work is to parametrize the neutrino flux using event rate measurements.

Neutrino Flux The neutrino fluxes are defined as

$$\frac{dN_{\nu_i}(E_\nu)}{dE_\nu}, \tag{2.2.1}$$

where one can see that the flux is binned in E_ν, the neutrino energy. Physically, this flux signifies the number of neutrinos passing through the detector for a given E_ν i.e. in a certain neutrino energy bin there are a certain number of neutrinos which is dependent on the detector.

Event Rate Calculations Having defined the binned neutrino flux, we want to express the number of neutrino CC scattering events as a function of this flux. This will be expressed as a function of the DIS variables

$$\left(x, Q^2, E_\nu\right) \tag{2.2.2}$$

For completeness, the event rates can also, completely analogously, be expressed in the observable quantities

$$\left(E_h, E_\ell, \theta_\ell\right), \tag{2.2.3}$$

which will be exploited in Section 5. The number of CC neutrino scattering events are binned in these three variables and can be expressed as:

$$N_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu) \equiv N_{\text{int}}^{(\nu_i)}\left(x_{\min} \le x \le x_{\max}, Q^2_{\min} \le Q^2 \le Q^2_{\max}, E_\nu^{\min} \le E_\nu \le E_\nu^{\max}\right). \tag{2.2.4}$$

As one can see, measuring the dependence of E_ν on the number of events will provide information on the incoming neutrino fluxes. In section 3 on the data used in the ML training, the optimal binning of the event rates as a function of E_ν will be discussed. The number of events is dependent on the

parametrisation of the neutrino flux and the relation between the two is as follows:

$$N_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu) = \int_{Q^2_{\min}}^{Q^2_{\max}} dQ^2 \int_{x_{\min}}^{x_{\max}} dx \int_{E_\nu^{(\min)}}^{E_\nu^{(\max)}} dE_\nu \widetilde{N}_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu)\,, \tag{2.2.5}$$

where the integrand is defined as:

$$\widetilde{N}_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu) \equiv n_T L_T \times \frac{dN_{\nu_i}(E_\nu)}{dE_\nu} \times \frac{d^2\sigma^{\nu_i A}(x, Q^2, E_\nu)}{dx dQ^2} \times \mathcal{A}(E_\ell, \theta_\ell, E_h)\,, \tag{2.2.6}$$

where

- n_T and L_T is the atomic density of the target material (Tungsten) and the length of the detector, respectively.
- $\frac{dN_{\nu_i}(E_\nu)}{dE_\nu}$ is the neutrino flux.
- The third term is the double differential DIS cross section which contains information on the neutrino-target interactions. This is given by

$$\frac{d^2\sigma^{\nu_i A}(x, Q^2, E_\nu)}{dx dQ^2} = \frac{G_F^2}{4\pi x \left(1 + Q^2/m_W^2\right)^2} \left[Y_+ F_2^{\nu p}(x, Q^2) - y^2 F_L^{\nu p}(x, Q^2) + Y_- x F_3^{\nu p}(x, Q^2)\right], \tag{2.2.7}$$

  where G_F^2 is the Fermi coupling constant, m_W the boson mass and where $Y_\pm$ is defined as $1 \pm (1 - y)^2$. The F terms in the square brackets are so-called structure functions. These functions encode the partonic structure of hadrons and thus the partonic behaviour in the scattering process. The neutrino DIS structure functions are given by

$$F_i^{\nu p}(x, Q^2) = \sum_{j=q,\bar{q},g} \int_x^1 \frac{dz}{z} C_{i,j}^{\nu N}\left(z, \alpha_s(Q^2)\right) f_j^{(p)}\left(\frac{x}{z}, Q^2\right), \quad i = 2, 3, L, \tag{2.2.8}$$

  where $f_j^{(p)}\left(\frac{x}{z}, Q^2\right)$ is the process-independent proton PDF mentioned in the introduction and $C_{i,j}^{\nu N}\left(z, \alpha_s(Q^2)\right)$ are the partonic coefficient functions. These coefficient functions are process dependent and describe the probability amplitude for each parton species to contribute to the structure function.[21]
  The double-differential DIS cross section in this work is computed up to next-to-leading-order (NLO) in QCD using the event generator POWHEG and the parton shower (PS) PYTHIA8, which together model the neutrino scatterings.

- $\mathcal{A}(E_\ell, \theta_\ell, E_h)$ is an acceptance factor modelling the acceptance region of the detector by imposing step functions. This term is evaluated by a Monte Carlo (MC) simulation, which will be explained in more detail below.

- The event rate measurements also depend on the integrated luminosity $\mathcal{L}$ of the pp collisions at the ATLAS IP. This is expressed in inverse femtobar(fb^{-1}) or inverse attobar(ab^{-1}). A higher integrated luminosity means more scattering events. Currently, the integrated luminosity for LHC Run-3 is 150 fb^{-1} and will be 3 ab^{-1} in the future in the so-called high luminosity era.

The double differential cross section and the acceptance factor can be combined into one, because they are both evaluated by the event generator. The combined term is denoted by $\left.\frac{d^2\sigma_{\text{NLO+PS}}^{\nu_i A}(x, Q^2, E_\nu)}{dx dQ^2}\right|_{\text{fid}}$,

where fid stands for the fiducial volume of the detector.

Binning of Events As mentioned above, the event rates are binned in three kinematic variables. Since the dependence of the DIS cross-section on x and Q^2 is assumed to be well-described by QCD the binned event rates will be given by

$$N_{\text{int}}^{(\nu_i)}(E_\nu, y_\nu) \equiv N_{\text{int}}^{(\nu_i)}\left(E_\nu^{(\text{min})} \leq E_\nu \leq E_\nu^{(\text{max})}\right), \tag{2.2.9}$$

and the integral now looks like

$$N_{\text{int}}^{(\nu_i)}(E_\nu) = \int_{E_\nu^{(\text{min})}}^{E_\nu^{(\text{max})}} dE_\nu, \int_{Q_0^2}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^{1} dx \widetilde{N}_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu), \tag{2.2.10}$$

where x and Q^2 are integrated over their kinematically allowed region. The optimal binning for the event rates depends on event statistics and the acceptance region of the detector. In Section 3 the specific choice of binning and the binning criteria are explained. The event rates in this equation are expressed as a function of E_ν, however, as mentioned above, they can also be expressed as a function of the $(E_h, E_\ell, \theta_\ell)$ kinematic variables.
As mentioned before, the goal of this work is to parametrize the neutrino flux using machine learning and training on event rate measurements. As one can see from the aforementioned equations, these two are related via a triple convolution. When using ML to parametrize the neutrino flux, one guesses a new shape for the flux at every iteration, thus having to perform these integrals and evaluating the DIS differential cross-section at each iteration. This is very time consuming and computationally expensive, since a ML training can easily comprise of a 1000 iterations. To shape a picture as to how time-consuming this task would be: evaluating the DIS and integral would take around a minute, thus one would have to wait ~ 16 hours to train a simple neural network. Since this is not feasible, the computationally efficient solution to this problem, called the FK-table formalism, will be explained below.

2.3 Fast Kernel-Table Formalism

The underlying idea of the FK-table formalism is to replace the three integrals(convolutions) in equation 2.2.10 by a single matrix to massively increase the computational efficiency of parametrizing the neutrino flux using ML. This matrix, stored in a precomputed fast interpolation table, will encapsulate the DIS cross section, acceptance factors and the binning. As mentioned in the introduction, this approach is equivalent to determining PDFs from DIS structure functions[24].
FK-tables in NNPDF The underlying idea in the NNPDF framework is to determine PDFs from double-differential cross sections, which are defined in terms of structure functions, which can be seen in equation 2.2.7. Starting from equation 2.2.8, the F_2 structure function can be expressed as:

$$F_2(x, Q^2) = \sum_{i,j}^{n_f} C_i(x, Q^2) \otimes \Gamma_{ij}(Q^2, Q_0^2) \otimes f_j(x, Q_0), \tag{2.3.1}$$

where the $\otimes$ replaced the integral and is called a Mellin convolution. $\Gamma_{ij}(Q^2, Q_0^2)$ is an evolution operator, which evolves the PDF from scale Q_0^2 to Q^2. The F_2 structure function is one of the observables that NNPDF uses to fit PDFs[13]. At each iteration the so-called Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations have to be solved to evolve the PDF from Q_0^2 to Q^2 and the convolution

also has to be solved at every iteration[13, 25]. As explained above, this is not feasible. This is why the coefficient functions and the evolution operator are precomputed and stored in an interpolation basis. The PDF can be expressed in an interpolation basis such as

$$f_i(x, Q^2) = \sum_\beta \sum_\tau f_i(x_\beta Q_\tau^2) I_\beta(x) I_\tau(Q^2), \tag{2.3.2}$$

where I_β are the interpolating functions and x_β and Q_τ^2 make up the new interpolation grid and are called grid nodes. Using this, the PDF evaluated at Q_0^2 can be related to this interpolated PDF using the evolution operator Γ in the following way:

$$f_{i,\beta\tau} = \sum_j \sum_\alpha \Gamma^\tau_{ij,\alpha\beta} f_j(x_\alpha, Q_0^2). \tag{2.3.3}$$

Now, the coefficient functions and the interpolation operator can be put into a precomputed grid:

$$FK_{i,\alpha}\left(x, x_\alpha, Q^2, Q_0^2\right) = C_i(x, Q^2) \sum_{\beta,\tau} \sum_j \Gamma^\tau_{ij,\alpha\beta} I_\beta(x) I_\tau(Q^2), \tag{2.3.4}$$

and the F_2 structure function is then given by

$$F_2(x, Q^2) = \sum_i^{n_f} \sum_\alpha^{n_x} FK_{i,\alpha}\left(x, x_\alpha, Q^2, Q_0^2\right) f_i(x, Q_0^2). \tag{2.3.5}$$

As can be seen the convolutions are replaced by matrix multiplications, thus drastically increasing computational efficiency[15],[26]. This is the underlying idea. Now, the specific implementation of this framework and idea will be worked out for the case of event rate measurements and parametrizing the neutrino flux.

Example of an Interpolation To make sure one properly understands the interpolation a concrete example will be given.
Suppose one wants to interpolate the function $f(x) = \sin(x)$ one has to go through 3 steps

1. Compute function values.
   In this case $f(0) = 0,\ f(\pi/4) = \sqrt(2)/2,\ f(\pi/2) = 1$

2. Compute the basis polynomials.
   In this case the Lagrange polynomials are used which are given by:

$$I_\alpha(z) = L_\alpha(z) = \prod_{\substack{\beta=1 \\ \beta\neq\alpha}}^{n_x} \frac{z - z_j}{z_i - z_j}. \tag{2.3.6}$$

   Where the first polynomial is given by: $L_0(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{(x-\pi/4)(x-\pi/2)}{\pi^2/8}$ and similarly $L_1(x) = \frac{16x(x-\pi/2)}{\pi^2}$, $L_1(x) = \frac{x(x-\pi/4)}{\pi^2/8}$

3. Constructing the interpolating polynomial.
   The polynomial is given by:

$$f(x) \approx f(x_0)L_0(x) + f(x_1)L_1(x) + f(x_2)L_2(x) = \sqrt{2}/2 * L_1(x) + L_2(x). \tag{2.3.7}$$

   It can be seen that the function is evaluated at the grid nodes $0, \pi/4, \pi/2$ i.e. an array of constant

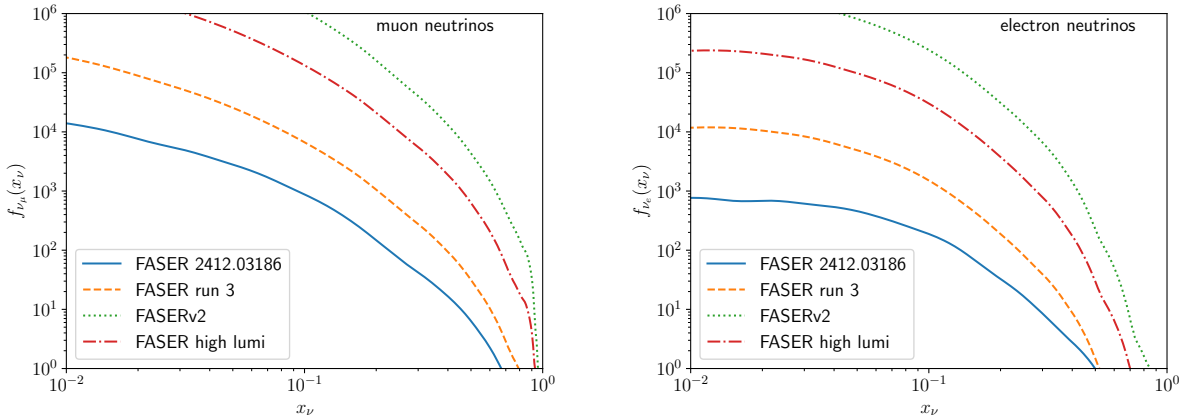


Figure 2.3. This figure shows electron and muon neutrino PDFs $f_{\nu_i}(x_\nu)$ for different FASER detector geometries.

values and the x dependence has been moved to the Lagrange polynomials.

FK-tables in This Work. In this work the FK-table will store the DIS cross-section, the acceptance factors and the binning in a precomputed grid and the integrals will be replaced by a matrix multiplication. In order to arrive at this result, let me first define the 'neutrino PDF' as

$$f_{\nu_i}(x_\nu) = \frac{s_{pp}}{2}\frac{dN_{\nu_i}(E_\nu)}{dE_\nu}, \quad i = e, \nu, \tau, \tag{2.3.8}$$

where x_ν is the momentum fraction and is defined as

$$x_\nu = \frac{2E\nu}{s_{pp}}, \quad 0 \leq x_\nu \leq 1, \tag{2.3.9}$$

where s_{pp} is the maximal center-of-mass energy at which the neutrinos could have been produced, which is equal to 13.6TeV for Run-3. This neutrino PDF is the normalised version of the neutrino flux[21]. Figure 2.3 shows muon and anti-muon neutrino PDFs for several FASER experiments to one an idea of what we are dealing with. The muon neutrino PDF is higher than the electron neutrino PDF. This is because there are more hadrons after hadronization decaying into muon neutrinos than into electron neutrinos. Muon neutrinos are produced from pion and kaon decays, electron neutrinos from charmed hadron decays. Since the first decay mechanism is much more likely to happen than the latter decay mechanism, the muon neutrino flux is higher compared to the electron neutrino flux. The tau neutrinos are solely produced from heavy hadron decays, which makes tau neutrinos quite rare. The name and definition of the neutrino PDF is, of course, derived from the problem of determining PDFs from DIS cross-sections. As can be seen from this equation and equation 7.2.4, also the neutrino PDF is dependent on the DIS cross-section, thus once again confirming the equivalence. Also the momentum fraction is defined in equivalence with PDFs as can be seen in equation 2.2.8. As is the case with PDFs, neutrino PDFs are also formatted in the LHAPDF framework, which stores the dependence of PDFs on energy and rapidity [27]. Both neutrino PDF and neutrino flux will be used throughout this work, they both refer to equation 2.3.8.
Now, the neutrino interaction event rates can be expressed in terms of this neutrino PDF

$$\widehat{N}_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu) \equiv n_T L_T \times \frac{2}{\sqrt{s_{\text{pp}}}} f_{\nu_i}(x_\nu) \times \left.\frac{d^2\sigma_{\text{NLO+PS}}^{\nu_i A}(x, Q^2, E_\nu)}{dx dQ^2}\right|_{\text{fid}}. \tag{2.3.10}$$

As explained above, the PDF, in this case the neutrino PDF, is expressed in terms of a linear expansion in an interpolation basis $\{I_\alpha(x)\}$:

$$f_{\nu_i}(x_\nu) \simeq \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) I_\alpha(x_\nu)\,, \tag{2.3.11}$$

where $x_{\nu,\alpha}$ are the grid nodes and n_x the number of grid nodes. One can choose several interpolation bases depending on the specific problem at hand. In this work the Lagrange polynomials are chosen, where the basis polynomials are given by

$$I_\alpha(z) = L_\alpha(z) = \prod_{\substack{\beta=1 \\ \beta\neq\alpha}}^{n_x} \frac{z - z_j}{z_i - z_j}\,. \tag{2.3.12}$$

So, for each node there is a specific polynomial interpolating the value of the nuetrino PDF on that grid node. Now the interpolated neutrino PDF 2.3.11 is plugged into equation 2.3.10, yielding

$$\begin{aligned}\widehat{N}_{\text{int}}^{(\nu_i)}(x, Q^2, E_\nu) &= n_T L_T \times \frac{2}{\sqrt{s_{\text{pp}}}} \left(\sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) I_\alpha(x_\nu) \right) \times \left. \frac{d^2\sigma^{\nu_i A}_{\text{NLO+PS}}(x, Q^2, E_\nu)}{dx dQ^2} \right|_{\text{fid}} \\ &= n_T L_T \times \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) \times \frac{2}{\sqrt{s_{\text{pp}}}} I_\alpha(x_\nu) \times \left. \frac{d^2\sigma^{\nu_i A}_{\text{NLO+PS}}(x, Q^2, E_\nu)}{dx dQ^2} \right|_{\text{fid}}\,.\end{aligned} \tag{2.3.13}$$

To show how the integrals are replaced by a matrix multiplication, equation 2.3.13 is plugged into equation 7.2.3, yielding:

$$N_{\text{int}}^{(\nu_i)}(E_\nu) = \int_{E_\nu^{(\min)}}^{E_\nu^{(\max)}} dE_\nu \int_{Q_0^2=1\,\text{GeV}^2}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^{1} dx \left[n_T L_T \times \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) \times \frac{2}{\sqrt{s_{\text{pp}}}} I_\alpha(x_\nu) \times \left. \frac{d^2\sigma^{\nu_i A}_{\text{NLO+PS}}(x, Q^2, E_\nu)}{dx dQ^2} \right|_{\text{fid}} \right].$$

Now, one has to realise that the neutrino PDF is evaluated at fixed grid nodes $\{f_{\nu_i}(x_{\nu,\alpha})\}$, thus it has become an array of constant values and can thus be moved outside of the integral. Because of this, the neutrino interaction event rates can be expressed as

$$N_{\text{int}}^{(\nu_i)}(E_\nu) = \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) \int_{E_\nu^{(\min)}}^{E_\nu^{(\max)}} dE_\nu \int_{Q_0^2=1\,\text{GeV}^2}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^{1} dx \left[n_T L_T \times \frac{2}{\sqrt{s_{\text{pp}}}} I_\alpha(x_\nu) \times \left. \frac{d^2\sigma^{\nu_i A}_{\text{NLO+PS}}(x, Q^2, E_\nu)}{dx dQ^2} \right|_{\text{fid}} \right],$$

where the neutrino PDF has been moved in front of the integral and has been replaced by the interpolation basis functions $\{I_\alpha(x_\nu)\}$. Now, it is finally time to define the FK-table as

$$\text{FK}_{\alpha,j} \equiv \frac{2 n_T L_T}{\sqrt{s_{\text{pp}}}} \int_{E_{\nu,j}^{(\min)}}^{E_{\nu,j}^{(\max)}} dE_\nu \int_{Q_0^2}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^{1} dx \left[I_\alpha(x_\nu) \left. \frac{d^2\sigma^{\nu_i A}_{\text{NLO+PS}}(x, Q^2, E_\nu)}{dx dQ^2} \right|_{\text{fid}} \right] \tag{2.3.14}$$

This FK-table is precomputed and only has to be computed once, and not for every iteration. This object is a 2 dimensional matrix, where one axis is labelled by the neutrino energy bin, and the other axis is labelled by the grid nodes. The event rates can now be expressed as a matrix multiplication between this FK-table and the neutrino PDF evaluated at the grid nodes:

$$N_{\text{int}}^{(\nu_i)}(E_{\nu,j}) = \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) \text{FK}_{\alpha,j}\,, \qquad j = 1, \ldots, n_{E_\nu}\,. \tag{2.3.15}$$

As one can see, instead of evaluating 3 integrals at every iteration only one matrix multiplication has to be performed. This of course greatly increases the computational efficiency when performing

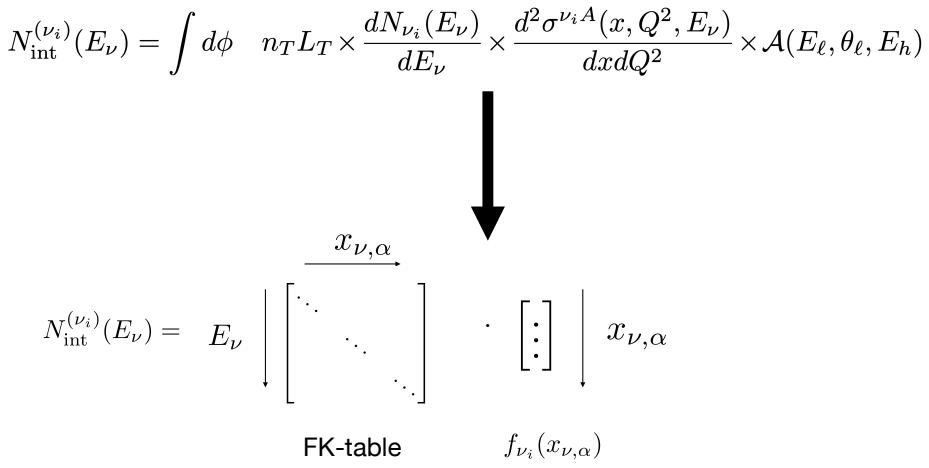


Figure 2.4. This is a visual representation of increasing computational efficiency in the ML fit. The triple integrals are replaced by a matrix multiplication by expressing the neutrino PDF in an interpolation basis. The matrix contains all fixed information i.e. DIS cross-section etc and the neutrino PDF has become an array of constant values.

a ML fit. Figure 2.4 provides a visual representation of how the increase in computational efficiency is achieved. In this derivation the event rates and the FK-table were expressed as a function of E_ν. This derivation can also easily be computed for the $(E_h, E_\ell, \theta_\ell)$ kinematic quantities, this is important since the event rates and FK-tables will also be expressed in terms of these quantities, which will also be used for parametrizing neutrino fluxes. The FK-tables are in principle also dependent on the neutrino flavour, however the mass of the outgoing leptons is much smaller than the cut imposed on the lepton energy i.e. 0.5MeV compared to 100GeV for electrons. This effect is neglected, thus the same FK-table is used to fit electron and muon neutrinos.

2.4 Neural Net Parametrization

Now the problem of computational inefficiency has been solved, the focus can be moved to parametrizing the neutrino PDF using ML. More specifically, we want to research how accurately the neutrino PDF can be parametrized. As mentioned before, close attention will be paid to the NNPDF framework, because of the equivalence between parametrizing neutrino PDFs and fitting PDFs from cross sections. The neutrino PDF will be parametrized as follows:

$$f_{\nu_i}(x_\nu) = x_\nu^{a_i}(1-x_\nu)^{b_i}\text{NN}_{\nu_i}(x_\nu)\,, \tag{2.4.1}$$

1. where $x_\nu^{a_i}(1-x_\nu)^{b_i}$ is a preprocessing function. This preprocessing function is designed to enforce the correct asymptotic behaviour of the neutrino PDF in small and large-x regions where there is little to no data[28] and helps to ensure a smooth and stable parametrization. The preprocessing function also facilitates the learning to decrease the computational load of the neural network. The preprocessing parameters are learned at the same time as the NN parameters.

