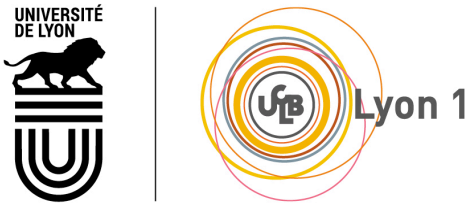
Kinematic search for tau neutrino appearance in the DUNE experiment

Thomas Kosc



N° d'ordre NNT : 2021LYSE1170

THESE DE DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N°52
PHAST : Physique et Astrophysique de Lyon

Spécialité de doctorat : Physique des Particules

Soutenue publiquement le 06/09/2021, par:
Thomas Kosc

Kinematic search for τ neutrino appearance in the DUNE experiment

Devant le jury composé de :

JOLLET	Cécile	Maître de Conférences - CENBG Bordeaux	Rapporteure
MARCHIONNI	Alberto	Senior Scientist - Fermilab (USA, Illinois)	Rapporteur
DAVESNE	Dany	Professeur des Universités - IP2I Lyon	Examinateur
NEDELEC	Patrick	Professeur des Universités - IP2I Lyon	Examinateur
AUTIERO	Dario	Directeur de Recherche CNRS - IP2I Lyon	Directeur de Thèse
CHEYNIS	Brigitte	Chargée de Recherche CNRS - IP2I Lyon	Invitée

Remerciements

Quand je regarde ce manuscrit rédigé, je pense à toutes les circonstances qui m'ont amenées là, à celles qui ont rendu ce travail de thèse possible, tout simplement. Je pense d'abord à toi, Dario, pour ces années d'apprentissage passées à tes côtés. Tu m'as souvent répété que tu m'enseignais l'autonomie, je réalise combien j'ai gagné sur ce point. Ta rigueur et ta culture ont façonné ce manuscrit. Aux personnes qui ont accepté de faire partie du jury de cette thèse, j'adresse mes profonds remerciements pour le temps qu'elles ont dédié. Aux personnes qui m'ont aidé pour la rédaction, Dario, Elisabetta, Slavic, Camille, Clément, merci du fond du coeur pour votre aide. Mes maladresses linguistiques n'avaient qu'à bien se tenir !

Mais ça va plus loin, j'ai pu bénéficier de ces années de thèse parce que les circonstances de ma vie l'ont permis. Mes parents qui m'ont toujours laissé la liberté d'aller étudier ce que je voulais et surtout soutenu dans chaque étape, avec calme et confiance. C'est un immense privilège. J'ai rencontré en chemin des enseignant.es merveilleux.ses qui ont su, je ne sais comment, éveiller en moi une flamme pour les sciences et me transmettre ce dont j'aurais besoin pour poursuivre l'aventure : connaissances, rigueur, curiosité... J'ai aussi une pensée pour les ami.es que j'ai rencontré.es sur la route, pour certain.es nous nous sommes perdus de vue, mais j'ai en moi une profonde gratitude pour la présence que vous avez pu être. Membres de Bërk et troupe de la Doua avec qui j'ai partagé maints repas (on a parfois déconné sur les sujets de conversation...) : j'ai beaucoup apprécié votre présence. Merci à vous.

Je connais la petite frustration qui apparaît lorsqu'on lit les remerciements d'une thèse et qu'on ne voit pas son nom apparaître. Ha ! C'est tout simplement impossible de citer chaque personne... Mais peu importe, chaque moment partagé a compté, chaque sourire, chaque échange... Une gratitude profonde apparaît en moi lorsque je pense, ou plutôt que j'essaye de penser à toutes celles et ceux que j'ai rencontré.es et dont la présence m'a été bénéfique. Cette thèse est le résultat d'un travail pour lequel j'ai été un des acteurs principaux, mais vous avez sûrement fait partie des acteurs secondaires (voire plus). Voyez qu'alors je me sens humble et subtilement relié à chacun.e d'entre vous. C'est un sentiment étrange et vertigineux qui donne les larmes aux yeux, et qui me fait réaliser combien j'ai été chanceux d'avoir eu la vie que j'ai eue jusqu'à ce jour. Ce sentiment procure une paix immense, c'est la magie de la gratitude.