2. $\text{NN}_{\nu_i}(x_\nu)$ is the neural network. There are many types of neural networks, for example convolutional NNs used in computer vision and variational autoencoders used to generate new data based on seen data. In this work a simple multi layer perceptron(MLP) is used[29]. Graphically it looks like figure 2.5. Most basically, a neural network maps an input to an output. A neural

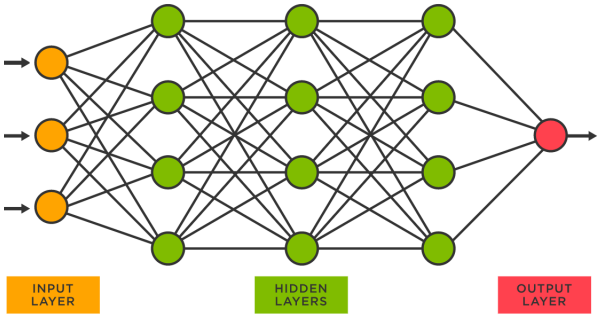


Figure 2.5. A graphical representation of a simple multi layer perceptron with size (3,4,4,4,1).

network consists of several layers: an input layer, hidden layers and an output layer. Every layer consists of a number of neurons or nodes. Each layer consists of weights and passes through an activation function. These weights are trained to best map the input to the output. The network learns by comparing its output to data, after which it computes a loss and adjusts the weights using stochastic gradient descent. This is called backpropagation. This process is repeated until the network correctly learns to map the input to a certain output. In this work the loss is defined as:

$$\chi^2 = \sum_{i,j}^{N_{\text{dat}}} (D-P)_i C_{ij}^{-1} (D-P)_j, \tag{2.4.2}$$

where D is the data, P the prediction made by the NN convoluted with the FK-table according to equation 2.3.15, and C the covariance matrix containing the correlations between data points and errors on data points. Contrary to the mean squared error loss, this loss optimizes the entire data range, because the the differences between D and P are multiplied by the inverse covariance matrix. For example, the difference between 10 and 5 is much smaller than the difference between 300 and 200, however, this loss makes sure that the NN pays almost equal amounts of attention to both differences because they are divided by the uncertainties of the data. By dividing the differences by the uncertainties on the data (covariance matrix), the loss becomes a relative one, or statistically meaningful, instead of an absolute one, which is for example the case for the mean squared error loss. The entire ML architecture can be seen in Figure 2.6. As is the case for ML models, there are hyperparameters which configure how the model is trained and set-up. Examples of hyperparameters are

- number of layers
- the number of nodes per layer
- the learning rate i.e. the step size taken in the direction of minimizing the error
- the optimizer: usually stochastic gradient descent, called Adam
- the activation function transforms its input and it is an important part of a NN, because it introduces non-linearities into the network enabling it to fit highly non-trivial functions.
- the number of iterations

In this work the (anti-) neutrino fluxes, and its anti-neutrinos are parametrized and trained by the same neural network, because this is convenient for the FASER event rate measurements. In this case there is one input neuron, namely the grid nodes, and two output layers, the neutrino and anti-neutrino flux. Fluxes for different neutrino flavours are trained in separate models.

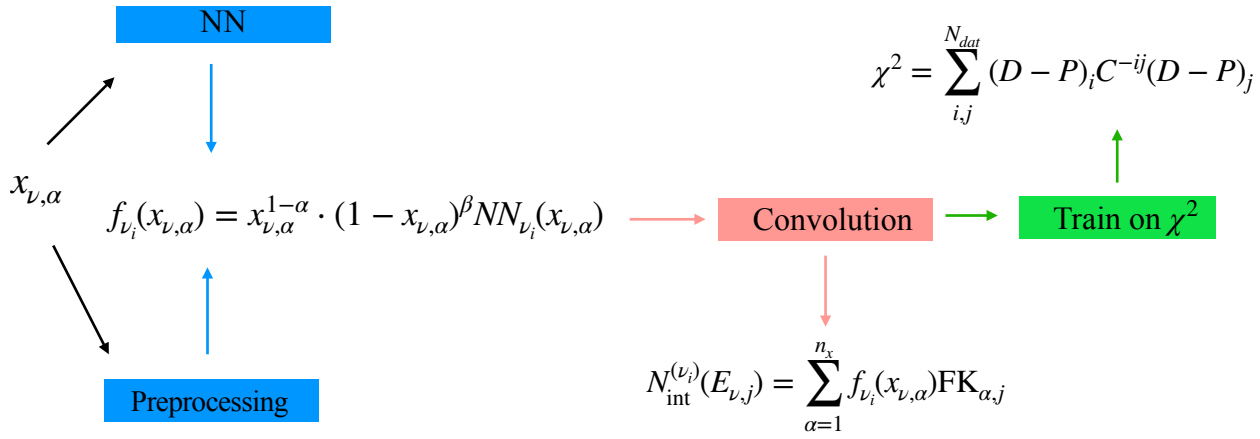


Figure 2.6. A graphical representation of the entire ML architecture used in this work. The grid nodes $x_{\nu,\alpha}$ are put into the MLP and into the preprocessing function. The outcome of the NN and the preprocessing function are then multiplied and convoluted with the FK-table. The event rates are then put into the loss together with the data and the model continues training.

3 Data

In order to train a NN one of course needs data. In this work there are two sources of data: measurements by the FASER experiment and pseudo data created by several MC event generators. In this section, details of how the data is gathered will be explained. The `POWHEG` and `PYTHIA8` frameworks will be explained in detail here, because both the FK-tables and events are generated using `POWHEG+PYTHIA`, even though also other event generators are used in this work. In Section 5, the other event generators used will be briefly explained. The explanation on `POWHEG` will also in part apply to the other generators, since there is overlap in methodology between generators.

3.1 `POWHEG` + `PYTHIA8`

Event generators and validation Before one can think about parametrising the neutrino flux using event rate measurements from FASER, the model needs to be tested and validated. This is where pseudo data comes along and saves the day, because it can generate both FK-tables and event rates. Using pseudo data one can validate the model and test if it is showing correct and physical behaviour. In the previous academic year, two master students, which also were under the supervision of Juan Rojo, integrated neutrino PDFs in the `POWHEG-BOX-RES` framework[21]. This framework implements neutrino DIS scattering and produces NLO predictions matched to `PYTHIA8`[30]. It can produce predictions for final-state distributions within detector acceptance (the FASER experiment for example). By generating events using the fiducial region of the FASER experiment and using the neutrino fluxes simulated in [31], which is the number of neutrinos passing through the FASER detector, the ML model presented in this work can be validated. This is because the result of the parametrization can directly be compared to the simulated neutrino PDF, since the event rate data is computed using this same neutrino PDF. In this section, the rough workings of the `POWHEG+PYTHIA` frameworks will be explained.

Brief Overview `POWHEG + PYTHIA` At the very core `POWHEG` is an MC event generator using collinear factorization to generate events. It was created to model hadron-hadron collisions at the LHC[32, 33, 34, 35]. An extension of this framework called `POWHEG DIS` was created to simulate DIS processes.

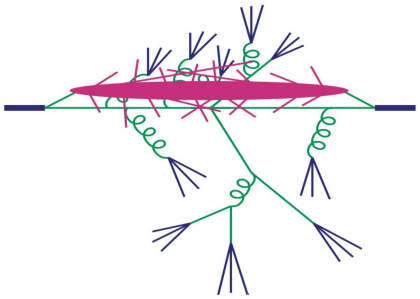


Figure 3.1. A graphical representation of the stages of the MC event generator simulating a pp collision. The pink line is the underlying event, green the jets and blue the hadronization. The latter two processes are simulated by `PYTHIA`.[38]

This framework also includes the process considered in this work and described in Figure 2.2. The starting point to generate events is the `POWHEG` cross section. This cross section is a sum of Born cross sections (the probability for a reaction to occur when 2 particles collide), integrated over phase space and summed over all partonic channels. This cross section can be computed in LO and NLO, the difference being that extra radiative emissions are also taken into account at NLO. Then it simulates an event by splitting the physical process into several stages iteratively to construct a more and more complex final state. After it has generated events, it is showered by `PYHTIA8` [36]. The role of the parton shower is to describe the describe the jets in DIS up to hadronization, because in the scattering process described in figure 2.2 several jets occur. Jets are narrow cones of hadrons and can also contain a substructures[37]. Figure 3.1 shows the several stages of an MC event generator simulating a pp collision, where `PYHTIA` models the green (jets)and blue parts(hadronization) and `POWHEG` the other stages. In this work, only NLO computations are considered.

Uncertainty There are two sources of uncertainty for pseudo data:

- Systematic uncertainty: The `POWHEG` framework provides theoretical uncertainties on the observables and the number of events. This is done by running the code with different renormalisation and factorisation scales independently. By varying these scales a maximum and a minimum value of the number of events is computed, which are used as error bands. The uncertainty on NLO results is expected to lie inside the error bands on LO results[39].

- Statistical uncertainty: Besides this theoretical uncertainty, there is also a statistical error on the number events computed by the usual $\sqrt{N_{int}}$. In this work only the statistical uncertainty will be taken into account, since the systematic uncertainty is negligible in comparison.

Acceptance Cuts and Simulation Settings In order to take the kinematic acceptance region of the detector into account, acceptance cuts are imposed in the code, which is given by the factor $\mathcal{A}(E_\ell, \theta_\ell, E_h)$ in equation 2.2.6 and included in the fiducial volume in equation 2.3.10. By imposing these cuts one can model the event yield for several different detector geometries, for example FASERν2 and FASERν by incorporating them in the generated neutrino PDF[21]. Concerning acceptance cuts, the cut Q_{min} = 2GeV is imposed, which is a standard DIS cut to ensure that one stays in the perturbative QCD region. This cut has little effect on the physics, since there are (almost) no events below this scale. Also cuts to the final-state variables are applied: $E_\ell > 100$ GeV, $n_{\text{tracks}} \geq 5$, and $\Delta\phi > \pi/2$. n_{tr} is the number of charged tracks i.e. the number of charged particles which have an energy above 2 GeV. $\Delta\phi$ is the angle between the lepton and the hadronic system after the scattering. Also, the following physical parameters are used:

$G_F = 1.16637 \times 10^{-5}$ GeV^{-2}	$m_W = 80.385$ GeV	$m_Z = 91.1876$ GeV
$\lvert V_{ud}\rvert = 0.97383$	$\lvert V_{us}\rvert = 0.2272$	$\lvert V_{cd}\rvert = 0.2271$
$\lvert V_{cs}\rvert = 0.975$	$\lvert V_{cb}\rvert = 42.21 \times 10^{-3}$	$\lvert V_{ub}\rvert = 3.96 \times 10^{-3}$
$\lvert V_{td}\rvert = 8.14 \times 10^{-3}$	$\lvert V_{ts}\rvert = 41.61 \times 10^{-3}$	$\lvert V_{tb}\rvert = 0.9991$

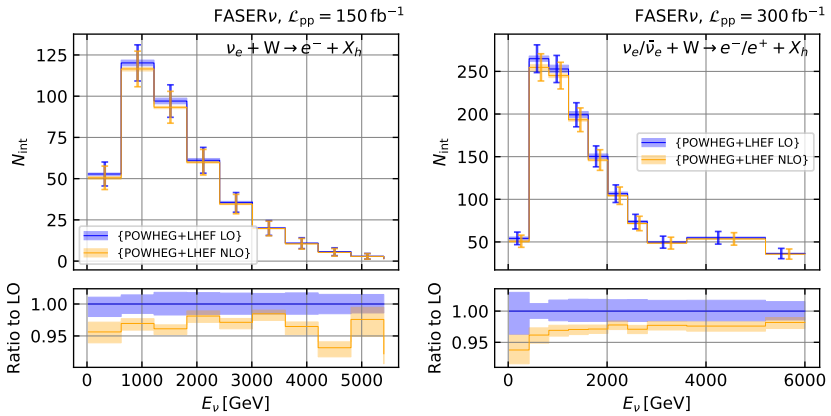


Figure 3.2. The number of events generated by `POWHEG + PYTHIA` for the electron neutrino as a function of the neutrino energy. Both LO and NLO is shown. The error bands are theoretical errors and the error bars are statistical errors. In the right plot, some of the bins in low- and high-E_ν are combined to make sure there are enough statistics in these regions.

Rebinning of Events When running `POWHEG+PYTHIA`, one can specify the binning of the events i.e. the binwidth of the kinematic variables. After running the code, a histogram shown in Figure 3.2 is created by reading the event files. As can be seen from the figure, the number of events drop off quite quickly at high neutrino energies. As mentioned above, the statistical errors scales with $\sqrt{N}$, thus the relative statistical error drastically increases when there are only a handful of events in a bin, or even less. This effect should be dampened considerably, because we want to accurately parametrize the neutrino flux and thus need enough statistics. This is where rebinning enters the stage and saves us from the demon of not having enough statistics. The idea of rebinning is to increase the binwidth and thus increasing the number of events in a bin. As a rule of thumb, at least 20 events per bin is imposed, because then a Poisonnian distribution can be approximated by a Gaussian distribution. The rebinning of events can be seen in Figure 3.2.

Pseudo data in this work In this work pseudo data will be simulated for several different MC event generators, detector geometries and detector observables, more on this in Section 5. The FK-tables will be different for the several kinematic observables and the fluxes will incorporate the detector geometry and the effects of the different MC generators. The event rates will be computed by convoluting the fluxes with their respective FK-tables, where the fluxes are thus generated by the different event generators, such as `POWHEG`.

3.2 Data from FASER(ν)

FASERν Detector The main goal of this work is to parametrise the neutrino flux using FASER event rate measurements. In this section the experimental setup of the FASER(ν) experiment will be discussed in order for the reader to get an idea as to how the cross sections and event rates are measured. The ForwArd Search ExpeRiment, or FASER for short, is a detector designed to measure light and weakly-interacting particles that escape the ATLAS detector and fly in the forward direction. The experiment is located about 480m from the ATLAS interaction point and is close to or in the

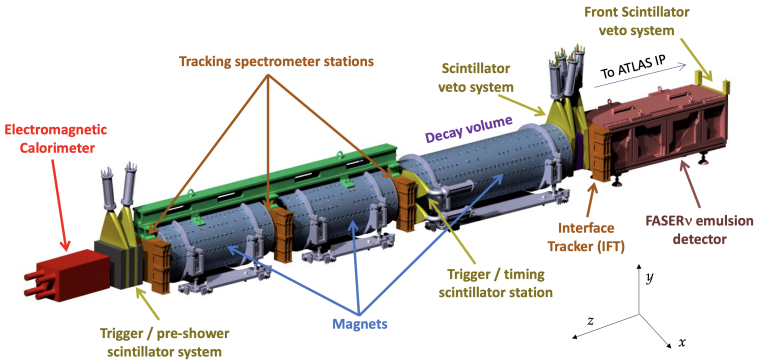


Figure 3.3. A sketch of the setup of the FASER(ν) detector. The FASERν emulsion detector is placed all the way at the beginning and the FASER detector is behind the emulsion detector.[40]

centre of the neutrino beam, thereby maximizing the event rate maesurements. The sub-detector FASERν is designed to detect all neutrino flavours. After travelling through stone and concrete where the particles are not interacting, they arrive at the detector where they decay into visible particles [40]. In figure 3.3 a sketch of the several parts of the detector are shown. The FASERν detector consists of tungsten plates alternated with emulsion films. In total there is 1.1 tonnes of tungsten in the detector. Here, neutrinos interact with the tungsten material via the CC DIS process described in section 1, where the emulsion films track the trajectories of the charged particles being created. These emulsion films can distinguish different lepton flavours, measure the momenta of hadrons and estimate the energy of the neutrinos. The tungsten target is chosen, amongst lots of other reasons, which are highlighted in [40], because of its high interaction rate whilst keeping the size of the detector small and thus manageable. In the end the emulsions will be analysed by a readout system in order to perform an event analysis. In order to get a proper and full grasp of the detailed workings of the detector, one might read [40], where every aspect of the detector is explained in thorough fashion.

Event Rate Measurements In December 2024 the FASER collaboration published a paper where they presented measurements of muon neutrino cross sections and the muon flux for the first time at the LHC[41]. The muon neutrinos were measured as a function of the neutrino energy and observed using an electronic detector which is located behind FASERν (FASER is an electronic detector) and was used as a spectrometer for muons produced in the neutrino interactions. Figure 3.4 shows the number of (anti-)muon neutrino interactions, including uncertainties and the simulated number of events. The binning was chosen such that there is a distinction between neutrinos and anti-neutrinos, and it was also chosen such that there is a comparable number of events in a bin. These measurements were conducted in 2022 and 2023 during run-3 of the LHC. The integrated luminosity is 65.6 fb^{-1}. In total, they observed:

$$n_{\nu,obs} = 338.1 \pm 19.0(\text{stat}) \pm 8.8(\text{sys}) \tag{3.2.1}$$

events. The FASER collaboration also provided a covariance matrix, including correlations between the datapoints, which will be used to compute the loss defined in equation 2.4.2. Using these event rate measurements from FASER a model-independent parametrisation of the neutrino flux will be performed.

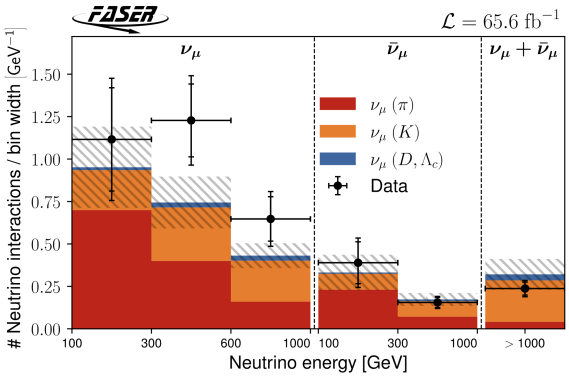


Figure 3.4. The unfolded(including background events) number of (anti-) muon neutrino interactions (black dots) compared to simulations (histogram). The error bars include both the statistical and systematic errors. The last bin has a width of 868.2 GeV and contains both muon and anti-muon neutrinos because from 1TeV the charge identification accuracy significantly decreases, thus one cannot distinguish the neutrinos coming from muons or anti-muons.

4 Numerical Framework

This section will describe all of the features, algorithms and numerical implementations used in this work alongside the several fits which were obtained using these implementations. First of all, the MC approach behind NNPDF closure tests will be explained. Second of all, the analysis machinery will be tested by carrying out two non-ML fits to see how well the uncertainties of the neutrino flux can be constrained. Also, the neural net machinery will be tested by directly fitting to the number of events i.e. without a convolution between the neutrino flux and the FK-table. Next, the complete numerical implementation of parametrising the neutrino flux using the NN will be presented. Finally, the neutrino PDF parametrised using FASER event rate measurements will be shown. All neutrino PDFs shown in these regions are muon neutrino PDFs and all the data shown are muon neutrino scattering events.

4.1 Implementation and Validation of FK-table Formalism

As was explained below, the neutrino PDF can be expressed in terms of a linear expansion in an interpolation basis:

$$f_{\nu_i}(x_\nu) = \sum_{\alpha=1}^{n_x} f_{\nu_i}(x_{\nu,\alpha}) I_\alpha(x_\nu)\,, \tag{4.1.1}$$

where the grid nodes(in neutrino momentum fraction) are given by

$$x_{\nu,\alpha} \in \{x_\alpha \mid 0 < x_\alpha \leq 1, \alpha = 1 \ldots n_{x_\nu}\}\,. \tag{4.1.2}$$

In this work the grid in $x_{\nu,\alpha}$ is divided in blocks:

$$B = [\alpha - \lfloor n/2 \rfloor, \alpha + \lfloor (n-1)/2 \rfloor]\,, \tag{4.1.3}$$

of size n. This interpolation approach was presented in the `EKO DGLAP` evolution package [14]. In this work 50 nodes are used for the interpolation, which is, as shown below, enough to reproduce the `POWHEG+PYTHIA` calculations within 1%. Using this division of x_ν into blocks, the Lagrange interpola-

tion polynomial is given by

$$I_\alpha(x_\nu) = \begin{cases} \prod_{\substack{i\neq\alpha \\ i\in B}} \frac{x_\nu - x_i}{x_\alpha - x_i} & \text{for } x_{\min_{j\in B}} \leq x_\nu \leq x_{\max_{j\in B}} \\ 0 & \text{else} \end{cases} \tag{4.1.4}$$

where the interpolation polynomial is only non-zero in regions where x_ν is close to the node α, which is numerically and computationally advantageous. In this work Lagrange polynomials are chosen to ensure numerical stability of the results. Figure 4.1 shows FK-table formalism is working, since the FK-table convoluted by neutrino PDF is within 2% of the `POWHEG` calculations in regions where there are enough statistics..

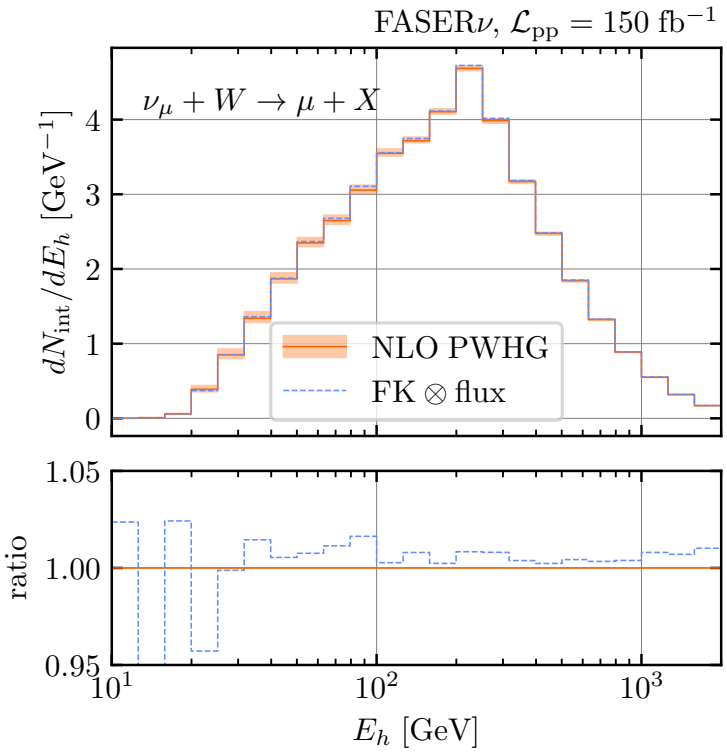


Figure 4.1. Comparison between the original NLO POWHEG simulation of neutrino event yields at FASERν (for $\mathcal{L}_{pp} = 150$ fb^{-1}) as a function of E_ν, with its counterpart based on the fast grid FK-table interpolation. The bottom panels display the ratio of the interpolated calculation with respect to the POWHEG result, showing that their differences are at the 2% level at most. The ratio in the low E_ν region increases due to low statistics.

4.2 Computing Uncertainties: NNPDF Closure Tests and MC replicas

In this work we want to accurately parametrise the neutrino flux. In order to have a sense of the accuracy of the flux, one has to compute the uncertainties fluxes parametrized using the NN. In the NNPDF framework a Monte Carlo approach has been developed to compute the uncertainties on ML fits. This is done by several so-called levels of closure tests. The rationale behind this idea is to generate MC replicas of the data by adding Gaussian fluctuations on top of the data. These replicas are then fitted using the NN, after which the mean and the error is computed using these several neutrino PDFs coming out of the fits[42]. In the NNPDF framework there are 3 levels of data:

- Level 0 data: this data is the raw data, and if one would try to fit a function to this data, one

would expect perfect agreement. The notation of level 0 data is as follows:

$$D^{(0)} = N_{int}^{(\nu_i)}(E_{\nu,j}), \tag{4.2.1}$$

i.e. level 0 data is equal to the event rates from the event generator or from FASER measurements. For pseudo data, a level 0 closure test should produce a χ^2 which is numerically zero and can be used to test the performance and correct behaviour of the fitting code. For event rate measurements, level-0 data should not produce a χ^2 of 0, because measurements inherently include fluctuations, and do not correspond to the central value, as is the case with simulated data.

- Level 1 data: this is where one adds a first layer of Monte Carlo fluctuations:

$$D_I^{(1)} = D^{(0)} + r_I^{\text{stat}}\sigma_I^{(\text{stat})} + r_I^{\text{sys}}\sigma_I^{(\text{sys})} \tag{4.2.2}$$

where I is a certain instance of the data, $\sigma_I^{(\text{stat})}$ and $\sigma_I^{(\text{sys})}$ are the statistical and systematic uncertainty of the data and r_I^{stat}, r_I^{sys} are the corresponding Gaussian random numbers. Each datapoint has its own random number and one can create several instances of level 1 data. Each of these fluctuations is called a replica.

For event rate measurements from FASER, level-1 data does correspond to MC replicas, because the measurements from the experiment inherently include fluctuations and do not correspond to central values. However, for pseudo data, a level-1 closure test will underestimate the uncertainties of the data. This is because the event rates from the MC generators correspond to the central value and a level-1 closure test will underestimate the uncertainties of the data, because the fit is performed on the same underlying data i.e. no Monte Carlo replicas are computed using equation 4.2.2.

- Level 2 data: Using level 1 data, N_{rep} Monte Carlo replicas are constructed by adding another level of Gaussian fluctuations to the level 1 data:

$$D_I^{2,k} = D_I^1 + r_I^{\text{stat},k}\sigma_I^{(\text{stat})} + r_I^{\text{sys},k}\sigma_I^{(\text{sys})}. \tag{4.2.3}$$

When using data from FASER, a level-1 closure fit is performed, and for the pseudo data a level-2 closure fit is performed. Now, one fits the ML model to every replica to construct a probability density in the space of ML models by computing the MC error.

By creating N_{rep} replicas, one will reproduce the uncertainties and correlations from the (experimental) covariance matrix. Each of these replicas will then be used to fit a neutrino flux to. In the end, the resulting flux will include the experimental uncertainties and correlations. For this method to work, one needs at least 100 replicas. In all cases the errors are fluctuated with the diagonal of the corresponding covariance matrix. For the pseudo data, only statistical uncertainties are included and for the FASER event rate measurements the experimental covariance matrix is used to generate the replicas.

4.3 Getting Hands-On with the Data: Gauging Uncertainties

4.3.1 Fitting Overall Flux Normalization

In order to test the analysis machinery and to get an idea of how well the overall normalization of the flux can be fitted. The FASER event rate measurements from FASER and pseudo data generated using the FASER detector geometry will be used. The pseudo data is generated using `EPOS+POWHEG`, where `EPOS` models light hadron production and `POWHEG` charmed hadron production. In Section 5, as short description of the event generator `EPOS` will be given. The first exercise to consider is to produce a fit using the following equation:

$$f_{\nu_i}(E_\nu) = \mathcal{A} \times f^{ref}_{\nu_i}(E_\nu), \tag{4.3.1}$$

where $f^{ref}_{\nu_i}(E_\nu)$ is the neutrino flux used to compute the pseudo data. As one can see, this equation is designed so that $\mathcal{A} \approx 1$, thus one might call this a trivial fit, which would not be far from the truth. We will consider two scenarios:

Fitting Using Identity FK-table

First a fit is performed where the assumption is made that the FK-table is an identity matrix, thus the event rates are related to the flux by:

$$N^{(\nu_i)}_{\text{int}}(E_{\nu,j}) = f_{\nu_i}(x_{\nu,j})\Delta E_{\nu,j} \quad i = 1, \ldots, n_{E_\nu}\,, \tag{4.3.2}$$

where $\Delta E_{\nu,j}$ is the binwidth in the neutrino energy. This fit comes down to constraining the normalization $\mathcal{A}$ using the number of events. The `SCIPY`[1] library is used to perform the fit and is in charge of minimizing the χ^2. For this fit, 100 Level-1 replicas have been computed using the event rate measurements from FASER. Figure 4.2 shows the result of the fit and tells us how well the normalization of the flux can be constrained. The resulting fit parameter is $\mathcal{A} = 1.00 \pm 0.06$ and the mean $\chi^2 = 1.67$.

Fitting Using FK-table Formalism

In order to test the FK-table formalism the same exercise has been done using pseudo data. In this case the number of events were related to the neutrino PDF by equation 2.3.15. We have chosen to bin the pseudo data in the same way as the data from FASER. Because of this, the normalization parameter fitted for the muon neutrino PDF and anti-neutrino PDF is the same. Both PDFs are then convoluted with their corresponding FK-table after which the number of events in the [1000,1900] GeV bin is added together. The result of this level 2 closure test can be seen in figure 4.3. This Figure shows the FK-table formalism is working properly, because the convolution of the neutrino PDF yields the event rates within uncertainty. The resulting fit parameter is $\mathcal{A} = 0.97 \pm 0.04$ and the mean $\chi^2 = 2.38$. The neutrino PDF from the fit is lower than the reference neutrino PDF because this specific instance of the level 1 data is on average smaller than the level 0 data.

4.3.2 Extension of Fitting Overall Normalization

Because the previous analysis was successful on all fronts, we can move on to an extension of the previous exercise. In this case the fitting machinery is extended to fit possible deviations in small- and large-x_ν limits of the neutrino fluxes. In this case the fit equation is given by

$$f_{\nu_i}(E_\nu) = \mathcal{A} \times x^{b-1}_{\nu_i}(1 - x_{\nu_i})^{c-1} \times f^{(\text{ref})}_{\nu_i}(E_\nu)\,, \tag{4.3.3}$$

[1] https://scipy.org

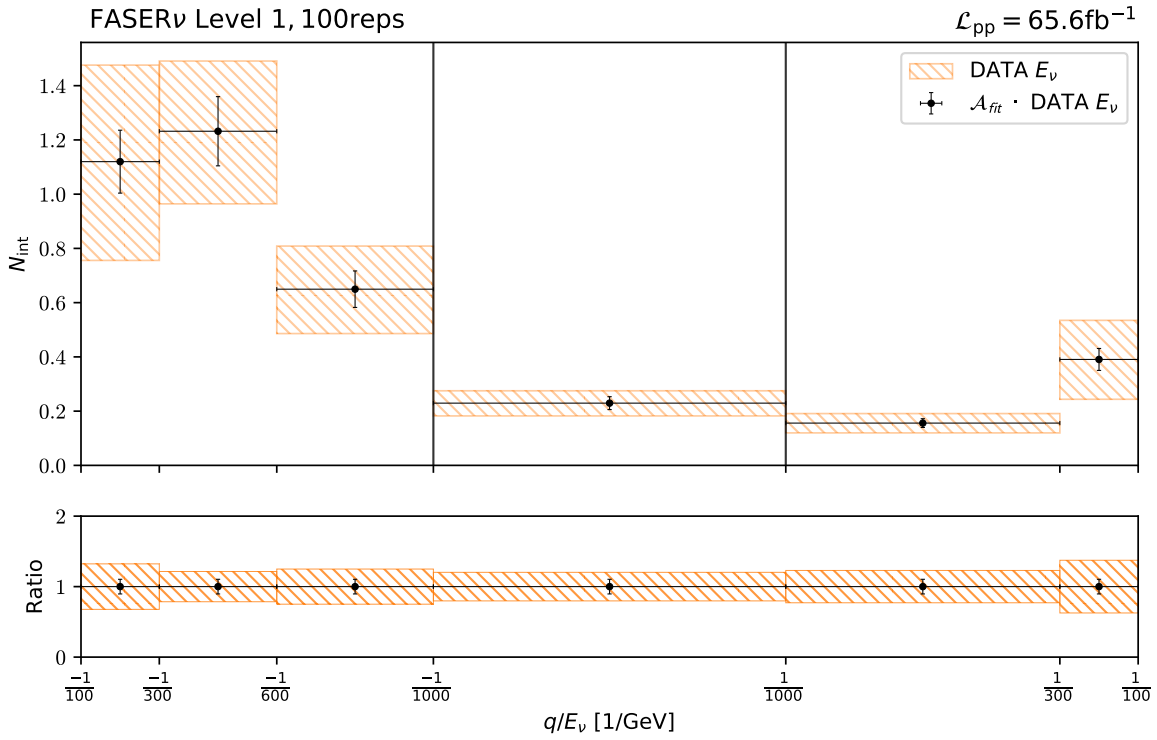


Figure 4.2. A level-1 closure fit probing the uncertainty of the normalization of the flux using equation 4.3.1. 100 replicas of the FASER data were computed and fitted. The normalization parameter was determined at a value of $\mathcal{A} = 1.00 \pm 0.06$ and the mean $\chi^2 = 1.67$. The normalization error falls in between the errors on the event rate measurements.

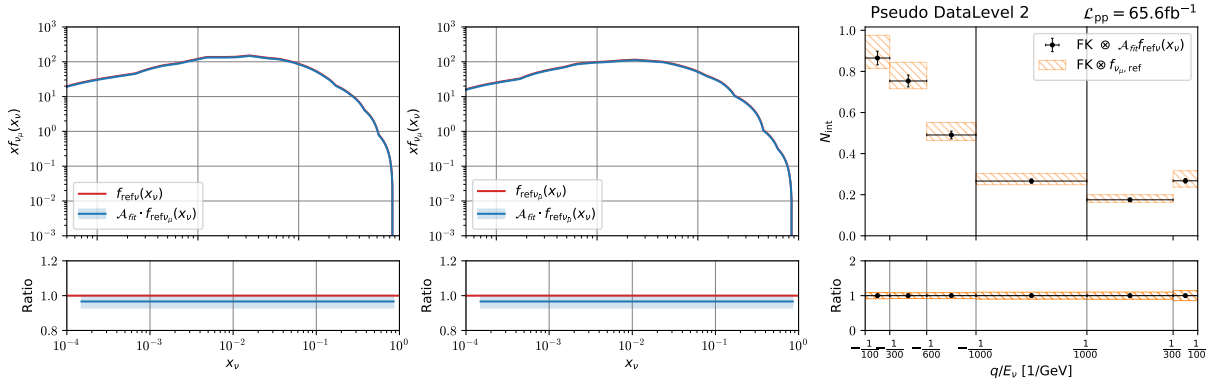


Figure 4.3. A level-2 closure fit probing the uncertainty of the normalization of the muon neutrino flux(left plot), and the anti-muon flux(middle plot), using equation 4.3.1. 100 replicas of the simulated data were computed and fitted. The normalization parameter was determined at a value of $\mathcal{A} = 0.97 \pm 0.04$. The normalization error falls in between the errors on the event rate measurements. In the upper right and the upper middle plots, the blue line is below the green and the red line. In the bottom left and bottom middle plot, the blue line is below the red line, because the level 1 data is on average smaller than the level 0 data in this instance.

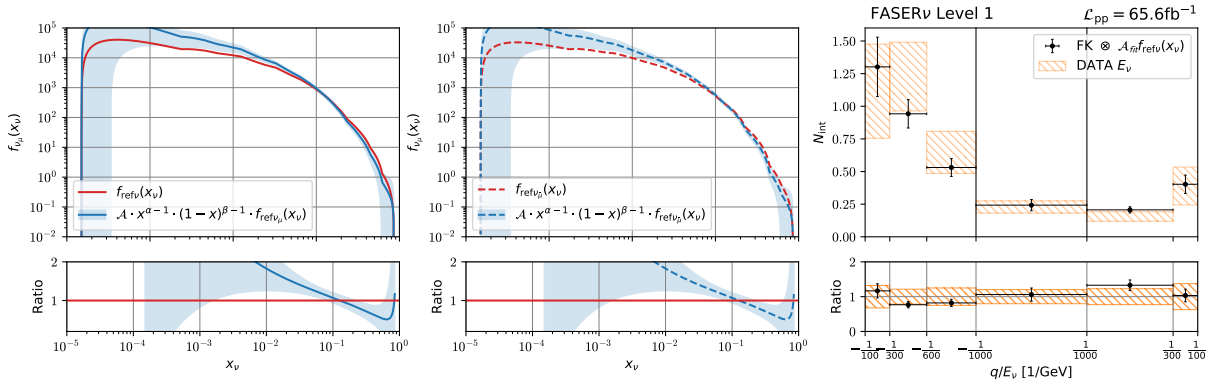


Figure 4.4. A level-1 closure fit probing the uncertainty of the flux using equation 4.3.3. 100 replicas of the FASER data were computed and fitted. The uncertainties of the fit correspond to the uncertainties of the data. This fit shows that a NN is needed to fit the neutrino flux well, since only the preprocessing function itself is not flexible enough.

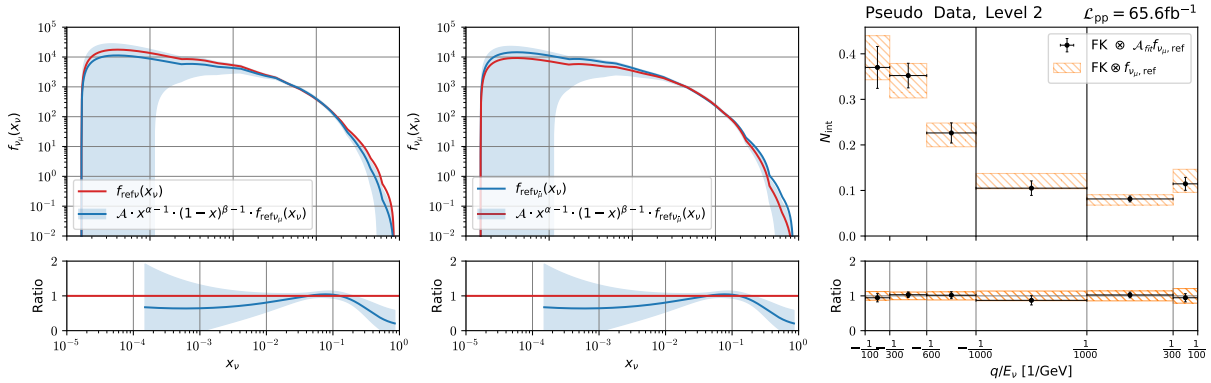


Figure 4.5. A level-2 closure fit probing the uncertainty of the muon neutrino flux using equation 4.3.3. 100 replicas of the simulated data were computed and fitted. The uncertainty on the event rates is similar to the normalization fit. For the neutrino PDF plot, one can see that the input neutrino PDF is obtained in regions where there is data, and that there is a large uncertainty in the region where there is no data. The χ^2 of this fit is lower compared to the normalization fit.

where by design the coefficients take the values: $\mathcal{A} = 1 = b = c$ in a level-0 closure test. The `SCIPY` library did not succeed in producing low χ^2 fits, several minimization methods were tried, but none of them yielded stable results. This is why, instead of `SCIPY`, the Adam optimizer was used. This optimizer yielded more stable results and lower χ^2s. The FK-table formalism will also be used to fit the event rate measurements from FASER, since it was deemed successful in the normalization fits. The results of level-1 closure test using event rate measurements from FASER can be seen in figure 4.4. It can be seen that the fitted flux is higher compared to the input flux, because the event rate measurements of FASER do not exactly correspond to pseudo data of course, but are higher, as can be seen from Figure 3.4. When looking at the right panel, this fit seems worse than the normalization fit. However, this is because the fit formula itself does not have enough flexibility(parameters) to fit the data well. This one of the reasons why a machine learning approach is used in this work. The result of the fit using pseudo data where the event rates and the neutrino PDF are related by a FK-table can be seen in Figure 4.5. Again, the uncertainties of the fit can be reasonably constrained and fall within the errors and the data. The parametrised neutrino flux reproduces the input flux in the region where there is data. The mean $\chi^2 = 1.18$, which is, as expected, lower compared to the normalization fit.

4.4 Gauging Uncertainties with Machine Learning Machinery

After gauging the normalization and overall uncertainties and testing the FK-table formalism, we move on to the ML fit. As mentioned above, the preprocessing function alone is not enough to accurately constrain the neutrino fluxes. To test the ML machinery, we start by considering an identity FK-table, thus directly fitting to the number of events. Because no neutrino PDF is convoluted in this case, no preprocessing function is involved.

4.4.1 Settings for ML parametrisations

For all ML parametrisations a similar stopping algorithm is employed as is done in the NNPDF framework. A graphical representation can be seen in Figure 4.6. The general idea of the stopping is to stop the training when the validation loss stops decreasing. When the validation loss stopped decreasing a patience algorithm is employed to make sure the training is not stopped to early by waiting a number of iterations to see if the loss is not decreasing further. In this work, enough data is not always available to split data into two sets, in these cases the validation loss is substituted by the training loss. In this work, the positivity constraint is employed by using the Softplus activation function after the final layer, whereas in NNPDF this constraint is implemented into the stopping algorithm.

ML FASER Event Rate Fit In order to get a reasonable parametrisation, one has to construct a well-performing NN. This is done by trying several number of hidden layers, number of nodes per layer, learning etc. It is very much a matter of trying and experimenting, engineering if you will, to get a NN which properly trains. To fit the event rates measurements from FASER without FK-tables, a simple 1-2-2-2-1 MLP is used, which includes 19 free parameters including 7 bias parameters. These are more than enough parameters to fit 6 points. Initially many more parameters were used, but the NN had a hard time converging. After realising the NN would probably not need many parameters to for 6 points, a smaller NN was used. A level-1 closure fit is performed using the event rate measurements from FASER. For this case, two separate networks are trained for the event rates coming from muon and anti-muon neutrinos. Both networks used the Adam optimizer with a learning rate of .001 and also the stopping algorithm was employed. The result of the fit can be seen in Figure 4.7. As can be seen from the plot, the uncertainty from the fit overlaps almost perfectly with the uncertainty of the data, confirming that the MC replica method is doing its job. This Figure also shows the NN can fit the data well. Both fitting using the FK-table formalism, as was shown before, and ML fits were successful, thus now one can move on to the full-fledged ML parametrisation of the data using the FK-table formalism.

4.5 ML Parametrisations: From Pseudo Data to FASER Event Rate Measurements

The main goal of this work is of course to parametrise the neutrino flux using ML and the FK-table formalism. First, to verify that the ML model is doing a proper job, pseudo data is used to make sure the input flux is recovered.

4.5.1 Why Not Implement Fitting Neutrino Fluxes in the NNPDF Framework?

The first approach to parametrising the neutrino flux was using the NNPDF framework. However, we quickly came to the conclusion it was better to write our own stand-alone code. This is because

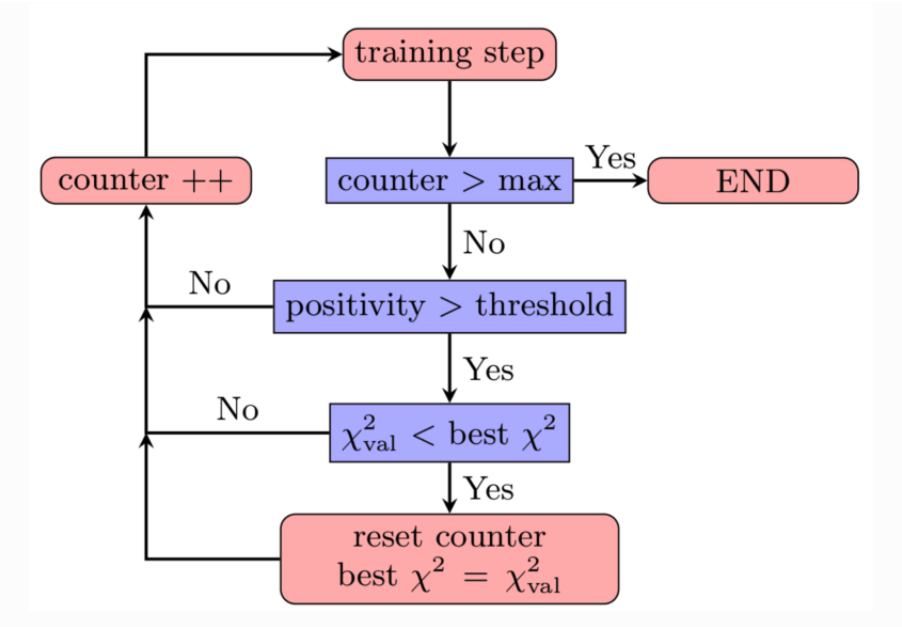


Figure 4.6. This shows how the stopping algorithm employed in the ML training works. Basically, the neural net is trained as long as the loss decreases. If the loss did not decrease, a patience alogorithm is employed to make sure the training is not stopped too early. The positivity constraint is employed by the activation function in this work.

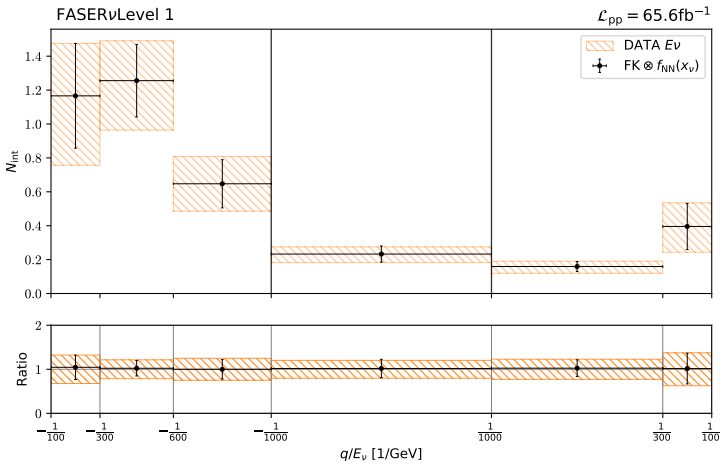


Figure 4.7. This Figure shows a parametrisation using an identity FK-table. As can be seen, using a level-1 closure test, the event rate measurements from FASER and its uncertainties can be reproduced, thus confirming that the MC replica method is working. A simple MLP was used (1,3,3,3,1) alongside ReLU activation functions and L2 regularisation.

NNPDF is simply not designed to be used for any other purpose than fitting PDFs. The NNPDF framework is a very useful code for fitting PDFs, but because of this it contains a lot of unnecessary and unfitting code for parametrising neutrino PDFs. For example, the framework contains several PDF bases and rotations, which is a great asset for parametrising PDFs, but this is quite inefficient and abundant for this purpose. One could potentially overcome this barrier by implementing a trivial neutrino PDF basis, but this would be only one of many duct tape fixes needed. Therefore, it is much more efficient to write our own code whilst relying heavily on the relevant parts of the NNPDF framework.

4.5.2 Complete ML Model and Parametrising Pseudo Data

Complete Description ML Model After deciding to write a stand-alone code, the focus shifted to parametrising the pseudo data. First a level-0 closure test was performed to test the performance of the model and thereafter a level-2 closure test is performed. In this work the python package `PyTorch`[2] is used to construct the ML model. After playing around with several model configurations e.g. number of layers, number of nodes per layer, learning rate, activation functions, optimizers and setting up the framework a successful level-0 closure fit was produced for the simulated data. In the preprocessing function, equation 2.4.1, the β parameter was constrained to be positive to ensure the neutrino PDF would not diverge in the low-x_ν region. A wide range of initial values were used for both the NN parameters and the preprocessing parameters, so that the model would explore the entire parameter space in regions where there is no data to ensure large uncertainties in those regions. For all fits a minimum of 100 replicas was used to ensure the uncertainties of the covariance matrix were correctly propagated into the neutrino flux. In this work a NN of size (1,4,4,4,1) was used, alongside the Softplus activation function, which is employed after each layer to allow for non-linearities and enforce positivity. The optimizer used in this work is, as usual, the Adam optimizer combined with L2 regularisation to ensure the weights of the NN do not blow up and to avoid divergent/non-smooth behaviour and overfitting. The learning rate, patience, total number of epochs and training/validation split(if applicable) are determined separately for each fit. This ML model will be used to parametrise all neutrino fluxes in this work.

Postfit Criteria In the NNPDF framework postfit criteria are imposed to remove certain fits which are considered outliers relative to the other fitted replicas. These are also imposed in this work. If for example the χ^2 after training deviates more than 4σ from the mean χ^2 of all fits, this replica is removed. In the NNPDF framework several postfit criteria are implemented: χ^2, integrability, positivity and arc length. In this work, the postfit criteria are based on the χ^2, and not on the other criteria because those are already enforced using different approaches e.g. activation functions and weight decay.

Results Parametrisations Pseudo Data In Figure 4.8, a level 0 closure fit for the (anti-)muon neutrino PDF is shown. The result of this fit tells us the model is working correctly, as it reproduces the input flux from the simulated data. Also, the neutrino PDF is showing physical behaviour by going to zero in low- and high-x regions and displaying large uncertainties in dataless regions. For all the fits in this work, the uncertainties grow in the low- and high-x regions due to the lack of data (constraints) in those regions, which the fit is reflecting.