Je ne finirai pas ces remerciements sans te mentionner, Clémentine, toi qui partages ma vie depuis le début de la thèse. Tu es une leçon de bienveillance, d'ouverture, de légèreté, de pardon, d'amour. J'admire beaucoup la personne que tu es, combien tu es spontanée et douce dans tes relations avec les autres. J'apprends beaucoup à tes côtés.

Je finirai par un des mantras les plus puissants et merveilleux qui existent : merci.

Résumé

Il m'a parut opportun de profiter de ce résumé français à faible contrainte pour présenter la thématique de cette thèse de façon générale et accessible à des non-spécialistes. J'y incluerai en plus un résumé technique en deuxième partie.

Le premier point à simplement aborder est la question (que l'on m'a si souvent posée...) : qu'est qu'un neutrino ? Voyez d'abord comment les physicien.nes pensent le monde. La matière est décrite comme composée de particules élémentaires (les fermions) interagissant via l'échange de bosons (par exemple les photons, les quantas de lumière). Le tout est structuré dans une théorie appelée le Modèle Standard de la physique des particules, une théorie de jauge locale qui rend compte des trois interactions fondamentales : électromagnétique, nucléaires forte et faible. En somme, elle fournit un cadre mathématique pour cataloguer les particules élémentaires et rendre compte de leurs interactions.

Le neutrino apparaît dans ce modèle comme une particule de matière électriquement neutre sensible seulement à l'interaction faible. Par conséquent, elle interagit relativement peu avec la matière dite "ordinaire", celle qui constitue les atomes. Ainsi c'est une forme de matière parfois qualifiée de subtile ou fantômatique pour la simple raison qu'elle interagit très peu avec "nous". D'une certaine manière, elle ne nous "voit" que très peu.

La question de la masse des neutrinos, que l'on m'a aussi souvent posée, est épineuse. La notion de masse nous est si familière qu'il est compliqué de penser que la masse des neutrinos n'est pas directement accessible par l'expérience. Des neutrinos massifs sont requis dans les modèles pour rendre compte d'un phénomène singulier qu'on appelle les oscillations de neutrinos, mais les masses en tant que telles ne sont pas directement mesurables. Il existe tout de même quelques lots de consolations : il est possible d'établir des contraintes sur la somme des trois masses de neutrinos par des arguments cosmologiques (faisant notamment intervenir la matière noire), les expériences sur les oscillations de neutrinos sont sensibles à la différence des carrés des masses, et des expériences de mesures précises de la queue de distribution de désintégration beta permettent de contraindre la masse effective du neutrino électronique.

Je parle de "masse effective" du neutrino électronique car, formellement, il n'a pas de masse définie. Quelques explications s'imposent. Il existe trois types de neutrinos liés à l'intéraction faible, que l'on appelle les états de saveur : électronique, muonique et tauique. Chaque saveur de neutrino est associée, par définition, au lepton chargé correspondant du Modèle Standard de la physique des particules : électron, muon ou tau. Un neutrino électronique est un neutrino produisant un électron lors d'une interaction faible (par courant chargé, mais je ne vais pas rentrer dans les détails ici). Seulement, il se trouve que la propriété de saveur n'est pas conservée lors de la propagation des neutrinos, alors qu'elle l'est pour les leptons chargés (un électron demeure un électron lorsqu'il se propage). Un neutrino électronique émit en un point de l'espace (coeur du Soleil, par exemple), peut très bien être mesuré sur Terre comme un neutrino doté

de la saveur muonique. Ce changement de saveur des neutrinos lors de la propagation s'appelle les oscillations des neutrinos. Il a été définitivement mis en évidence en 1998 par l'expérience Super-Kamiokande et en 2001 par l'expérience SNO.

Ce phénomène nécessite d'ajouter au Modèle Standard une façon de rendre compte de la masse des neutrinos, un problème à part entière dans lequel je n'entrerai pas. En admettant que l'on dispose d'un tel modèle, on peut alors définir des états propres de masse comme combinaison des états de saveur. Chaque neutrino massif est un mélange des trois états de saveur, on les appelle neutrino 1, 2 et 3. Ces neutrinos là ont une masse, mais n'ont pas de saveur bien définie. Inversement les neutrinos de saveur n'ont pas de masse bien définie. Or expérimentalement, on mesure les neutrinos lorsqu'ils intéragissent via l'interaction faible, ce sont donc les états de saveur que nous mesurons, pas les états de masse.