[2] https://pytorch.org

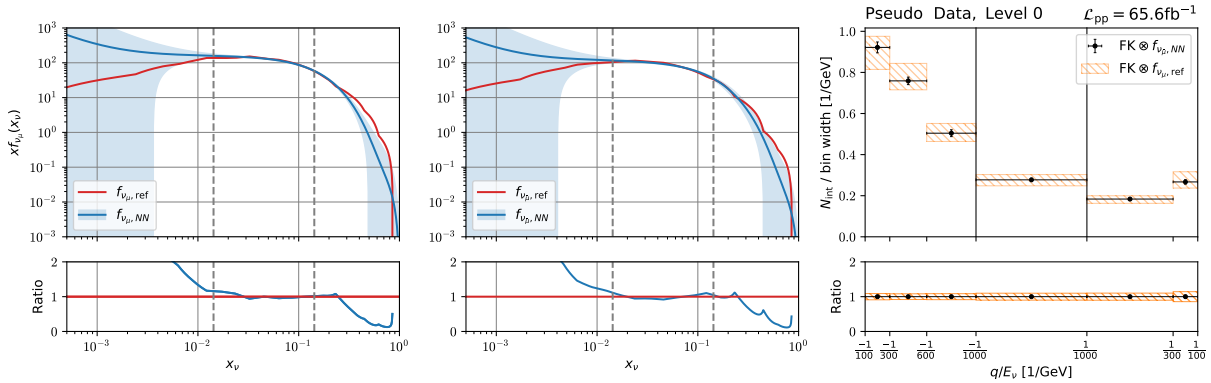


Figure 4.8. This Figure shows a successful level 0 closure test for the muon neutrino. In this plot event rates are generated using the FASER neutrino PDF. The neural net reproduces this input flux. Both the FK-table formalism and a small (1,3,3,3,1) MLP was used, alongside the Adam optimizer including L2 regularisation.

After getting a confidence boost from this parametrisation, the level-2 closure fits can be computed. From Figure 4.9, it can be seen that, once again, the closure test passed with flying colours. The input flux and the data is reproduced within the 1σ error bars and the neutrino PDF shows physical behaviour in the asymptotic regions.

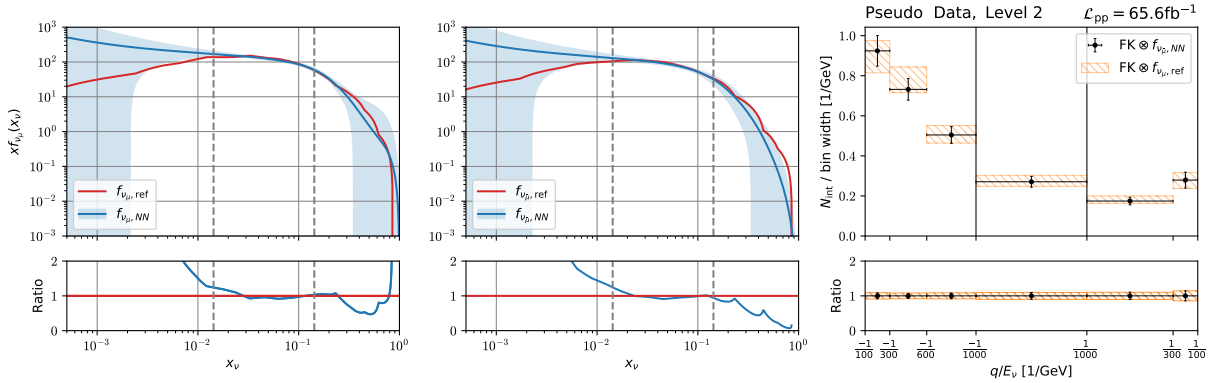


Figure 4.9. Same as Figure 4.8, however a level 2 closure test is performed instead of a level 0 closure test.

4.6 Extracting Neutrino Fluxes From FASER Event Rate Measurements

Now that we have established that the ML model works as intended, because it can reproduce the input flux within uncertainties and fit the event rates correctly, lets move on the pinnacle of this work: parametrising the (anti-) muon neutrino PDF to event rate measurements from FASER using the FK-table formalism and a feed-forward NN. Because the FASER experiment could not discriminate between neutrinos and anti-neutrinos at energies above 1000 GeV, the NN parametrising the flux contains two output layers, meaning the neutrino and anti-neutrino flux were fitted at the same time by a single NN. For this fit a 1-4-4-4-2 network was used alongside the Softplus activation function to keep the flux smooth, a weight decay of 1e-3, a patience of 100 and a total number of 2000 epochs. It was found that these settings made sure that both flux was smooth and the event rates were recovered within uncertainties. The result of this level-1 closure test can be seen in Figure 4.10. This plot shows the comparison between the NN fit and the FASER data. This is the first time the full $E\nu$ dependence on the muon neutrino flux is constrained. The event rates are binned in E_ν and leptonic charge q and the uncertainties correspond to 1σ. The uncertainties are correctly propagated through the MC method, since the uncertainties from the fit correspond to the uncertainties on the event rate measurement. It can be seen that the muon neutrino flux is higher than the flux from

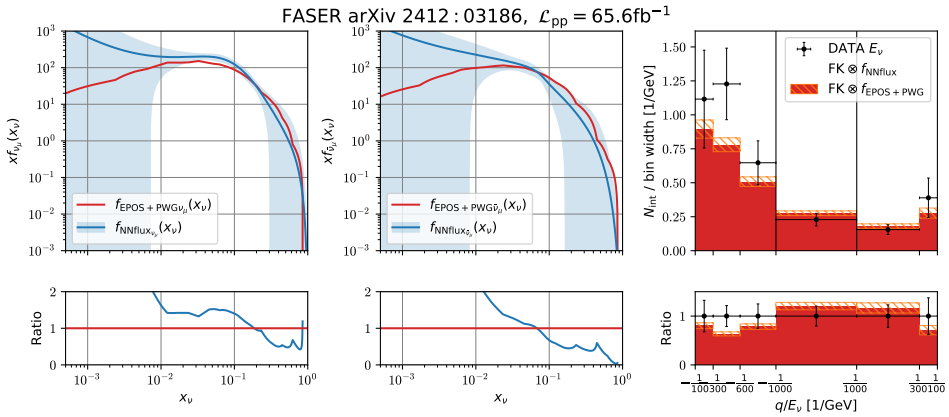


Figure 4.10. This plot shows the muon neutrino flux obtained using the FASER 2024 event rate measurements of [3]. It can be seen that the NN prediction is reproducing the uncertainties of the data in the rightmost plot. In the left and in the middle plot the parametrised flux is compared to the `POWHEG+EPOS` flux. In the right panel, also the event rates prediction by `POWHEG+EPOS` are shown for reference and completeness.

`EPOS+POWHEG` and thus that the method presented in this work is a significant improvement compared to the `EPOS+POWHEG` prediction. The neutrino PDF from the fit is between 20% and 40% larger than the `EPOS+POWHEG` flux and agrees within the fit uncertainties in almost the entire x_ν range.

4.7 Hyperparameters

The settings for the neural net described above were capable of producin satisfactory level-0, level-1 and level-2 fits. In this subsection, the rationale behind the neural net structure will be explained.

- **Preprocessing function**. As mentioned in the theory section the preprocessing function is used to make sure the neutrino PDF goes to zero in high- and low-x regions and to increase computational efficiency. In the closure test above, very few data points are available to train the NN on, and that is why the preprocessing function is also serving as a tool to get stable results. Since data is only available in a very limited x-range one should expect the parametrised flux to have large uncertainty in the large regions where there is no data at all. The preprocessing function in this work is used to explore the entire parameter space and can thus yield the expected uncertainty of the neutrino PDF in regions where there is no data. This can be seen in Figure 4.9 and can be achieved by varying the initial parameters of the preprocessing function sufficiently. When only using a NN, it is very hard (impossible) to create a NN which goes to zero in low- and high-x regions, displays large uncertainties in regions where there is no data, does not overfit, is computationally inexpensive and produces converged χ^2 values. An attempt to create large uncertainties can be seen in Figure 4.11. To generate this plot, a large NN was used and the initial parameters of the NN were also varied 10 times as much as the default option. An uncertainty band can be seen, but it is not as large as it should be, because this plot seems to know for sure the neutrino flux is in the shaded region. However, there is no data in that region. When further increasing the size or the initial parameters or using a different activation function the result either would not change, or the data would not be fitted correctly anymore.

- **Activation functions** In this work the Softplus activation function is used for a number of reasons. First of all the neutrino flux should be positive and the Softplus activation function, which is given by $f(x) = \frac{1}{\beta} * \ln(1 + exp(\beta * x))$, ensures that negative values are mapped to very

small values i.e. the Softplus is a smooth version of the Relu activation function which is given by $f(x) = \max(0, x)$. Because the neutrino PDFs need to be smooth, the Softplus activation function was found to be more suitable, although the Relu activation function is also capable of producing a successful closure test. This can be seen in Figure 4.12. Also, since the values of the neutrino PDFs are not between [-1,1], activation functions such as Sigmoid and Tanh are not suitable. Not only because of the range, but also because of the vanishing gradient problem are these activations functions not suitable for deep neural networks.

- **L2 regularisation**. L2 regularisation is a sensible feature to implement and a no-brainer, because it discourages large weights in the NN. This feature helps to avoid overfitting, makes sure the output of the NN is smoother and does not diverge in regions where there is no data and thus also helps in making the parametrisation more physical correct.

- **Learning rate and the training length** The learning rate and the size of the NN do not have a large effect in this work on the outcome of the fit. Generally, a larger learning rate causes faster, but also more unstable convergence of the loss. A smaller/larger learning rate is accompanied by more/less patience and epochs.

- **Size of the NN** In this work, relatively small NNs are used to parametrise the data, because the complexity of the neutrino PDF is low. Larger NNs could be used in this work, but this would only result in a slower convergence, more overfitting and an increase of dead neurons (neurons which are not contributing to the learning process because their weights are not updating due to redundancy).

- **Runcards** For all the fits computed in this work a separate runcard is used. In these runcards, one can specify the model architecture, hyperparameters(patience, learning rate, number of layers etc.), if one wants to use postfit measures/criteria etc.

- **Hyperparameter optimization** In this work the size of the datasets were not large (the largest dataset consisted of around 250 datapoints). Because of this a hyperparameter optimization algorithm was not used to find the combination of parameters which yielded the best fit quality. Instead hyperparameter optimization was done by trying a couple of combinations of parameters, usually after which the fit quality was up to our standards. NNPDF has implemented a hyperoptimization algorithm to find the best parameters for the learning of the NN. NNPDF uses a technique called K-folding cross-validation to perform the hyperparameter scan. This technique makes sure the NN produces a low χ^2 on both seen and unseen data, maintains stability despite random variations and remains dependable even with a limited number of points. The core idea of this algorithm is to create partitions, so-called k-folds, of the data. Then, for every combination of hyperparemeters, k-fits are computed where each time one k-fold is left out. The partition left out is used to evaluate the performance of the fit using a figure of merit, such as the hyperparameters which produced the best average χ^2 i.e. $L_{\textbf{hyperopt}} = \frac{1}{N_k} \sum \chi^2$ [15, 28]. Figure 4.13 shows a neat overview of this algorithm from NNPDF. Bayesian optimization is employed to find the set of hyperparameters which yield the lowest χ^2. In principle, one can also employ another NN to find the best set of hyperparameters, however in practice Bayesian optimization is found to be computationally less expensive and well suited to find a minimum[43]. This hyperoptimization algorithm is also implemented in this framework and can be used to find the best set of hyperparameters.

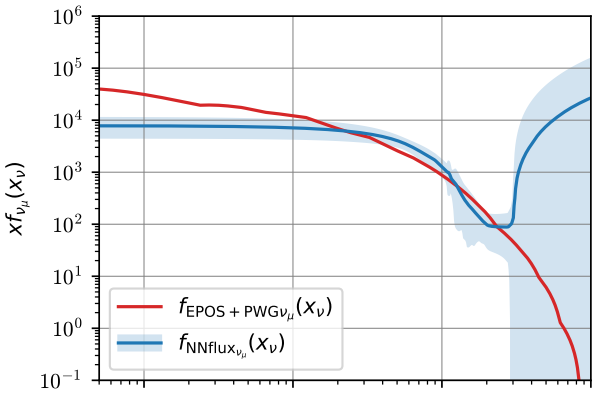


Figure 4.11. This Figure shows a fit to the FASER event rate measurements using a large NN (7 hidden layers, with 7 nodes per layer). The closure test is successful, however the error band is not increasing in regions where there is no data. Also, the neutrino PDF is not smooth and the fit was computationally expensive. All in all, the preprocessing function is a very powerful tool which can solve all these issues at once.

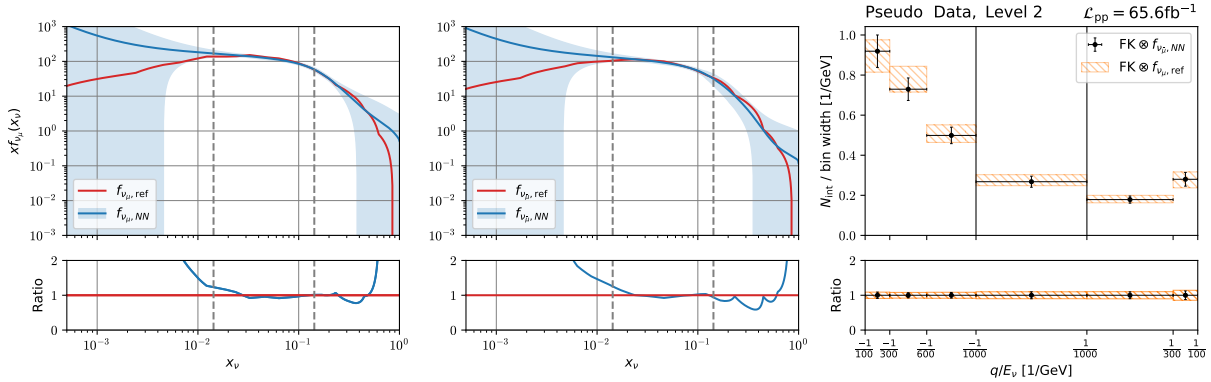


Figure 4.12. This Figure shows a successful level 2 closure test for the muon neutrino fitted to pseudo data. Instead of the Softplus activation function in Figure 4.9, the ReLU activation function was used. The resulting neutrino PDF is the same.

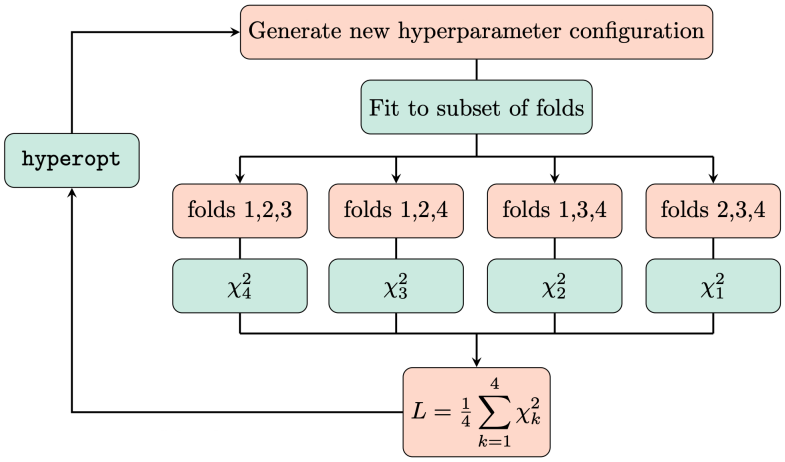


Figure 4.13. This is an overview of the hyperparameter optimization algorithm is implemented in the NNPDF framework[28]. The main idea is to use k-fold cross validation to determine the best set of hyperparameters combined with Bayesian optimization[43].

4.8 Assessing the Quality and Stability of the Parametrisations

In the NNPDF framework there are several so-called postfit measures to assess the quality of the fits [13, 28, 10], which are also implemented in this work. The data used to compute the postfit measures is ideally not used in the fit itself to make sure the measures are actually useful in telling the quality of the fits, however, in this work this cannot always be achieved due to the limited data.

- When performing closure tests, the following measure can be computed:

$$\varphi_{\chi^2} \equiv \sqrt{\langle \chi^2[N_{\text{fit}}, D]\rangle - \chi^2\langle [N_{\text{fit}}], D\rangle}. \tag{4.8.1}$$

  This measure computes the spread of all replicas in terms of data uncertainty. When fitting to level 0 data it is expected that all replicas converge(provided enough epochs are employed), since level 0 data does not contain any fluctuations and the only difference between replicas are the initial values of parameters of the model. Thus, it is expected that φ_{χ^2} goes to zero when the number of epochs are increased. For level 1 and level 2 closure tests, this quantity can also be computed with respect to level 0 and level 1 data to quantify the spread between the fitted replicas.

- When performing a level-2 closure test a specific instance of level-1 data is created (one experiment), however there is nothing holding us back to compute several level-2 closure tests using different level-1 instances. Actually, to for example get a probability distribution for the following postfit measures, one needs to execute several level-2 closure tests with different level-1 data instances, since these postfit measures are stochastic variables.

- It is to be expected that the average neutrino PDF obtained by level-1 and level-2 fits reproduces the input PDF. Thus the χ^2 of the average neutrino PDF convoluted with an FK-table and compared to level-1 data should be approximately the same as the χ^2 of the input PDF convoluted with an FK-table and compared to level-1 data. This is described by

$$\Delta_{\chi^2} = \frac{\chi^2[D_1, \langle N_{\text{fit}}\rangle] - \chi^2[D_1, N_{\text{data}}]}{\chi^2[D_1, N_{\text{data}}]}, \tag{4.8.2}$$

  and is a quantitative measure of how close the result of the closure fit is to the theoretical prediction (a measure of how successful the fit is). It is designed such that a Δ_{χ^2} of 0 corresponds to perfect learning of the underlying law, $\Delta_{\chi^2} < 0$ to overlearning i.e. also fitting noise and $\Delta_{\chi^2} > 0$ to underlearning.

  As mentioned above, an expectation value of this measure can be computed across different level-1 shifts i.e.

$$\Delta_{\chi^2} = \frac{1}{N_{\text{fits}}} \sum_{j=1}^{\text{fits}} \frac{\chi^2[D_1^j, \langle N_{\text{fit}}^j\rangle] - \chi^2[D_1^j, N_{\text{input}}]}{\chi^2[D_1^j, N_{\text{input}}]}. \tag{4.8.3}$$

- Another measure to measure the success of a level-2 closure fit is the accuracy which is given by:

$$\xi_{n\sigma} = \frac{1}{N_x} \frac{1}{N_{\text{fits}}} \sum_{j=1}^{N_x} \sum_{l=1}^{N_{\text{fits}}} I_{[-n\sigma_{\text{fit}}^{(l)}(x_j), n\sigma_{\text{fit}}^{(l)}(x_j)]} \left(\langle f_{\text{fit}}^{(l)}(x_j)\rangle - f_{\text{in}}(x_j) \right). \tag{4.8.4}$$

  This measure was created so that one can assess the accuracy of the uncertainties of the fitted neutrino PDFs with respect to the input neutrino PDF. The idea is that the the values of the

input neutrino PDF fall 68.3% of the time in the one sigma uncertainty band of the fitted neutrino PDF. So, $\xi_{1\sigma} \approx 0.68$. Again, one averages over several fits in this equation, which means one has to compute multiple fits each consisting of at least 100 replicas. However, because this can be computationally intensive an approximation would be to assume the neutrino PDF uncertainties between different closure test do not vary significantly. Then the accuracy becomes:

$$\xi_{n\sigma} = \frac{1}{N_x} \sum_{j=1}^{N_x} I_{[-n\sigma_{\text{fit}}(x_j), n\sigma_{\text{fit}}(x_j)]} \left(\langle f_{\text{fit}}(x_j) \rangle - f_{\text{in}}(x_j) \right), \tag{4.8.5}$$

where the x-points consist of 10 logarithmically spaced points between 10^{-4} and 0.1 and 10 linearly spaced points between 0.1 and 1. In the equation, $\langle f_{\text{fit}}(x_j) \rangle$ and $\sigma_{\text{fit}}(x_j)$ is the mean PDF and the standard deviation of the neutrino PDF obtained by the closure fit, respectively. The factor $I_{[-n\sigma_{\text{fit}}(x_j), n\sigma_{\text{fit}}(x_j)]}$, is an indicator function. This function is one if the fitted neutrino PDF evaluated on x_j falls within 1σ of the input neutrino PDF, and zero otherwise.

- The final postfit measure implemented in this work is the bias-to-variance ratio. This measure is implemented to assess the faithfulness of the PDF uncertainties and to see whether the uncertainties of the data are correctly propagated in the uncertainty of the model predictions. Again, the expectation value of these quantities is computed using different level-1 shifts, however, one can also compute this quantity for a single level-2 closure test. The bias is defined as:

$$\text{bias} = \frac{1}{N_{\text{rep}}} \frac{1}{N_{\text{dat}}} \sum_k \left[\sum_{ij} \left(\langle N^k_{\text{rep}} \rangle - N_{\text{input}} \right)_i C^{-1}_{ij} \left(\langle N^k_{\text{rep}} \rangle - N_{\text{input}} \right)_j \right], \tag{4.8.6}$$

and represents the difference between the mean predictions of the model and the theoretical predictions in units of the covariance matrix. The variance is given by:

$$\text{variance} = \frac{1}{N_{\text{dat}}} \frac{1}{N_{\text{rep}}} \sum_k \left[\sum_{ij} \left(\langle N_{\text{rep}} \rangle - N^k_{\text{rep}} \right)_i C^{-1}_{ij} \left(\langle N_{\text{rep}} \rangle - N^k_{\text{rep}} \right)_j \right] \tag{4.8.7}$$

and signifies the fluctuations of each fitted replica with respect to the mean prediction in units of the covariance matrix i.e. how much each fitted replica differs from the mean prediction. Now, when dividing the bias by the variance, this creates a measure which tells one how well the uncertainty of the data is propagated to the PDF predictions and checks if the average distance between the mean prediction and the theoretical prediction is the same as the variance across replicas. Thus if

$$\frac{\text{bias}}{\text{variance}} = 1, \tag{4.8.8}$$

it means the uncertainties of the fitted neutrino PDFs are 'faithful' and can thus be trusted.

Now the postfit measures of the fit shown in Figure 4.9 can be computed to assess the quality of the fit. From Table 1 it can be seen that most postfit measures are indicating that there is some degree of overfiting. However, considering there are 6 data points it is hard to make a distinction between fitting the noise and fitting data. One can be certain about one thing: the model is accurately fitting the data. The accuracy metric is indicating the accuracies of the input PDF are correctly reproduced. Besides computing these postfit measures, several level 2 closure tests with different level 1 instances are also executed to see if this will reproduce the level 0 data. Figure 4.14 shows a comparison between the flux from a level 0 closure test, the underlying flux used to generate the pseudo-data and the mean

Estimator	Value
Δ_{χ^2}	-0.88
ξ_σ	0.67
$\langle \mathrm{TL} \rangle_{\mathrm{rep}}$	1924
$\langle \chi^2_{\mathrm{tot}} \rangle_{\mathrm{rep}}$	0.65
$\varphi^2_{\chi^2}$	5.03
$\frac{\text{bias}}{\text{variance}}$	0.12

Table 1. The postfit estimators corresponding to a single level 2 closure test displayed in Figure 4.14. $\langle \mathrm{TL} \rangle_{\mathrm{rep}}$ is the average training length.