Depuis la découverte des oscillations des neutrinos, un large programme scientifique s'est déployé pour mesurer précisement ce phénomène. Mon travail de thèse s'inclue dans ce contexte avec la future expérience DUNE (Deep Underground Neutrino Experiment). Je me suis intéressé en particulier à la détection de la saveur tauique, largement non-triviale comparée aux deux autres. Le résumé ci-dessous est une traduction du résumé anglais qui se trouve plus loin dans le manuscrit.

DUNE (Deep Underground Neutrino Experiment) est une future expérience sur faisceau de neutrinos sur longue distance qui aura pour objectif, entre autres, d'approfondir l'étude des oscillations des neutrinos. Elle s'appuiera sur la création d'un faisceau de neutrinos muoniques au Fermilab (Illinois), caractérisé avec un détecteur proche situé lui aussi au Fermilab, et utilisé pour les mesures d'oscillation avec des détecteurs lointains situés 1300 km plus loin au Sanford Underground Research Facility dans le Dakota du Sud. Cette installation souterraine comprendra quatre gigantesques détecteurs de neutrinos de 10 kilotonnes chacun (masse de détection) employant la technologie des chambres à projection temporelle à argon liquide, dotée d'une grande résolution spatiale et d'une grande performance calorimétrique.

Les oscillations des neutrinos sont un phénomène spécial qui leur permet de changer de saveur. Il a été découvert en 1998 et a depuis fait l'objet d'une activité de recherche expérimentale intensive. Trois ans après cette découverte, le neutrino tau était observé directement pour la première fois. C'est aujourd'hui la particule du Modèle Standard la moins bien étudiée avec seulement une vingtaine de candidats directement observés. DUNE sera d'une sensibilité sans précédent concernant l'apparition de neutrinos tau grâce au large canal d'oscillations des neutrinos muoniques en neutrinos tauiques. Un taux d'événements de cette saveur d'environ 30 neutrinos par an et par détecteur lointain y est attendu. Les étudier permettra de considérablement améliorer la compréhension des interactions de neutrinos avec la matière baryonique et offrira un test pour le paradigme à trois saveurs de neutrino.

L'identification de la saveur d'un neutrino s'obtient par la détection du lepton chargé de la même famille créé au vertex d'interaction. Or, la désintégration rapide du lepton chargé tau rend l'identification de cette saveur de neutrino bien plus ardue que pour les neutrinos électroniques et muoniques pour qui l'identification directe de l'électron et du muon est possible. De plus, la grande masse du lepton chargé tau (1.78 GeV), comparable à l'énergie des neutrinos de faisceau de DUNE (3 GeV), rend la section efficace en courant chargé des neutrinos tau défavorable d'un point de vue cinétique. L'identification d'un échantillon de neutrinos tau pour DUNE requiert donc une méthode non-triviale.

Cette thèse présente une analyse de recherche de neutrinos tau sur des événements simulés pour l'expérience DUNE en s'appuyant sur des critères cinématiques. La reconstruction précise

des traces dans les chambres à projection temporelle à argon liquide du point de vue de la résolution spatiale et de la résolution en énergie rend l'expérience particulièrement adaptée pour la recherche de désintégrations de leptons chargés tau par des critères cinématiques. Trois canaux de désintégration sont étudiés avec pour chacun une analyse dédiée dans le but d'isoler des neutrinos tau et de rejeter les bruits de fond associés.

Le premier canal étudié est celui par lequel le lepton tau se désintègre en électron et deux neutrinos (chapitre 3). La topologie des événements associés contient donc un électron dans l'état final, ce qui les rend similaires à des interactions de neutrinos électroniques par courant chargé. En se plaçant dans le plan transverse de l'interaction où la cinématique est connue, il est cependant possible de distinguer les deux types d'interactions. En effet, les interactions liées aux neutrinos tau apportent deux neutrinos dans l'état final qui ne seront pas reconstruits par les détecteurs. La conservation de l'impulsion dans ce plan transverse révèle alors une énergie manquante plus élevée (emportée par les deux neutrinos de l'état final) que pour des interactions de neutrinos électroniques. D'autres variables cinématiques en plus de cette énergie manquante permettent de mettre en place une analyse par rapport de vraisemblance pour distinguer les interactions des neutrinos tau de celles des neutrinos électroniques.

Le rapport d'embranchement de la désintégration du tau en électron est d'environ 18%. Une étude complémentaire portant sur des canaux de désintégration semi-leptoniques est déployée. En particulier, le tau possède un large canal de désintégration en méson rho (25%) qui se désintègre presque immédiatement en un pion neutre et un pion chargé. La signature de cette désintégration est en particulier identifiable par le calcul de la masse invariante des deux particules filles. Cela offre la possibilité de rechercher des interactions de neutrinos tau par ce canal de désintégration, ce qui constitue l'objet du quatrième chapitre de cette thèse. Un dernier chapitre, plus court, propose une extension de l'analyse du chapitre 4 pour un mode de désintégration du tau plus exclusif composé de seulement un pion chargé (et d'un neutrino tauique non détecté). Enfin et surtout, il rapporte la sensibilité de ces trois canaux de désintégration combinés.

Deux chapitres d'introduction permettent une mise en contexte précise de cette thèse. Le chapitre 1 donne une présentation générale de la physique des neutrinos avec un accent marqué sur le formalisme permettant de décrire leurs oscillations. L'expérience DUNE y est également exposée ainsi les enjeux liés à la physique de la saveur tau dans le secteur des neutrinos. Le chapitre 2 présente quant à lui les outils et concepts clefs pour comprendre la description et la simulation des interactions des neutrinos. Une partie est également consacrée à la simulation de la réponse des détecteurs de manière générale, et en particulier de ceux s'appuyant sur la technologie des chambres à projection temporelle à argon liquide.

⋆ ⋆ ⋆

Table of Contents

List of Figures

List of Tables

Introduction

The discovery of neutrino oscillations in 1998 by the Super-Kamiokande collaboration was a major breakthrough of recent particle physics. It proved neutrinos to be massive, which can not be accounted for by the current Standard Model of particle physics. A minimal extension adding three Dirac right-handed neutrino fields allows describing massive neutrinos. It also allows incorporating a neutrino mixing between the flavour states and the mass states and it is carried by the PMNS (Pontecorvo–Maki–Nakagawa–Sakata) matrix, the CKM equivalent mixing matrix of the neutrino sector.

The PMNS matrix is parametrized with three mixing angles θ_{12}, θ_{13} and θ_{23} and one complex phase δ_{CP}. Over the last two decades an extensive experimental program aimed at determining and measuring the PMNS matrix elements. In addition neutrino oscillations depend on the three square mass differences $\Delta m^2_{ij} = m^2_j - m^2_i$ (with $j > i$) of the neutrino mass states among which only two are mathematically independent. Historically the study of solar and reactor neutrinos at long-baselines allowed to precisely measure θ_{12} and Δm^2_{21} while the study of atmospheric neutrinos allowed to measure θ_{23} and Δm^2_{32}. In the last decade the mixing angle θ_{13} was found to be non-zero and opened the path to the search for the constraining of the last unknown parameter which is the complex phase δ_{CP} via the study of the sub-leading $\nu_\mu \to \nu_e$ oscillations in the atmospheric neutrinos sector.

DUNE (Deep Underground Neutrino Experiment) is a major forthcoming next generation of long-baseline neutrino experiment which will study this point by exploiting the world's most intense 1.2 MW neutrino beam associated to a near detector facility at Fermilab (Illinois, US). A far detector complex located 1285 km away at Sanford Underground Research Underground in South Dakota (US) is designed to host four liquid argon detector modules each of about 10 kt fiducial mass.