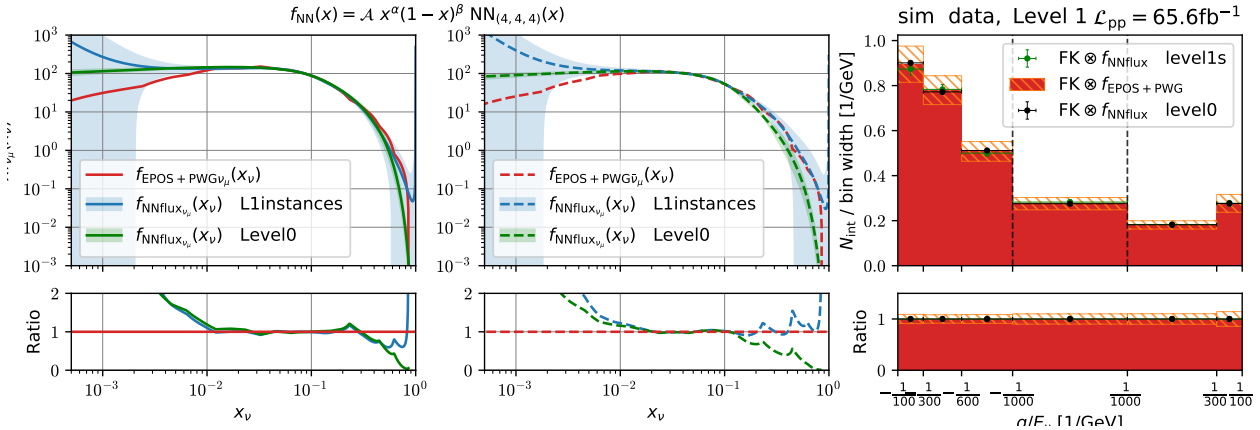


Figure 4.14. Left and middle plot: comparison between the flux from a level 0 closure test, the underlying flux used to generate the pseudo-data and the mean flux of ten level-2 closure tests using different level 1 instances. The fluxes from the closure test are obtained by fitting to the same pseudo data as was done above. The bottom-left and bottom-middle panels display the ratio between the fitted and input fluxes, where the bands indicate the 68% CL interval in the extracted fluxes. Right: comparison between the input pseudo-data for the N_{int} distributions as a function of q/E_ν entering the closure test with the predictions based on the fitted fluxes and the convolution of the fitted fluxes and the FK-table. For the input pseudo-data, uncertainties are assumed to be purely statistical. For the theoretical predictions, the error band is the 68% CL range evaluated from the $N_{\mathrm{rep}} = 100$ Monte Carlo replicas.

flux of ten level 2 closure tests using different level 1 instances. The fluxes from the closure test are obtained by fitting to the pseudo data. The bottom-left and bottom-middle panels display the ratio between the fitted and input fluxes, where the bands indicate the 68% CL interval in the extracted fluxes. The right panel of the plot shows the comparison between the input pseudo-data for the N_{int} distributions as a function of q/E_ν entering the closure test with the predictions based on the fitted fluxes and the convolution of the fitted fluxes and the FK-table. From Fig 4.14, it can be seen that the average over several level 2 closure tests with different level 1 instances agrees with a level 0 closure test. This is as expected, since several level 2 closure tests can be seen as a level 0 closure test. Another expectation which should be verified is the one of homogeneity between level 2 closure tests. To verify this expectation, two plots have been produced, which can be seen in Figure 4.15 and 4.16. From these Figures, it can be seen that the χ^2s are fluctuating around 0.4 and the relative uncertainty barely fluctuates in the x_ν-region where there is data. All in all, the Figures show the closure tests are successful and yield stable results.

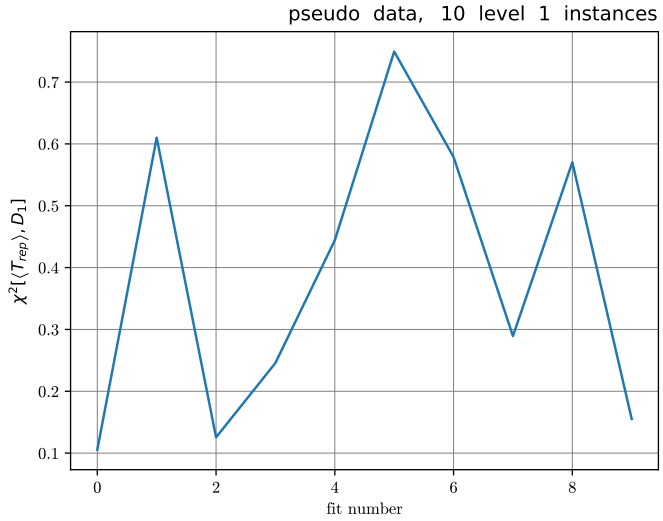


Figure 4.15. This plot shows the χ^2s for several L2 closure tests with different level 1 instances. The χ^2s are computed between the mean prediction and the corresponding level 1 data instance.

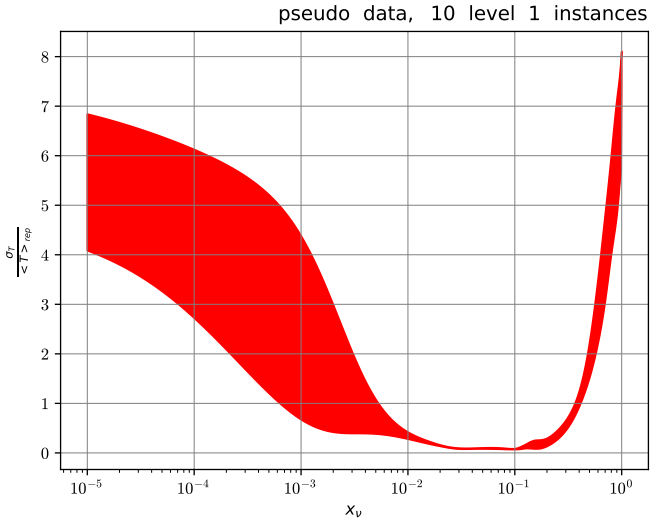


Figure 4.16. This plot shows the relative uncertainties for several L2 closure tests with different level 1 instances. The relative uncertainty is given by $\frac{\sigma_{L2\ \text{fits}}}{\langle L2\ \text{fits}\rangle}$. It can be seen that the variance in neutrino PDFs between several level 2 closure tests is low in regions where there is data.

5 Applying the Numerical Framework to Forward Particle Production at the LHC

Now that we a data-driven parametrisation of the neutrino flux using event rate measurements from FASER is performed, the framework presented above can be applied to several applications. First of all, the determination of the LHC neutrino fluxes can be used to constrain models of forward hadron production, since LHC neutrinos are primarily decay products from pions, kaons and charm hadrons. Forward hadron production is a process which produces hadrons in the forward region after a proton-proton collision i.e. at small angles with respect to the beam line. As mentioned in the introduction, constraining models of forward hadron production is a way to understand QCD in the small-x regime and can contribute to improved predictions for the neutrino telescopes looking for prompt neutrino fluxes. In this section it will be shown that by parametrising neutrino fluxes, certain models are disfavoured. Not only fluxes from different generators are compared, also fluxes from different detector geometries and kinematic observables are compared, which can tell us about the suitability of the kinematic observables and the several detector geometries. Another application of the framework that was researched, is looking at BSM enhanced decays of neutral mesons into neutrinos and comparing these BSM scenarios to SM scenarios for several detector geometries. The final application considered is the effect of the charm quark in the proton on the for of neutrino fluxes. For all applications, the constraints that FASERν(2) can provide on the fluxes is also researched.

5.1 Testing Forward Event Generators and Detector Geometries

In this section the framework developed will be applied to test several forward event generators for several detector geometries for several kinematic observables ($E_h.E_\ell, E_\nu, \theta$).

5.1.1 Detector Geometries

Several detector geometries will be considered to analyse the potential to constrain neutrino fluxes for current and proposed detectors. The following detectors will be considered:

- **FASER 2024** This detector geometry was used to measure the event rates presented in [3] with an integrated luminosity of 65.6fb^{-1}. The geometry of this detector consists of a cylindrical fiducial volume with a radius of 10 cm and a depth of 88.5 cm of tungsten.
- **FASERν at Run 3** This detector uses the FASERν emulsion detector to measure neutrinos and it has dimensions of 25 cm x 30 cm x 80 cm of tungsten. The integrated luminosity at the end of run 3 is expected to be around 150 fb^{-1}.
- **FASERν at High Luminosity LHC** The FASER experiment will continue to take data during LHC run 4, meaning we also consider the FASERν detector in the high luminosity area meaning an integrated luminosity of 3 ab^{-1}
- **FASERν2** If the proposed FPF will take place, there will be a FASERν2 detector, which will be sensitive to many more neutrinos because of its larger size. This experiment will be located 620 m from ATLAS and have dimensions 40 cm x 40 cm x 6.8 m. Since this experiment is designed to detect neutrinos in the high luminosity, an integrated luminosity of 3 ab^{-1} is assumed.

5.1.2 Event Generators

Several different QCD models will be considered and, using the parametrisation of the neutrino PDF computed in the previous section, improved predictions will be derived. In the literature there are a range of Monte Carlo generators available which are used to simulate and describe forward hadron production. Using these models one can generate several neutrino fluxes for different detector geometries. These fluxes are then compared to the neutrino flux obtained in the previous section from event rate measurements from FASER. In doing this comparison one can constrain and thus disfavour certain models of forward hadron production. Besides comparing the models to the neutrino flux parametrisation using the FASER measurements, they will also be compared to each other for several detector geometries and kinematic observables. A range of Monte Carlo generators and theoretical models are used to describe and simulate forward hadron production, depending on the physics focus. Here's a breakdown:

- **EPOS-LHC** [44] stands for **E**nergy conserving quantum mechanical approach, based on **P**artons, parton ladders, strings, **O**ff-shell remnants, and **S**plitting of parton ladders. `EPOS` is an MC event generator for hadronic interactions and can model both heavy ion interactions and cosmic ray collisions. `EPOS` uses both a parton based description i.e. simulating parton interactions as well as parts of the Regge theory. `EPOS` uses a so-called Parton-Based Gribov-Regge Theory, which is an effective theory. The main idea of this theory is that the elementary interactions are described by so-called Pomeron exchanges. Pomerons represent elementary interactions and is a phenomenological object in the Gribov-Regge theory. Contrary to this theory, `EPOS'` energy conservation is considered in calculations of cross-sections and hard-processes are implemented in a natural way compared to the parton model.

- **QGSJET 2.04** [45] is an acronym for Quark-Gluon String model with JETs and is, similarly to `EPOS`, also based on the Regge theory and was created to have a framework which can reliably compute interactions at energies above collider energies. Also in the `QGSJET` Pomerons represent the elementary interactions. The Quark-Gluon string comes from the fact that `QGSJET` in the process of a proton-proton collision two quark-gluon strings form. This is because the protons in the collision are modeled as consisting of a quark and a diquark. In the collision the quark from one proton interacts with the diquark of the other proton and vice versa. This process then leads to the formation of two quark-gluon strings.

- **Sibyll 2.3d** [46] (no acronym) is an event generator which was created to accurately simulate high energy cosmic ray cascades based on a physical model. Similarly to `QGSJET` and `EPOS` it uses the Regge theory, however, `Sibyll` is based on the dual parton model using the minijet and Lund models. The details of these models do not fit the tone of this thesis and, for the interested reader, explanations of these models can be found in the literature.

- **DPMJET 2019** [47] (Dual Parton Model JET). `DPMJET` is based on the Dual Parton Model (DPM), which itself is grounded in Regge theory. Similar to `EPOS` and `QGSJET`, the model uses Pomeron exchanges to represent elementary interactions. In the DPM framework, each Pomeron corresponds to a colorless exchange leading to the formation of two color strings, which then hadronize into observable particles. `DPMJET` extends the basic DPM formalism by including multiple scattering processes and the treatment of both soft and hard interactions.

Different parts of the event generation are modelled by different MC generators depending on the

physics. The light hadron production is the dominant component of particle production in high-energy collisions and the production of the charm quarks are heavier and less abundant, but crucial for certain types of neutrino flux, like prompt neutrinos. The generators mentioned above will be used to simulate the production of light hadrons and `SIBYLL 2.3d` and `DPMJET 2019` will also be used to model forward charm production. For `QGSJET 2.04` and `EPOS-LHC`, the event generator `POWHEG` will be used to model the production of charm hadrons matched with the `PYTHIA 8.3` [48] parton shower. In short, pseudo will be computed for the following models:

- `EPOS`(π,K)+`POWHEG`(c)
- `QGSJET`(π,K)+`POWHEG`(c)
- `Sibyll`(π,K,c)
- `DPMJET`(π,K,c)

and for the following detector geometries:

- **FASERν2**
- **FASERν at High Luminosity LHC**
- **FASERν at Run 3**
- **FASER 2024**

FK-tables will be generated for four observables: E_h, E_l, E_ν, and θ. Using these four observables one can parametrise four fluxes and compare them to each other to assess the relative suitability of each observable.

5.1.3 Updated Methodology

Training/Validation split Up until this point the number of data points that the neutrino flux is fitted to was very limited, thus also limiting the machine learning techniques that were implemented. Fortunately, the number of events drastically increases to about $O(10^2)$ for the detector geometries with an integrated luminosity of 3 ab^{-1}. This means one can split the dataset into a training and a validation set, which is critical to avoid overfitting when using ML on large(r) datasets. In this work, a validation set of 10% is used for the high luminosity detector geometries. As a consequence the validation loss is used in the stopping algorithm, instead of the training loss used previously. Also the postfit measures are computed using the validation set instead of the training set.

Fitting electron fluxes The FASER experiment could distinguish scattering events coming from μ and $\bar{\mu}$ neutrinos up to 1000 GeV. However, distinguishing between electron and anti-neutrinos is impossible, because their decay products in terms of showers look exactly the same. For the fits muon neutrino fits μ and $\bar{\mu}$ were fitted in one network, but were considered separate. In order to fit the electron neutrino flux and in order to stay true to the event rate measurements from FASERν in the future, it was decided to fit $f_{\nu_e} + f_{\bar{\nu}_e}$ instead of fitting them separately, and thus make the assumption $f_{\nu_e+\bar{\nu}_e} = f_{\nu_e} + f_{\bar{\nu}_e}$. The event rates therefore contained both the neutrino and anti-neutrino scattering events. The output of the NN was convoluted with a combined (summed) FK-table. The plots below thus show the combined neutrino PDF of the (anti-) electron neutrino.

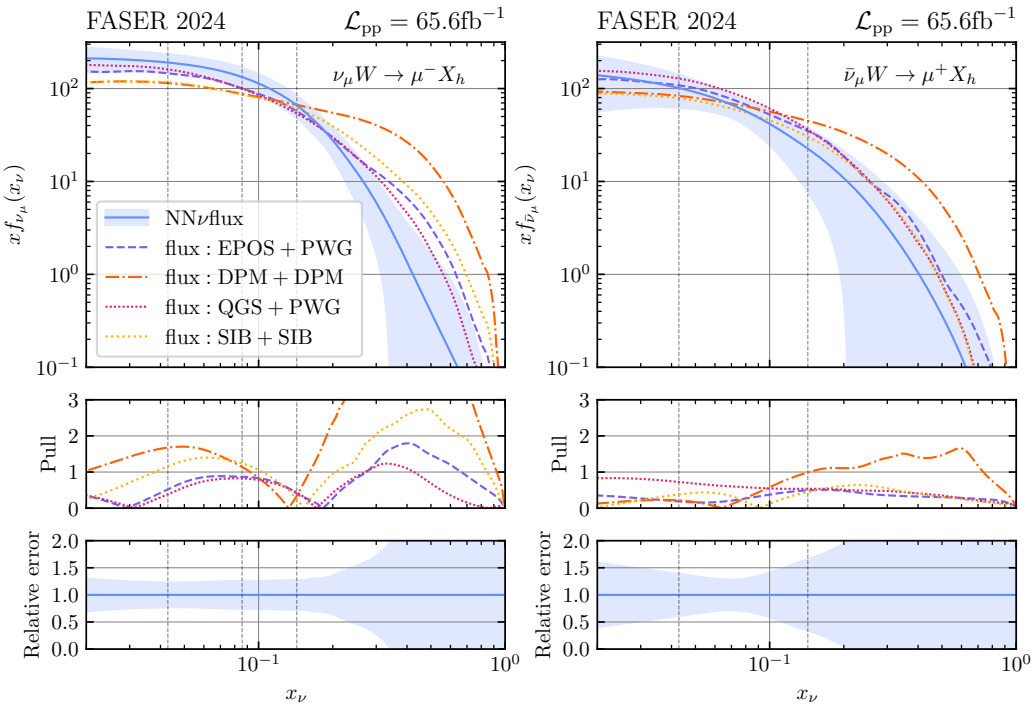


Figure 5.1. The results of the determination of the muon neutrino and antineutrino PDFs, $x_\nu f_{\nu_\mu}(x_\nu)$, from the FASER 2024 measurements of the event yields binned in E_ν, where the band indicates the 68% CL uncertainties. The results of the fit are compared with the predictions based on four different fluxes as discussed above `EPOS+POWHEG` (baseline flux in [3]), `DPMJET 2019`, `QGSJET+POWHEG` and `SIBYLL`. The middle panels indicates the pull between each of these four flux predictions and the fit results, see text for more details. The bottom panels indicated the relative error of the fit result.

5.1.4 Comparison Neutrino Flux from FASER Event Rate Measurements to MC Event Generators

The neutrino PDF that was extracted using the FASER event rate measurements for muon neutrino scattering can be compared to the several forward hadron production models, which are described above. This Figure can be seen in 5.1. The upper panel shows that the uncertainties on the flux is around 25%, corresponding to the uncertainties on the FASER event rate measurements. The bottom panels show how many sigmas the models differ from the fit result. It can be seen that all models except the `DPMJET 2019` agree well with the neutrino PDF extracted from the data within uncertainties. The `DPMJET 2019` model diverges from the data in the final bin i.e. in $x_\nu > 0.2$ ($E_\nu > 1$)TeV, because it predicts a much higher rate of neutrinos than what is observed by FASER. From this one can conclude that the `DPMJET 2019` model is disfavoured by the FASER data, however, more event rate measurements are needed to fully exclude the model. This model treats charm quarks as massless and because of this, more charms are produced, especially in the forward region. Since most neutrinos in the forward region are decay products of charmed hadrons, `DPMJET` displays an excess of neutrinos compared to the other event generators. From this, it can be concluded that the `DPMJET` model is less suitable as an event generator, not only in this work, but also as a theoretical model for the several neutrino telescopes. Also the `SIBYLL` model predicts an excess of neutrinos, although this difference is a maximum of 3σ. This comparison plot is an excellent result considering we are dealing with low statistics.

5.1.5 Results Comparison Detector Geometries, Event Generators and observables

In this section, plots will be shown which compare several parametrised neutrino PDFs for different detector geometries, event generators and observables. First of all, Figures 5.2 and 5.3 show a comparison between muon and electron neutrino fluxes obtained from four different observables. It can be seen that all closure tests are successful, irrespective of the kinematic input. In both plots the middle panel shows that the E_ν observable produces the flux with the smallest uncertainty and that θ_ℓ produces the 'worst' fit and the flux with the highest uncertainty. The closure tests for the θ_ℓ observable are all successful, however it seems that the input flux is not always reproduced. This makes sense, since E_ν is directly related to the input neutrino flux and the lepton scattering angle θ_ℓ is is only loosely correlated with the E_ν dependence on the flux i.e. the fk-table for this kinematic observable is much less diagonal than for the other kinematic observables. Or, in other words the flux as a function of the normalized neutrino energy obtained is relatively loosely correlated to the event rates as a function of the lepton scattering angle. In short, θ_ℓ is not a suitable kinematic observable to fit the neutrino flux to. The distributions as a function of E_h E_ℓ yield comparable uncertainties to E_ν signifying that these kinematic variables have, as mentioned in the introduction, a much stronger correlation to the neutrino energy.

Figures 5.4 and 5.5 show plots comparing muon and electron fluxes for different event generators for all four kinematic observables. As showed in Figure 5.1 and mentioned in [3], it can be seen that the `DPMJET 2019` model is incompatible with the other event generators from the bottom panel. In this panel the Pull between the several generators and the `EPOS+POWHEG` is plotted, which is defined as:

$$P_{\text{Gen}}(x_\nu) = \frac{\left|f^{(\text{EPOS})}_{\nu_e+\bar{\nu}_e}(x_\nu) - f^{(\text{Gen})}_{\nu_e+\bar{\nu}_e}(x_\nu)\right|}{\sqrt{\sigma^{(\text{EPOS})2}_{\nu_e+\bar{\nu}_e} + \sigma^{(\text{Gen})2}_{\nu_e+\bar{\nu}_e}}}\,. \tag{5.1.1}$$

If the P=1, it means that the two event generators models agree with each other within uncertainties. This measure can quantify the statistical compatibility between the several fluxes, because it takes both the fit uncertainty from bith fits into account. From these Figures it can be seen that `DPMJET` is incompatible with the other event generators, because a pull up to ~ 20 can be seen. The highest pull is seen between $x_\nu = 0.1$ and 0.6. When future data will be available from the high luminosity era for either the FASERν or FASERν2 detector geometry, the model will probably be fully excluded up to several sigmas.

Figures 5.6 and 5.7 show plots comparing the neutrino fluxes for different detector geometries. From these plots, the projected neutrino PDF for different detector geometries can be seen. It is clear that the flux increases with a higher luminosity and for the larger FASERν2 detector at the FPF. Because of this, the relative uncertainty of the neutrino flux drastically decreases i.e. the flux is more accurately constrained. As mentioned above, this will help to constrain and exclude models of forward hadron production. These plots also show that the framework again correctly propagates the uncertainties of the data to the neutrino fluxes.

All these plots show that the neutrino fluxes obtained from the fits reproduce the input fluxes. The fits are also successful in reproducing the L1 instances of the pseudo data within uncertainties. An example of a successful closure test is shown in Figure 5.8, where the L1 data and its uncertainties are recovered. Also, we have verified that several L2 CL using different L1 seeds yield the same result as a L0 CL. As mentioned above, there are also statistical estimators to further quantitatively assess the quality of the fit. Table 2 displays the several statistical estimators for three fits, which show the fits are successful.

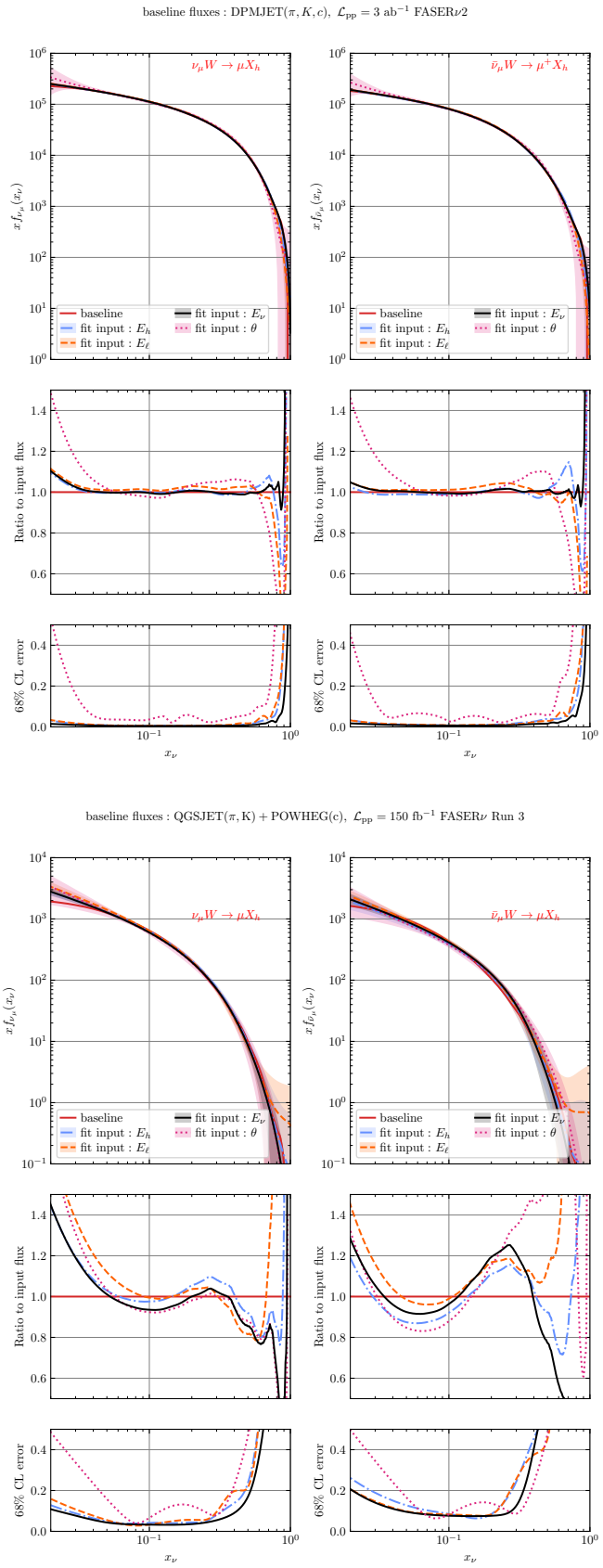


Figure 5.2. These plots compare muon neutrino fluxes obtained from four different observables. The upper plot compares observables for the FASERν2 detector geometry and the `DPMJET 2019` model. The lower plot compares observables for the FASERν detector geometry and the `QGSJET+POWHEG` model. The upper panel shows the neutrino PDFs and the baseline flux used to generate the data. The middle panel shows the ratio between the baseline and the parametrised neutrino PDF and the lower panel shows the 68% closure test error. In both plots it can be seen that the $E\nu$ observable produces the flux with the lowest uncertainty and that θ_ℓ produces the flux with the highest observable.

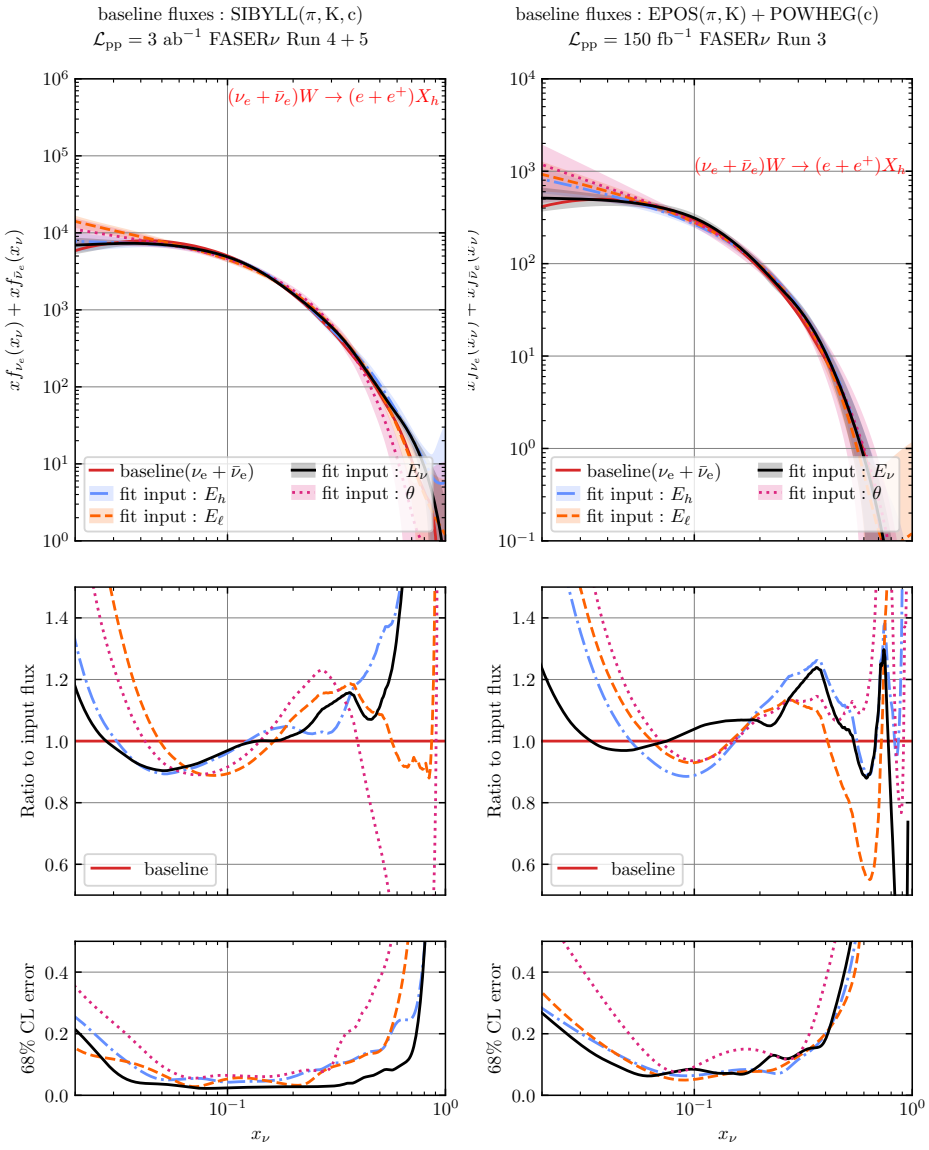


Figure 5.3. This Figure shows two plots comparing electron neutrino fluxes obtained from four different observables. The left plot compares observables for the FASERν detector geometry 3 ab^{-1}and the `SIBYLL` model. The right plot compares observables for the FASERν detector geometry 150 fb^{-1}and the `EPOS+POWHEG` model. The upper panel shows the neutrino PDFs and the baseline flux used to generate the data. The middle panel shows the ratio between the baseline and the parametrised neutrino PDF and the lower panel shows the 68% closure test error. In both plots it can be seen that the $E\nu$ observable produces the flux with the lowest uncertainty and that θ_ℓ produces the flux with the highest observable.

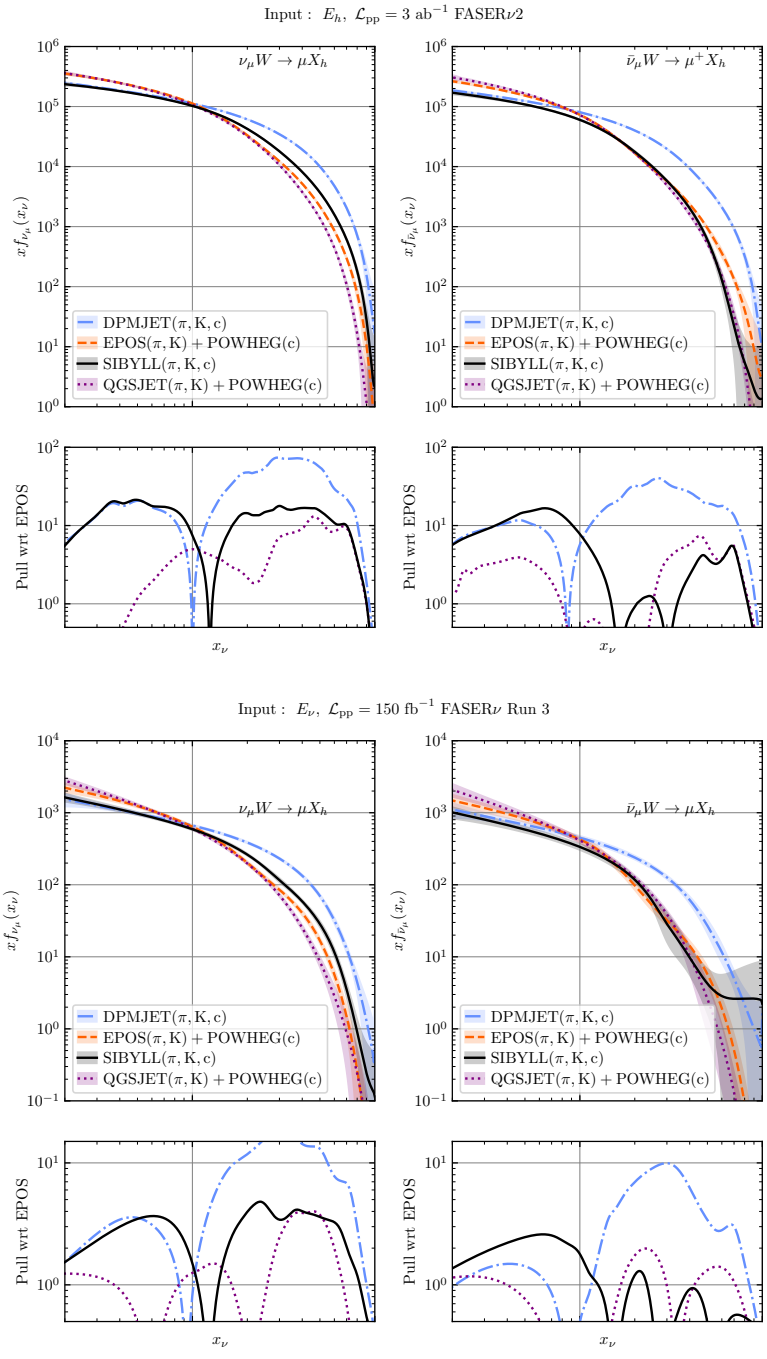


Figure 5.4. This Figure shows two plots comparing muon neutrino fluxes obtained from four different MC generators. The upper plot compares observables for the FASERν2 geometry and the E_h observable. The lower plot compares observables for the FASERν geometry and the $E\nu$ observable. The upper panel shows the fitted neutrino PDFs for all generators. The lower panel shows the pull between the different fitted fluxes and the fitted `EPOS+POWHEG` flux. In the lower panel it can be seen that the `DPMJET 2019` model is incompatible with the other models, because of a high pull.

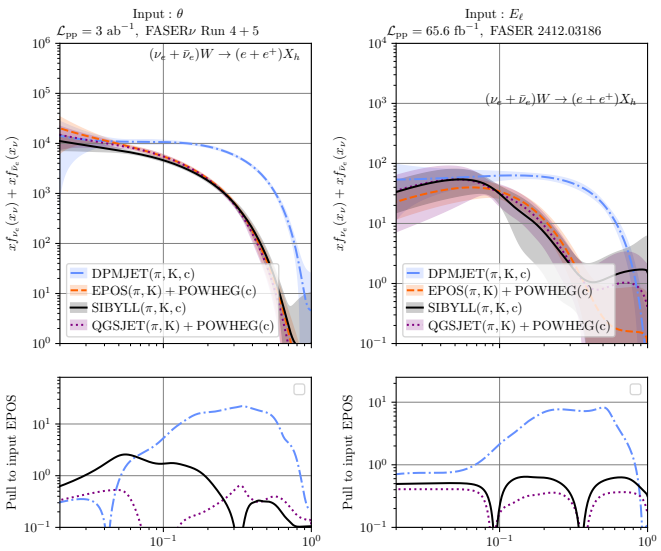


Figure 5.5. This Figure shows two plots comparing electron neutrino fluxes obtained from four different MC generators. The left plot compares observables for the FASERν geometry high lumi and the θ_ℓ observable. The right plot compares observables for the FASER 2024 geometry and the E_ℓ observable. The upper panel shows the fitted neutrino PDFs. The lower panel shows the pull between the different fitted fluxes and the fitted `EPOS+POWHEG` flux. In the lower panel it can be seen that the `DPMJET 2019` model is incompatible with the other models.

5.2 Researching BSM Enhanced Decays of Neutral Mesons into Neutrinos

Another physics application of the framework will be to constrain BSM scenarios with enhanced branching fractions of neutral hadrons into neutrinos. The branching fractions of these decays in the SM are quite low compared to the current bounds on the branching ratios. For this decay:

$$\pi \to \nu_\mu + \bar{\nu}_\mu, \tag{5.2.1}$$

the SM constraint is zero, since the decay is forbidden by angular momentum conservation with massless left-handed neutrinos. However, when using massive neutrinos, the SM constraint is 10^{19} times lower than the current bound. Since an increased branching ratio produces significantly more neutrinos in the forward region at the LHC, the FASER collaboration could improve current constraints on these BSM decays. Here, we will consider three decays with enhanced neutrino production:

- $\eta \to \nu_\mu + \bar{\nu}_\mu(\nu_e + \bar{\nu}_e)$ $\quad$ BR$(\eta \to \nu_e + \bar{\nu}_e) \leq 10^{-4}$
- $\eta\prime \to \nu_\mu + \bar{\nu}_\mu(\nu_e + \bar{\nu}_e)$ $\quad$ BR$(\eta' \to \nu_e + \bar{\nu}_e) \leq 6 \times 10^{-4}$
- $\pi \to \nu_\mu + \bar{\nu}_\mu(\nu_e + \bar{\nu}_e)$ $\quad$ BR$(\pi^0 \to \nu_e + \bar{\nu}_e) \leq 1.7 \times 10^{-6}$

The branching ratios (BR) are be taken according to the Particle Data Group (PDG) 2024 limit [49]. Figure 5.9 shows the number of interacting neutrinos at FASERν for the several BSM scenarios. It can be seen that for the electron neutrino flux, a significant contribution from these BSM decays to the flux can be seen. More specifically, the contribution from the BSM decays even exceeds the SM contribution around $E_\nu = 3$ TeV. This effect would definitely be visible by the FASER experiments, which could, as mentioned above, improve current constraints on the branching ratios. For the muon neutrinos the contribution of the BSM decays at high energies is much smaller due to the large contribution from kaon decays, as was mentioned in the Section 2. The branching ratios for η and $\eta\prime$

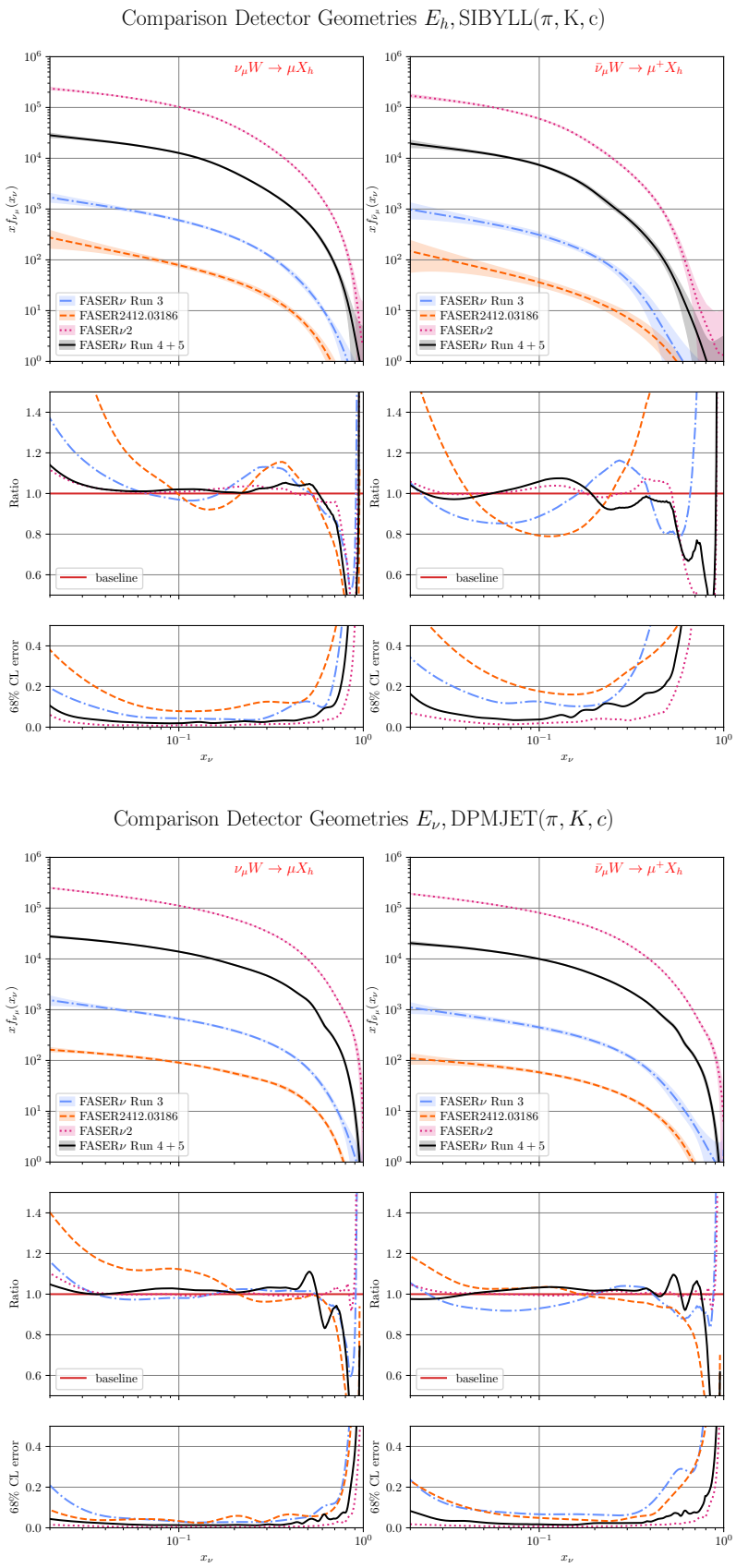


Figure 5.6. This Figure shows two plots comparing muon neutrino fluxes obtained from four different detector geometries. The upper plot compares observables for the `Sibyll` generator and the E_h observable. The lower plot compares observables for the `DPMJET 2019` model and the E_ν observable. The upper panel shows the neutrino PDFs and the baseline flux used to generate the data. The middle panel shows the ratio between the baseline and the parametrised neutrino PDF and the lower panel shows the 68% closure test error. In the lower panel it can be seen that the relative error decreases for the high luminosity neutrino fluxes.

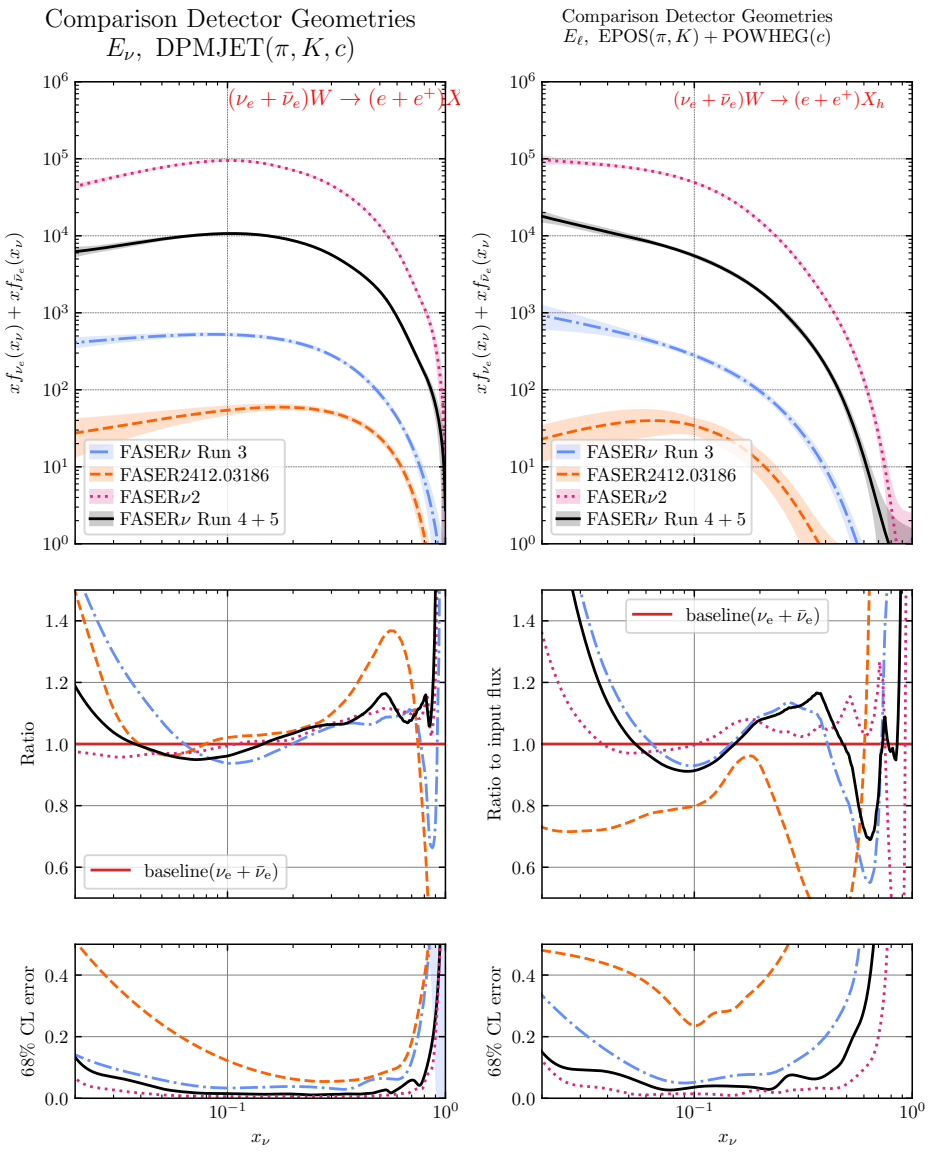


Figure 5.7. This Figure shows two plots comparing electron neutrino fluxes obtained from four different detector geometries. The upper plot compares observables for the `DPMJET 2019` generator and the E_ν observable. The lower plot compares observables for the `EPOS+PWOHEG` model and the E_ℓ observable. The upper panel shows the neutrino PDFs and the baseline flux used to generate the data. The middle panel shows the ratio between the baseline and the parametrised neutrino PDF and the lower panel shows the 68% closure test error. In the lower panel it can be seen that the relative error decreases for the high luminosity neutrino fluxes.

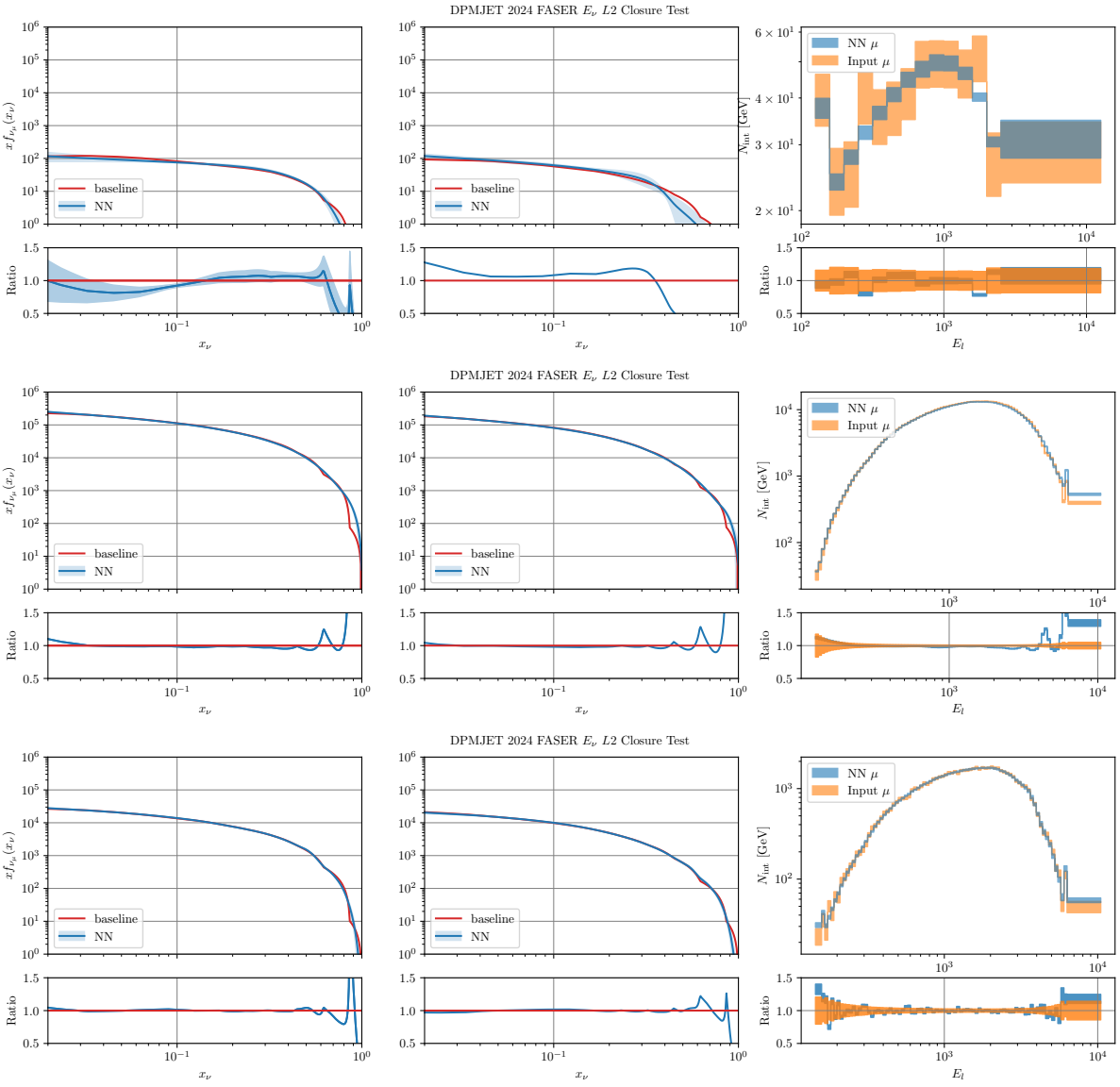


Figure 5.8. These plots show successful L2 closure tests. The events were generated using the `DPMJET 2019` model. Both on the level of the neutrino PDF as well as on the level of the event rates, the CL is successful, since it reproduces the input within uncertainties.