The establishment of non-0 and non-π values for the phase δ_{CP} would establish the Charge Parity (CP) symmetry violation in the leptonic sector. Cosmologists are in particular interested in this measurement since it could be the key to understand the underlying mechanism for the observed matter/antimatter asymmetry in the universe. DUNE will in addition have an unprecedented sensitivity to τ neutrinos thanks to the large $\nu_\mu \to \nu_\tau$ oscillation probability along its baseline with an expected event rate of about 30 ν_τ charged current interactions per 10 kt per year. As for comparison the combined DONuT and OPERA results, which were past experiments searching for τ neutrino interactions, isolated a total of 18 candidates. Super-Kamiokande and more recently IceCube each reported an almost 5σ sensitivity to the τ neutrino appearance through atmospheric neutrino oscillations. The study of the $\nu_\mu \to \nu_\tau$ oscillation channel will allow to perform a 3-flavour phenomenology in which the τ neutrino appearance is combined with the electron neutrino appearance and the muon neutrino disappearance to better constrain the oscillation parameters. In addition it will make it possible to perform crucial tests of the 3-flavour paradigm and of the unitarity of the PMNS mixing matrix.

However the identification of the τ neutrino flavour relies on the observation of the coupled charged lepton $\tau^{\pm}$ which in DUNE is expected to decay too promptly for its track to be identifiable. The search for τ neutrino appearance then relies on the possibility to identify $\tau^{\pm}$ decays in the detectors. The chosen technology for the far detectors is the Liquid Argon Time Projection Chamber (LArTPC). It makes it possible to picture precisely the neutrino interactions with a few millimeter spatial resolution and at the same time to perform an excellent tracking, identification and energy measurement of final state particles. Such a performance opens the path to search for τ neutrino interactions relying on kinematical criteria like in the pioneering NOMAD experiment. The study of kinematical criteria in the DUNE far detectors to search for τ neutrino charged current events at the simulation level is the subject of this thesis. DUNE will have also the possibility to change the beam configuration in order to enhance the detection of charged current ν_τ interactions.

The content of the thesis is organized in five chapters. Chapter 1 contains a general introduction to neutrino physics. Some non-exhaustive key milestones of this field are described. A formal introduction to the neutrino oscillations formalism to set the context of DUNE is provided. A final section presents the current key motivations for doing τ neutrino physics. Chapter 2 reports the main analysis tools and the basic ingredients to describe neutrino interactions at the simulation level. A discussion is aimed at understanding the key concepts which allow simulating the detector response and the energy reconstruction at the level of single particles. Chapters 3 & 4 report the ν_τ appearance search analysis based on kinematical criteria respectively for the $\tau^- \to e^-\nu_\tau\bar{\nu}_e$ and $\tau^- \to \rho^-\nu_\tau \to \pi_0\pi^-\nu_\tau$ decay modes which both have a large (18% and 25%) branching ratio. Chapter 5 presents an extension of chapter 4 to a more exclusive decay mode $\tau^- \to \pi^-\nu_\tau$. A final chapter closes the thesis with a combined analysis of the three decay modes and a conclusion.

$\star$ $\star$ $\star$

Chapter 1

Neutrinos and DUNE

Contents

The history of neutrino physics is fascinating and this chapter first offers a brief overview of the key moments which marked this research domain. Given the tiny interaction rate of these particles in comparison to standard matter, the technology deployed in the experiments always had to be at the forefront. This chapter in particular proposes to shed light on some past and current experiments to discuss their design and their contribution to the field. Then a formal introduction to the description of neutrino oscillations both in vacuum and in matter is given. This allows to set the context for the future long-baseline DUNE (Deep Underground Neutrino Experiment) for which a conceptual design is described. Finally the chapter closes with a special care given to the τ neutrino. It is the least observed of the three neutrino flavours and the least known particle of the Standard Model of particle physics. It is expected that DUNE contribute significantly to increase the available data of the τ neutrino, which motivates the need for a τ search analysis in DUNE.

1.1 Major milestones

1.1.1 Discoveries of neutrino flavours

The history of neutrinos is about one century long. It started with the issue of the beta radioactivity in 1914: James Chadwick and his team discovered that the beta radioactivity spectrum was continuous [1], unlike the alpha radioactivity which had a monochromatic spectrum. Thus, the energy carried by the emitted electron in the process of a single beta decay (the differentiation between β^+ and β^- was not known yet) ${}^A_Z X \rightarrow {}^A_{Z+1} X + e^-$ seemed to be allowed to take any energy value between 0 and the energy of the reaction. The interpretation of the spectrum itself divided physicists. For instance, Meitner explained the spectrum shape with secondary effects: the first emitted electron of beta disintegration was responsible for other electron emissions of variable energy, via $e^- e^-$ scatterings, resulting in the observation of a continuous spectrum. In 1927, Ellis and Wooster proved Meitner wrong: the continuous spectrum corresponded to the spectrum of the electron directly emitted by the disintegration [2]. An other interpretation, offered by Niels Bohr, suggested that energy conservation would only make sense on a statistical point of view.