Estimator	Closure Test		
	FASER 2024 (E_ℓ)	FASER HL-LHC (E_ν)	FASERν2 (E_ν)
χ^2	0.15	0.9	0.73
Δ_{χ^2}	−0.93	−0.9	−0.98
ξ_σ	0.67	0.5	0.54
$\langle \mathrm{TL} \rangle_{\mathrm{rep}}$	3105	13018	14928
$\langle \chi^2_{\mathrm{tot}} \rangle_{\mathrm{rep}}$	1.14	1.9	2.02
$\varphi^2_{\chi^2}$	3.6	4.38	5.56

Table 2. Statistical estimators computed of a representative subset of the closure test studies validating our fitting methodology and shown in Figure 5.8. In all cases the input neutrino flux used to generate the synthetic data is the same, namely `DPMJET 2019`. From both the closure tests and from this Table it can be seen that the fits are successful.

are about 100 times higher than for π. This explains why there is no significant increase in neutrinos for the enhanced branching ratio of the decay of pions to neutrinos.
Since the BSM should be clearly visible in the neutrino flux, pseudo data was generated using Figure 5.9 for the FASERν Run 3 detector geometry. The neutrino fluxes were generated using `EPOS+POWHEG` in the SM and the three enhanced BSM fluxes were added separately to these fluxes. Then, closure tests for the electron and muon neutrino PDFs were carried out using this data. Figure 5.10 shows the results of the neutrino PDFs from the fits for the E_ℓ kinematic variable. First of all, it can be seen that all closure tests are successful, since the input flux agrees with the fitted flux within uncertainties in regions where there is data in the second panel. It can also be seen that fitted BSM fluxes do not show a clear enhancement compared to the SM flux, as was expected from Figure 5.9. However, the electron neutrino flux is clearly several sigmas higher than the SM flux in a large x_ν region. The bottom panel shows the pull between the BSM and the SM fit and, from Figure 5.10, it can be seen that the electron flux has a pull around 4-5σ in the x_ν region between 0.15 and 0.4 (or between $E_\nu =$ 1 and 3 TeV) for the η and $\eta\prime$ BSM decays, while the π^0 decay agrees with the SM fit. This effect is even larger for the FASERν detector geometry in the high luminosity era, where a pull of 10 can be seen for the η and $\eta\prime$ BSM decays. This is because the uncertainties on the neutrino PDFs decrease, as is shown in Figure 5.11. All in all the parametrised fluxes show that the measurements of electron fluxes at FASERν can potentially improve the current PDG bounds on enhanced decays of η and $\eta\prime$ mesons into neutrinos.

5.3 The effect of intrinsic charm in Forward D-Meson Production

The final physics application considered in this work is the effect of intrinsic charm in forward D-meson production. The NNPDF collaboration published a paper [50] in which they fitted PDF sets including charm. They found that assuming that the production of charmed hadrons at the LHC is dominated by gluon fusion to charm quarks: $g + g \to c + \bar{c}$, does not hold. The contribution from quark-initiated processes such as $c + g \to c + X$ i.e. an intrinsic charm component, cannot be neglected when PDF sets including charm are fitted[51]. This contribution from gluon-charm scattering to D-meson production thus becomes sizeable in the forward region compared to the gluon fusion channel[52]. Here, we obtain

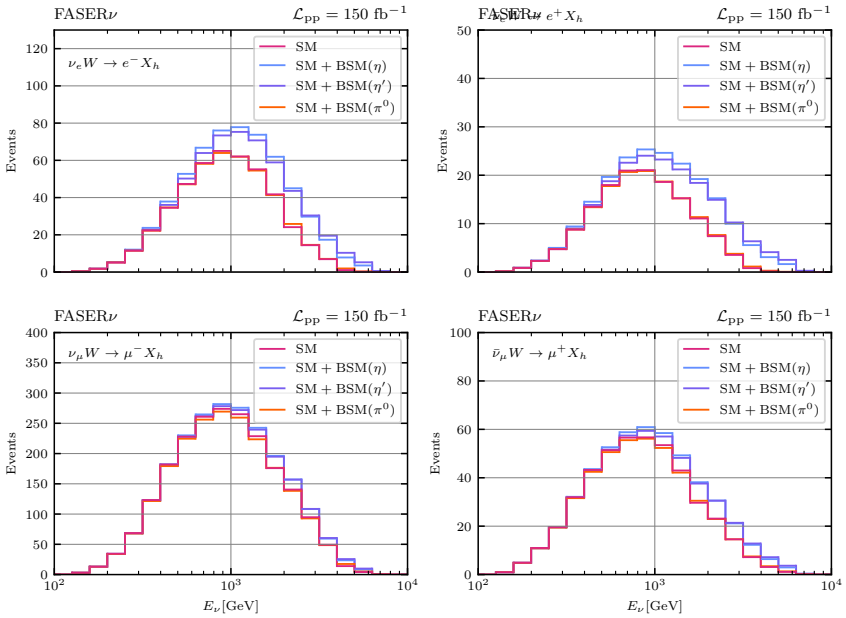


Figure 5.9. The number of interacting neutrinos at FASERν evaluated with `EPOS-LHC`. The BSM scenarios, where the branching ratios of the light neutral mesons $(\pi^0, \eta, \eta^{'})$ have been taken according to he PDG bounds, are compared to the SM prediction. It can be seen that for the electron neutrino the BSM scenario has a significant effect on the number of interacting neutrinos.

an estimate of the sensitivity of FASER to the effect of intrinsic charm in forward D-meson production using synthetic data. We also try to constrain the charm quark content of the proton by looking at the parametrised electron and muon fluxes.

Unfortunately the `POWHEG` framework could not be used to generate events including intrinsic charm, since charm is not included in the initial state. This is why in this work the forward neutrino production from the $g + c$ channel from [53] is added to the `POWHEG+EPOS` predictions of [54]. Because this section is used to explore a rough estimate of the sensitivity of FASER to intrinsic charm, this approach will do for now. However, when comparing to experimental data from FASER it is important to do a proper job and use NLO calculations to account for intrinsic charm.

Figure 5.12 shows the number of neutrino charged current interactions for the FASER$\nu(\mathcal{L}_{pp} = 3\ \text{ab}^{-1})$ and FASERν2 detector geometries for electron and muon neutrinos. Again, the baseline flux is generated using `EPOS` for light hadrons and `POWHEG` for heavy hadron production. On top of this the contribution from $Q + g \to Q + X$ [53] is added where the intrinsic charm carries $\sim 1\%$ of the proton's momentum. From the plots it can be seen that the expected number of events almost doubles for electron neutrinos for both geometries in the region around $E_\nu = 1$TeV. This effect is smaller for muon neutrino scattering events, because most muon neutrinos are produced from light hadrons and the contribution from D-mesons plays a subdominant role in this.

Figure 5.13 shows the fitted electron neutrino PDF to the event rates showed in Figure 5.12 for the FASER$\nu(\mathcal{L}_{pp} = 3\ \text{ab}^{-1})$ and FASERν2 detector geometries compared to the fitted electron neutrino PDF without the $c + g$ contribution. In this case the pull is defined as:

$$P_{\rm IC}(x_\nu) = \frac{\left| f^{(\rm gQ,IC)}_{\nu_e+\bar{\nu}_e}(x_\nu) - f^{(\rm no\ gQ)}_{\nu_e+\bar{\nu}_e}(x_\nu) \right|}{\sqrt{\sigma^{(\rm gQ,IC)2}_{\nu_e+\bar{\nu}_e} + \sigma^{(\rm no\ gQ)2}_{\nu_e+\bar{\nu}_e}}} \,. \tag{5.3.1}$$

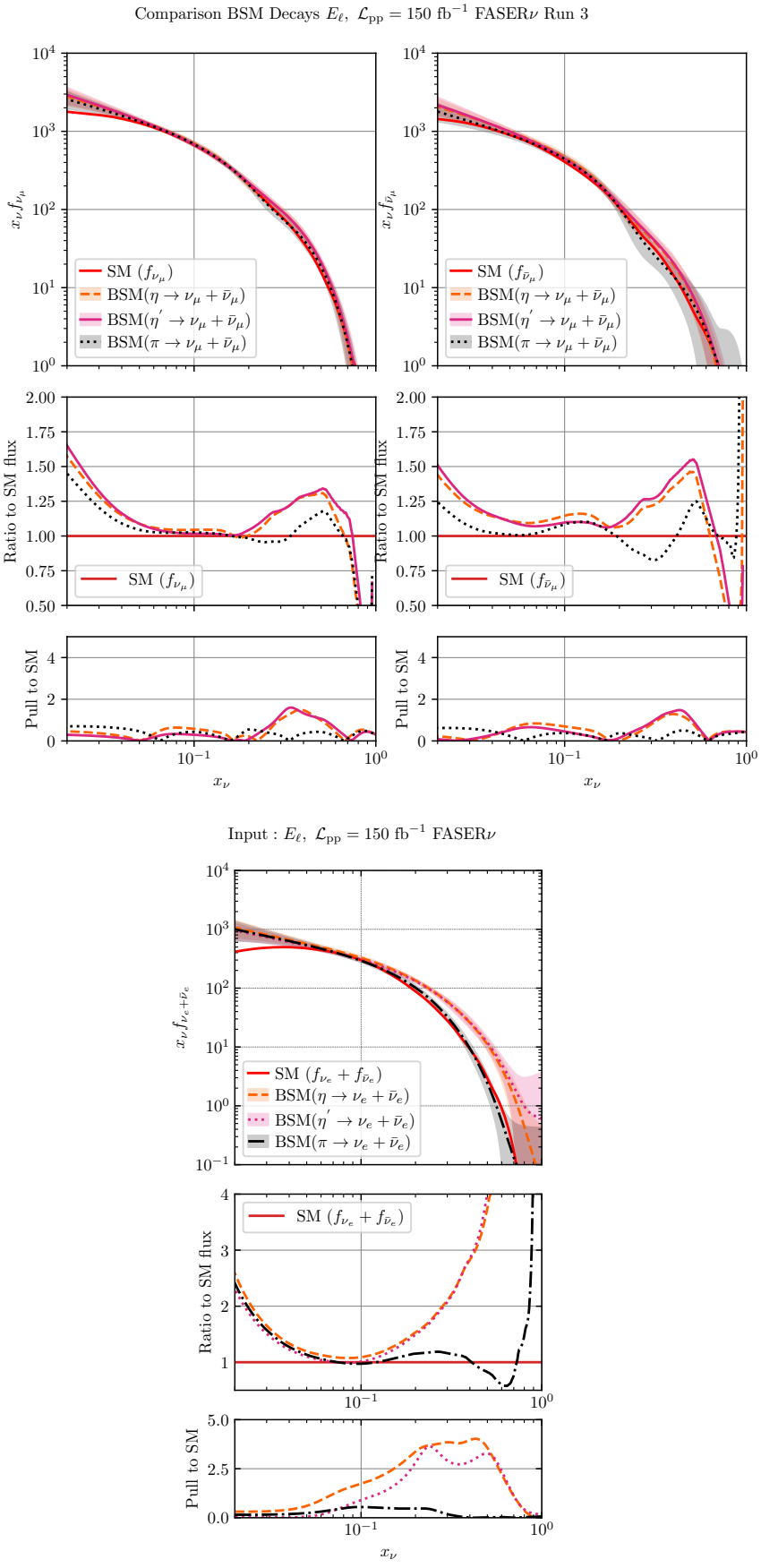


Figure 5.10. These plots show projected neutrino PDFs fitted to event rate measurements which include BSM enhanced decays of neutral mesons into neutrinos. The event rates used for the fits were as a function of the E_ℓ observable. The EPOS flux is the SM prediction and the lower panel shows the comparison between the fitted BSM fluxes and the SM flux. For the muon neutrino flux it can be seen that the pull between the BSM flux and the SM flux does not reach 2. Thus, so no clear excess of muon neutrinos is seen for the enhanced BSM decays. However, the electron neutrino flux is clearly several sigmas higher (around 4 for the η and $\eta\prime$ enhanced decays) than the SM flux in a large x_ν region.

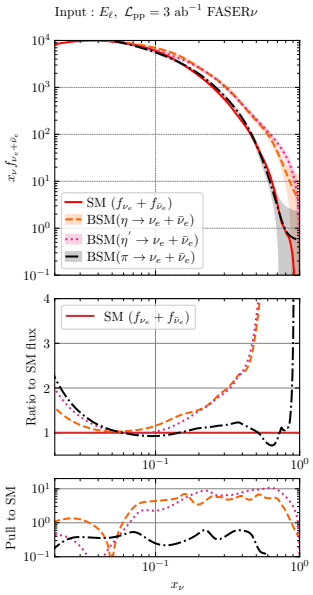


Figure 5.11. Same as Figure 5.10, but for the FASERν detector geometry in the high luminosity era. Only the fitted electron flux is shown, since from Figure 5.10 it can be seen that the enhanced BSM decays have a much smaller effect for the production of muon neutrinos. In the high luminosity era it can be seen that the pull for the η and $\eta\prime$ BSM enhanced decays is around 10 between x_ν= 0.1 and 0.8.

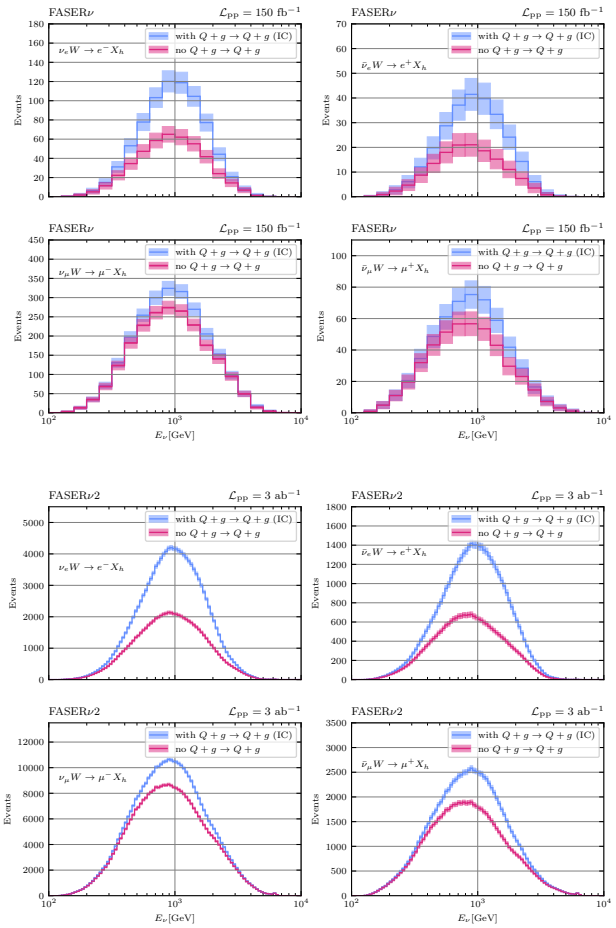


Figure 5.12. These plots show the predicted number of interacting electron and muon neutrinos events at FASERν($\mathcal{L}_{pp}$ = 3 ab^{-1}) and FASERν2 when intrinsic charm is included. The baseline flux prediction is computed using `EPOS+POWHEG` and the contribution from $Q+g \to Q+X$ production is added from [54]. It can be seen that especially for electron neutrinos intrinsic charm doubles the number of events for both detector geometries.

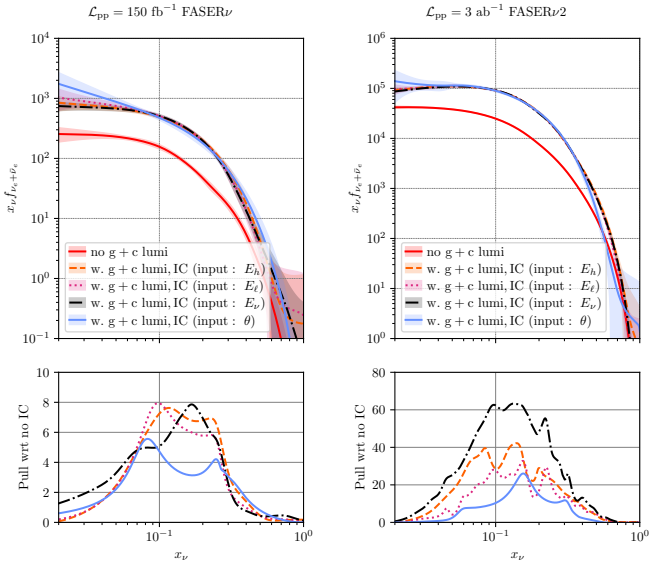


Figure 5.13. This Figure shows two plots showing the results for the electron neutrino IC fits.The upper panel shows the fitted fluxes with and without IC, the bottom panel shows the pull between the fitted IC fluxes and the non-IC fitted `EPOS+POWHEG` flux. It can be seen that the pull with respect to the no-IC flux is around 10σ for FASERν run-3 detector and around 40-60σ for the FASERν detector in the high-lumi era.

where $f^{(\text{no gQ})}_{\nu_e+\bar{\nu}_e}(x_\nu)$ is the fitted flux using the event rates from `EPOS+POWHEG`. In the left plot, the lower panel shows a Pull of around 8 between the fluxes with and without the $c+g$ between 700 and 1500 GeV for all kinematic observables except θ_ℓ, which agrees with the previous observations. The pull projection is optimistic, since only statistical uncertainties are taken into account. As expected, the pull increases to around 40-60 for the FASERν2 detector geometry, since the relative uncertainties in the fitted fluxes decrease. All in all, future data from FASERν at $\mathcal{L}_{pp} = 3$ ab^{-1} should be sensible to intrinsic charm being present in the initial state of the pp.

6 Code availability

6.1 From physicists' code to sustainable code

Physicists write a lot of code to solve their problems numerically. These codes are designed to solve a problem and good code practices are not always taken into account. This is why in this work time has been spent on trying to write the framework in a user-friendly and future proof way. This was also done to make sure the framework is flexible enough that future improvements(which are mentioned below in the outlook) can be implemented relatively easy. After the framework was developed and all the results were computed using this framework, the code was rewritten so that also other people could in principle use and add to the code. The code is dubbed NNfluxν and can be accessed here . On top of this documentation has been written using mkdocs[3] and can be read here. On this website, instructions for installing the code and how to use the code are available. Also, the structure of the framework is explained and a part of the documentation is written using the docstrings of the classes and functions used in the code. In short, users can install the code by cloning it from GitHub and the framework uses poetry to install the required packages. The LHAPDF library needs to be installed separately, since this is not compatible with poetry or pip. After installing the code, users can generate pseudo data, or use event rate measurements from FASER/SND@LHC to parametrise neutrino PDFs. When users want to run a fit, they can specify all the (hyper)parameters in yaml files. The postfit calculations are also included in the framework as well as the hyperparameter optimization algorithm.

6.2 Code used to produce plots

The git repository that was used to produce all the plots can be found on Github. My code can be found in the ML_fit_enu/src/ML_fit_neutrinos directory. There are separate directories for all fits produced in this work and all fits use the same files to for example generate data, perform the fit, plot the result etc. For each fit a separate control file is made with its hyperparameters, because these differ per fit. After a fit is performed, the neutrino PDFs, event rates and the postfit analysis are written to a txt file. A proof that every single closure test produced in this work, also the ones which are not shown in this thesis, is successful can be found here. All plots comparing generators, geometries, observables, enhanced BSM decays and IC vs no IC can also be found here.

[3] https://www.mkdocs.org

7 Summary and Outlook

7.1 Summary

In this work, a framework is developed which is capable of parametrising neutrino PDFs using machine learning from neutrino scattering event rate measurements measured at the LHC. The framework is largely based on the NNPDF framework, which extracts PDFs from measurements on DIS structure functions. Using the framework, for the first time the full E_ν range dependence on the muon neutrino flux is constrained using neutrino event yields from FASER. This neutrino flux is compared to neutrino flux predictions computed by several different event generator, which showed that the neutrino flux extracted from FASER data disfavoured the neutrino flux computed by the `DPMJET 2019` model. The framework is also capable of directly comparing theoretical models for forward hadron production in proton-proton collisions. To facilitate this comparison, several projections of neutrino fluxes were made for a combination of detector geometries (both current and future setups) and forward hadron production models (`EPOS,POWHEG,DPMJET,SIBYLL,QGSJET`), which showed that the high luminosity data from the LHC will constrain the neutrino PDF more accurately and that the `DPMJET` model is incompatible compared to the other models. The framework was also used to research three different BSM enhanced decays of neutral mesons into neutrinos. When including BSM decays, the extracted neutrino flux shows a clear deviation from the SM neutrino flux for muon neutrinos i.e. an increase in the flux is seen when including BSM decays. The compariosn also showed that the FASERν detector, especially in the high luminosity era, is sensitive to this effect. Finally, the effect of intrinsic charm on the neutrino flux was researched by adding gluon-charm scattering to D-meson production. This showed that the FASERν(2) detector geometries is sensitive to intrinsic charm being present in the initial state in proton-proton collisions.

7.2 Future Work and Outlook

7.2.1 Tau neutrinos

In this work, projections for muon and electron neutrino fluxes were made. However, tau neutrinos were not considered, partly due to the lack of statistics, because the neutrino flux of tau neutrinos is much lower compared to electron and muon neutrinos. However, the framework of course can also very easily be extended to fit tau neutrino fluxes. Using this framework to fit tau neutrinos is left for future work.

7.2.2 Rapidity

As explained in the Theoretical Framework, the effect of neutrino rapidity when relating the measurement final state variables to the DIS kinematic variables is assumed to be small for LoS detectors such as FASER. However, for future work rapidity can and should be taken into account which means pseudo rapidity will enter the equations given in Section 2. Including rapidity in the equations and in the framework will make the framework more general since it will also be suitable for data taken by the SND@LHC experiment, since this experiment is off-axis.
The neutrino flux including rapidity is defined as

$$\frac{d^2 N_{\nu_i}(E_\nu, y_\nu)}{dE_\nu dy_\nu}, \tag{7.2.1}$$

where y_ν is the pseudo rapidity defined by equation 2.1.5. The number of neutrino CC scattering events will then be expressed by four variables: $\left(x, Q^2, E_\nu, y_\nu\right)$ and is given by:

$$N_{\rm int}^{(\nu_i)}(x,Q^2,E_\nu,y_\nu) \equiv N_{\rm int}^{(\nu_i)}\left(x_{\rm min}\le x\le x_{\rm max}, Q^2_{\rm min}\le Q^2\le Q^2_{\rm max}, E_\nu^{\rm min}\le E_\nu\le E_\nu^{\rm max}, y_\nu^{\rm min}\le y_\nu\le y_\nu^{\rm max}\right). \tag{7.2.2}$$

The relation between the number of events and the neutrino flux is the given by:

$$N_{\rm int}^{(\nu_i)}(x,Q^2,E_\nu) = \int_{Q^2_{\rm min}}^{Q^2_{\rm max}} dQ^2 \int_{x_{\rm min}}^{x_{\rm max}} dx \int_{E_\nu^{\rm (min)}}^{E_\nu^{\rm (max)}} dE_\nu \int_{y_\nu^{\rm (min)}}^{y_\nu^{\rm (max)}} dy_\nu\, \widetilde{N}_{\rm int}^{(\nu_i)}(x,Q^2,E_\nu,y_\nu)\,, \tag{7.2.3}$$

where the integrand is defined as

$$\widetilde{N}_{\rm int}^{(\nu_i)}(x,Q^2,E_\nu,y_\nu) \equiv n_T L_T \times \frac{d^2N_{\nu_i}(E_\nu,y_\nu)}{dE_\nu dy_\nu} \times \frac{d^2\sigma^{\nu_i A}(x,Q^2,E_\nu)}{dxdQ^2} \times \mathcal{A}(E_\ell,\theta_\ell,E_h)\,, \tag{7.2.4}$$

These equations apply for detectors which have access to rapidity such as FASERν2. The FK-table formalism should then also include rapidity dependence, since the flux is both dependent on rapidity and neutrino energy. The 2D neutrino PDF is defined as

$$f_{\nu_i}(x_\nu,y_\nu) \equiv \frac{\sqrt{s_{\rm pp}}}{2}\frac{dN_{\nu_i}(E_\nu,y_\nu)}{dE_\nu dy_\nu}, \qquad i=e,\mu,\tau, \tag{7.2.5}$$

and in terms of the interpolating functions it is given by:

$$f_{\nu_i}(x_\nu,y_\nu) \simeq \sum_{\alpha=1}^{n_x}\sum_{\beta=1}^{n_y} f_{\nu_i}(x_{\nu,\alpha},y_{\nu,\beta}) I_\alpha(x_\nu) I_\beta(y_\nu), \tag{7.2.6}$$

where I are the interpolation polynomials, and x_ν and y_ν are the grid nodes. Now the event rates are given by:

$$\begin{aligned}
N_{\rm int}^{(\nu_i)}(E_{\nu,j},y_{\nu,k}) &= \int_{E_{\nu,j}^{\rm (min)}}^{E_{\nu,j}^{\rm (max)}} dE_\nu \int_{y_{\nu,k}^{\rm (min)}}^{y_{\nu,k}^{\rm (max)}} dy_\nu \int_{Q_0^2=1}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^1 dx \\
&\quad\times\left(n_T L_T \times \frac{2}{\sqrt{s_{\rm pp}}} f_{\nu_i}(x_\nu,y_\nu) \times \left.\frac{d^2\sigma^{\nu_i A}_{\rm NLO+PS}(x,Q^2,E_\nu)}{dx\,dQ^2}\right|_{\rm fid}\right) \\
&= \sum_{\alpha=1}^{n_x}\sum_{\beta=1}^{n_y} f_{\nu_i}(x_{\nu,\alpha},y_{\nu,\beta}) \int_{E_{\nu,j}^{\rm (min)}}^{E_{\nu,j}^{\rm (max)}} dE_\nu \int_{y_{\nu,k}^{\rm (min)}}^{y_{\nu,k}^{\rm (max)}} dy_\nu \int_{Q_0^2=1}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^1 dx \\
&\quad\times\left(n_T L_T \times \frac{2}{\sqrt{s_{\rm pp}}} I_\alpha(x_\nu) I_\beta(y_\nu) \times \left.\frac{d^2\sigma^{\nu_i A}_{\rm NLO+PS}(x,Q^2,E_\nu)}{dx\,dQ^2}\right|_{\rm fid}\right) \\
&= \sum_{\alpha=1}^{n_x}\sum_{\beta=1}^{n_y} f_{\nu_i}(x_{\nu,\alpha},y_{\nu,\beta})\,{\rm FK}_{\alpha,\beta,j,k}
\end{aligned} \tag{7.2.7}$$

The FK-table has become a rank-4 tensor, where two indices label the bins in E_ν and y_ν and the other two indices are labelled by the grid nodes. The expression for the FK-table is given by:

$${\rm FK}_{\alpha,\beta,j,k} = \frac{2n_T L_T}{\sqrt{s_{\rm pp}}} \int_{E_{\nu,j}^{\rm (min)}}^{E_{\nu,j}^{\rm (max)}} dE_\nu \int_{y_{\nu,k}^{\rm (min)}}^{y_{\nu,k}^{\rm (max)}} dy_\nu \int_{Q_0^2=1}^{2m_N E_\nu} dQ^2 \int_{Q^2/2m_N E_\nu}^1 dx\, I_\alpha(x_\nu) I_\beta(y_\nu) \left.\frac{d^2\sigma^{\rm NLO+PS}_{\nu_i A}(x,Q^2,E_\nu)}{dx\,dQ^2}\right|_{\rm fid}. \tag{7.2.8}$$

.

Besides including rapidity in the theoretical framework, it should also be included in the numerical

framework. Again, when generating pseudo data, the FK-table convoluted with the flux should recover the event rates within a few percent. The rapidity dependence can also be implemented in the LHAPDF framework and `POWHEG` is also capable, when including some minor changes, of producing event rates as a function of both neutrino energy and rapidity. For the ML fits, the preprocessing function will not change, because there are no asymptotic behaviours considering the dependence of the neutrino PDF on rapidity. Considering the structure of the NN, there would be two instead of one input layers, because there are two vectors of grid nodes. The output of the NN, the neutrino PDF, would be a 2D object and the number of events in a certain energy and rapidity bin, would be a matrix instead of a vector. Unfortunately due to lack of time, rapidity is not yet implemented in the framework. However, in the future it will be implemented by means of a future release.

7.2.3 ML algorithms and the NNPDF framework

The framework presented in this work is based on NNPDF and can produce ML parametrisation of neutrino PDFs from event rate measurements. It has adopted most features from the NNPDF framework, however there are still features/algorithms/implementations that are not yet adopted by our framework, such as:

- First of all, NNPDF fits all separate PDFs together in one MLP. In this work, separate networks are used for different flavours. This is because currently there are no event rate measurements for electron neutrinos from FASER or SND@LHC. However, in the future, when this data will be available, the framework should be adjusted in a way so that all neutrino flavours are fitted at once.

- PineAPPL[25]. In NNPDF, the library PineAPPL is used to generate fast-interpolation grids and to convolute the grids with PDFs and the advantage of PineAPPL is its speed of the operations, since it is written in Rust. Our framework does not yet have an interface with PineAPPL, however this will be the case in a future release.

- Bayesian reweighing[55]. In NNPDF there is a way to incorporate new data from the LHC to existing PDF sets using Bayesian Reweighing. One starts off by having N equally likely PDF sets, before adding the new data. Then, by using Bayes' Theorem, weights are computed and using these weights the probability distribution is updated to N reweighed PDF sets. This is not incorporated into our framework, partly due to the lack of data. However, this method is an elegant way to incorporate new data into neutrino PDFs without having to redo the NN fits.

- In the future, when there is more data available from FASER and SND@LHC, it would be possible to combine the NNPDF framework and this framework to fit PDFs and neutrino PDFs simultaneously. This would be beneficial, because in this way the NN can learn that different observables, detectors, and kinematic regions put constraints on both PDFs and neutrino PDFs simultaneously.

7.2.4 Plethora of Physics Applications

In this work several applications to extracting neutrino PDFs from data have been researched. However, one can think of many more physics applications for this framework such as:

- With the current neutrino PDF extracted from FASER event rate measurements, one can constrain the K-π-swap model of strangeness enhancement. Strangeness enhancement is the phenomenon in which strange particles like kaons and Ω baryons are observed more abundantly in Pb-Pb collisions than in pp collisions[56]. The K-π-swap model is a model which tries to mimic this behaviour by creating strangeness enhancement by hadronic rescattering processes. This model would then also predict a number of neutrinos, which can be compared to the neutrino flux parametrisation from the FASER data.

- In this work the effect of intrinsic charm on the neutrino production is also investigated. There are many other possible physics effects that can be researched, for example gluon saturation. Gluon saturation suggests that the gluon density is no longer growing freely at small Bjorken-x, which can also have an effect on the number of neutrinos produced from pp collisions. This is because the forward particle production is sensitive to the small Bjorken-x region of the target proton.

- Another application that this framework is suited for is to constrain low-x PDFs for charm quark production. NNPDF has multiple different PDF sets which include different physics and data, such as LHCb data and PDFs with and without resummation. This framework can compare several PDFs, because different PDFs yield a different number of scattering events per neutrino energy.

- DGLAP vs BFKL. DGLAP and BFKL are both equations describing how PDFs evolve with energy. DGLAP is mainly used for high-x and to describe how PDFs evolve with large Q^2, while BFKL is used in the low-x region and describe how PDFs evolve with small-x[57]. Again, these two equations will yield different different forward hadron production and will thus yield a different number of scattering events per neutrino energy. The framework described in this work, will be able to differentiate the two approaches to evolving PDFs when comparing neutrino PDFs.

- In the future, when rapidity dependence will be implemented in this framework and when the neutrino PDF parametrisation including rapidity is also available, the prompt neutrino fluxes can be predicted for the KM3NET and IceCube neutrino telescopes. The advantage of these predictions would be their accuracy, since the PDFs display large uncertainties, which also work through in the production of charm in the forward region and thus in the prediction of these prompt neutrino fluxes. In short, FASER event rate measurements can, using the framework presented in this work, actively be used to accurately predict prompt neutrino fluxes for neutrino telescopes.

7.2.5 Future Data

The numerical framework in this work is build to parametrise neutrino fluxes using event rate measurements from detectors at the LHC. The FASER collaboration has measured event rates for muon neutrino scattering as a function of neutrino energy [3] and as showed above, this work has, for the first time, constrained the full E_ν dependence on the muon neutrino flux. Currently, and also in the high luminosity area, FASER will continue to take data. The numerical framework in this work will be ready to use this data to extract more precise neutrino fluxes for more flavours as has been showed using pseudo data in Section 5. If the proposed FPF will be build and fully realised, this means even

more data and thus even more accurate neutrino flux parametrisations. The fluxes extracted from event rate measurements can then also be compared to several event generators, as is done in this work, to further constrain and perhaps fully exclude models. In conclusion, future data from the LHC promises an exciting and important role for this framework in unravelling QCD and the structure of the proton.

8 Acknowledgements

This thesis would not have been possible without my supervisor Prof. Dr. Juan Rojo. I thank him for making this entire project possible and his guidance throughout the project. In particular, he has showed me what is needed to conduct high-quality research and what is needed to publish a paper by his sharp comments on my work. Also, I would like to thank my (daily) supervisor Peter Krack for generating the pseudo data and FK-tables as well as some of the plots provided in this work. He has truly played an integral part in the progress of this work. Felix Kling was also an integral part of this project. His expertise of FASER and his knowledge on high-energy physics in general has helped my understanding of the FASER experiments and collider physics a lot. Furthermore, I am grateful to Jelle Koorn for being my partner and friend throughout this project. Having discussions with him were very valuable for my understanding and the progress of this work. Finally, I want to thank the entire Nikhef theory group for being very warm and friendly towards not only me but all the master students. You have showed me that doing research in physics can actually be combined with a sense of belonging, community and kindness in general.

References

[1] A. Unknown, "Looking for neutrinos: Nature's ghost particles," *Smithsonian Magazine* (2009) . https://www.smithsonianmag.com/science-nature/looking-for-neutrinos-natures-ghost-particles-64200742/. Accessed: 2024-12-11.

[2] H. e. a. Abreu, "First direct observation of collider neutrinos with faser at the lhc," *Physical Review Letters* **131** no. 3, (July, 2023) . http://dx.doi.org/10.1103/PhysRevLett.131.031801.

[3] **FASER** Collaboration, R. Mammen Abraham *et al.*, "First Measurement of the Muon Neutrino Interaction Cross Section and Flux as a Function of Energy at the LHC with FASER," arXiv:2412.03186 [hep-ex].

[4] R. e. a. Albanese, "Observation of collider muon neutrinos with the snd@lhc experiment," *Physical Review Letters* **131** no. 3, (July, 2023) . http://dx.doi.org/10.1103/PhysRevLett.131.031802.

[5] J. L. e. a. Feng, "The forward physics facility at the high-luminosity lhc," *Journal of Physics G: Nuclear and Particle Physics* **50** no. 3, (Jan., 2023) 030501. http://dx.doi.org/10.1088/1361-6471/ac865e.

[6] J. L. Feng, F. Kling, M. H. Reno, J. Rojo, D. Soldin, L. A. Anchordoqui, J. Boyd, A. Ismail, L. Harland-Lang, K. J. Kelly, *et al.*, "The forward physics facility at the high-luminosity lhc," *Journal of Physics G: Nuclear and Particle Physics* **50** no. 3, (2023) 030501.

[7] M. e. a. Aartsen, "Observation of high-energy astrophysical neutrinos in three years of icecube data," *Physical Review Letters* **113** no. 10, (Sept., 2014) . http://dx.doi.org/10.1103/PhysRevLett.113.101101.

[8] e. a. Adrián-Martínez, S, "Letter of intent for km3net 2.0," *Journal of Physics G: Nuclear and Particle Physics* **43** no. 8, (June, 2016) 084001. http://dx.doi.org/10.1088/0954-3899/43/8/084001.

[9] J. Adhikary *et al.*, "Science and Project Planning for the Forward Physics Facility in Preparation for the 2024-2026 European Particle Physics Strategy Update," arXiv:2411.04175 [hep-ex].

[10] J. Gao, L. Harland-Lang, and J. Rojo, "The structure of the proton in the lhc precision era," *Physics Reports* **742** (May, 2018) 1–121. http://dx.doi.org/10.1016/j.physrep.2018.03.002.

[11] J. J. Ethier and E. R. Nocera, "Parton distributions in nucleons and nuclei," *Annual Review of Nuclear and Particle Science* **70** no. 1, (Oct., 2020) 43–76. http://dx.doi.org/10.1146/annurev-nucl-011720-042725.

[12] R. Placakyte, "Parton distribution functions," *arXiv preprint arXiv:1111.5452* (2011) .

[13] R. D. Ball, S. Carrazza, L. Del Debbio, S. Forte, Z. Kassabov, J. Rojo, E. Slade, and M. Ubiali, "Nnpdf," *Parton distributions from high-precision collider data," Eur. Phys. J. C* **77** (2014) .

[14] A. Candido, F. Hekhorn, and G. Magni, "Eko: evolution kernel operators," *The European Physical Journal C* **82** no. 10, (Oct., 2022) . http://dx.doi.org/10.1140/epjc/s10052-022-10878-w.

[15] T. N. Collaboration, "Nnpdf documentation." https://docs.nnpdf.science/. Accessed: 2025-02-11.

[16] A. Candido, A. Garcia, G. Magni, T. Rabemananjara, J. Rojo, and R. Stegeman, "Neutrino structure functions from gev to eev energies," *Journal of High Energy Physics* **2023** no. 5, (May, 2023) . http://dx.doi.org/10.1007/JHEP05(2023)149.

[17] P. D. G. Collaboration, S. Bethke, G. Dissertori, and et al., "Review of particle physics," *Progress of Theoretical and Experimental Physics* (2023) . https://pdg.lbl.gov/2023/reviews/rpp2023-rev-qcd.pdf.

[18] F. Kling and L. J. Nevay, "Forward neutrino fluxes at the lhc," *Physical Review D* **104** no. 11, (2021) 113008.

[19] F. Collaboration, A. Ariga, T. Ariga, J. Boyd, F. Cadoux, D. W. Casper, F. Cerutti, S. Danzeca, L. Dougherty, Y. Favre, *et al.*, "Technical proposal for faser: Forward search experiment at the lhc,".

[20] M. E. Peskin and F. Halzen, "An introduction to quantum field theory, by michael e. peskin and daniel v. schroeder," 1995.

[21] M. van Beekveld, S. Ferrario Ravasio, E. Groenendijk, P. Krack, J. Rojo, and V. S. Sánchez, "A phenomenological analysis of LHC neutrino scattering at NLO accuracy matched to parton showers," *Eur. Phys. J. C* **84** no. 11, (2024) 1175, arXiv:2407.09611 [hep-ph].

[22] CERN, "GHM Lecture 2," 2015. https://indico.cern.ch/event/388234/attachments/1129101/1613186/GHM_Lecture_2.pdf. Accessed: 2025-02-12.

[23] E. Daw, "Lecture 7-rapidity and pseudorapidity," *Lecture Notes* (2012) 1–8.

[24] T. KLUGE, K. RABBERTZ, and M. WOBISCH, "Fast pqcd calculations for pdf fits," in *Deep Inelastic Scattering DIS 2006*. WORLD SCIENTIFIC, Jan., 2007. http://dx.doi.org/10.1142/9789812706706_0110.

[25] S. Carrazza, E. R. Nocera, C. Schwan, and M. Zaro, "Pineappl: combining ew and qcd corrections for fast evaluation of lhc processes," *Journal of High Energy Physics* **2020** no. 12, (Dec., 2020) . http://dx.doi.org/10.1007/JHEP12(2020)108.

[26] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali, "A first unbiased global nlo determination of parton distributions and their uncertainties," *Nuclear Physics B* **838** no. 1–2, (Oct., 2010) 136–206. http://dx.doi.org/10.1016/j.nuclphysb.2010.05.008.

[27] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, "Lhapdf6: parton density access in the lhc precision era," *The European Physical Journal C* **75** no. 3, (Mar., 2015) . http://dx.doi.org/10.1140/epjc/s10052-015-3318-8.

[28] R. D. Ball, S. Carrazza, J. Cruz-Martinez, L. Del Debbio, S. Forte, T. Giani, S. Iranipour, Z. Kassabov, J. I. Latorre, E. R. Nocera, *et al.*, "The path to proton structure at 1% accuracy: Nnpdf collaboration," *The European Physical Journal C* **82** no. 5, (2022) 428.

[29] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems* **8** no. 7, (2009) 579–588.

[30] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, *et al.*, "A comprehensive guide to the physics and usage of pythia 8.3," *SciPost Physics Codebases* (2022) 008.

[31] F. Kling and L. J. Nevay, "Forward neutrino fluxes at the lhc," *Physical Review D* **104** no. 11, (Dec., 2021) . http://dx.doi.org/10.1103/PhysRevD.104.113008.

[32] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP* **11** (2004) 040, arXiv:hep-ph/0409146.

[33] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP* **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[34] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP* **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[35] T. Ježo and P. Nason, "On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower," *JHEP* **12** (2015) 065, arXiv:1509.09071 [hep-ph].

[36] E. Groenendijk, "Predictions for high-energy neutrino scattering at the large hadron collider," Master's thesis, Vrije Universiteit Amsterdam & Nikhef, 2023. Supervisor: Prof. Dr. Juan Rojo, Second examiner: Dr. Lydia Brenner.

[37] K. Rabbertz, *Jet Physics at the LHC*. Springer, 2016.

[38] M. H. Seymour and M. Marx, "Monte carlo event generators," in *LHC Phenomenology*, pp. 287–319. Springer, 2014.

[39] V. C. S. S´anchez, "Nlo predictions for neutrino deep inelastic scattering at the future forward physics facility," Master's thesis, Vrije Universiteit Amsterdam & Nikhef, 2023. Supervisor: Prof. Dr. Juan Rojo, Second examiner: Dr. Lydia Brenner.

[40] H. Abreu, E. A. Mansour, C. Antel, A. Ariga, T. Ariga, F. Bernlochner, T. Boeckh, J. Boyd, L. Brenner, F. Cadoux, *et al.*, "The faser detector," *Journal of Instrumentation* **19** no. 05, (2024) P05066.

[41] R. M. Abraham, X. Ai, J. Anders, C. Antel, A. Ariga, T. Ariga, J. Atkinson, F. U. Bernlochner, T. Boeckh, J. Boyd, *et al.*, "First measurement of the muon neutrino interaction cross section and flux as a function of energy at the lhc with faser," *arXiv preprint arXiv:2412.03186* (2024) .

[42] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, *et al.*, "Parton distributions for the lhc run ii," *Journal of High Energy Physics* **2015** no. 4, (2015) 1–148.

[43] P. I. Frazier, "A tutorial on bayesian optimization," 2018. https://arxiv.org/abs/1807.02811.

[44] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "Epos lhc: Test of collective hadronization with data measured at the cern large hadron collider," *Physical Review C* **92** no. 3, (Sept., 2015) . http://dx.doi.org/10.1103/PhysRevC.92.034906.

[45] S. Ostapchenko, "Monte carlo treatment of hadronic interactions in enhanced pomeron scheme: Qgsjet-ii model," *Physical Review D* **83** no. 1, (Jan., 2011) . http://dx.doi.org/10.1103/PhysRevD.83.014018.

[46] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, "Hadronic interaction model sibyll 2.3d and extensive air showers," *Physical Review D* **102** no. 6, (Sept., 2020) . http://dx.doi.org/10.1103/PhysRevD.102.063002.

[47] S. Roesler, R. Engel, and J. Ranft, *The Monte Carlo Event Generator DPMJET-III*, p. 1033–1038. Springer Berlin Heidelberg, 2001. http://dx.doi.org/10.1007/978-3-642-18211-2_166.

[48] M. Fieg, F. Kling, H. Schulz, and T. Sjöstrand, "Tuning pythia for forward physics experiments," 2023. https://arxiv.org/abs/2309.08604.

[49] **Particle Data Group** Collaboration, S. Navas *et al.*, "Review of particle physics," *Phys. Rev. D* **110** no. 3, (2024) 030001.

[50] R. D. Ball, A. Candido, J. Cruz-Martinez, S. Forte, T. Giani, F. Hekhorn, K. Kudashkin, G. Magni, and J. Rojo, "Evidence for intrinsic charm quarks in the proton," *Nature* **608** no. 7923, (Aug., 2022) 483–487. http://dx.doi.org/10.1038/s41586-022-04998-2.

[51] T. Hobbs, J. Londergan, and W. Melnitchouk, "Phenomenology of nonperturbative charm in the nucleon," *Physical Review D* **89** no. 7, (Apr., 2014) . http://dx.doi.org/10.1103/PhysRevD.89.074008.

[52] R. e. a. Aaij, "Study of bosons produced in association with charm in the forward region," *Physical Review Letters* **128** no. 8, (Feb., 2022) . http://dx.doi.org/10.1103/PhysRevLett.128.082001.

[53] R. Maciuła and A. Szczurek, "Far-forward production of charm mesons and neutrinos at forward physics facilities at the lhc and the intrinsic charm in the proton," *Physical Review D* **107** no. 3, (Feb., 2023) . http://dx.doi.org/10.1103/PhysRevD.107.034002.

[54] L. Buonocore, F. Kling, L. Rottoli, and J. Sominka, "Predictions for neutrinos and new physics from forward heavy hadron production at the lhc," *The European Physical Journal C* **84** no. 4, (Apr., 2024) . http://dx.doi.org/10.1140/epjc/s10052-024-12726-5.

[55] Y. Wang, A. Kucukelbir, and D. M. Blei, "Robust probabilistic modeling with bayesian data reweighting," 2018. https://arxiv.org/abs/1606.03860.

[56] M. Hanafy, O. S. A. Qandil, and A. G. Shalaby, "Strangeness enhancement at lhc energies using the thermal model and eposlhc event-generator," 2021. https://arxiv.org/abs/2105.14303.

[57] A. Kotikov and L. Lipatov, "Dglap and bfkl equations in the n=4 supersymmetric gauge theory," *Nuclear Physics B* **661** no. 1–2, (June, 2003) 19–61. http://dx.doi.org/10.1016/S0550-3213(03)00264-5.