Then it came the neutrino hypothesis. It consisted in postulating that β-radioactivity was more complex than α-radioactivity, which is a two-body decay. The atomic nucleus in the final state is much heavier than the emitted α-particle, thus carrying away negligible kinetic energy in the process. The expected energy spectrum of the α-particle should then be monochromatic. The continuous spectrum could, by extension, be explained if the β-decay was a three-body process, involving an undetected third particle in the final state. This is the idea Pauli expressed in what is now a famous letter [3] in 1930. For completeness, though we will not study this point here, Pauli actually reconciled in one hypothesis two issues: the nuclear structure and the continuous beta spectrum. Indeed, the two atomic nuclei (initial and final state) in a beta decay have the same number of nucleons, hence the spin variation of the atomic nucleus is integer, while the emitted electron was known to have a half-integer spin, following the observation of anomalous Zeeman effect in 1898. At that point in history, the neutrino would be a fermion (spin 1/2) with a null electric charge, and a mass at most comparable to the one of the electron and obey Fermi-Dirac statistics.

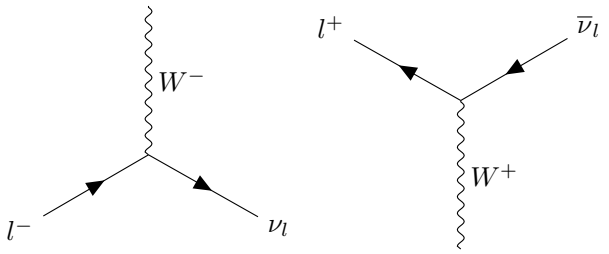
Fermi first suggests in 1933 a theory describing the process of a β-decay including the hypothetical neutrino. The electron is not considered bound inside the nuclei in an unknown manner. Instead, it is treated like photons in the emission of light quanta by excited atoms. Thus, the

electron (and the neutrino) becomes a particle which can be emitted and absorbed, so that the total number of electrons is not constant anymore. The theory describes the β emission as a neutron turning into a proton, coupled to the creation of an electron and a neutrino (which today is actually known to be an electron antineutrino). The interactions is modelled as a 4-point fermion vector current (no intermediate boson). The first evaluation of the coupling constant by Fermi was $3.25 \times 10^{-6}\text{GeV}^{-2}$, in comparison to the modern Fermi constant $G_F = 1.17 \times 10^{-5}\text{GeV}^{-2}$. The theory also allowed to estimate the cross section of the process where the neutrino would interact on a proton to produce a neutron (*i.e* the inverse process). For a 1 MeV neutrino interacting on a proton at rest, the cross section reaches $G_F^2 E_\nu^2 \approx 10^{-43}\text{cm}^2$, where E_ν denotes the neutrino energy. As for comparison, the Thomson cross section of a photon scattering on an electron is typically of the order of 10^{-24}cm^2, 19 orders of magnitude above the neutrino scattering. The original paper of Fermi was published in his native language [4], an English version can be found here [5]. Today, the β-decay is well described in the electroweak framework of the Standard Model, and at low energy (in comparison to the W gauge boson mass), one can approximate the model to Fermi's theory.

Detecting neutrinos straight away appeared to be a difficult task. The main challenge was to find a suitable neutrino source, in a way that a large flux would compensate the expected tiny cross section of neutrino interactions. This challenge was solved by using the Savannah River (USA) nuclear reactor in 1956 by Frederick Reines and Clyde L. Cowan [6] (the original design of the experiment planned to use an atomic bomb...). Thanks to a detector filled with 4200 l of liquid scintillator, they observed the coincidence of the positron and the neutron signatures that were hypothetically given by the inverse beta process $p + \bar{\nu} \rightarrow n + e^+$. The pair gathered signals (about 3 neutrino-like signals per hour) that could only be explained by the absorption of a neutrino by an atomic nucleus (inverse beta decay).

Today this result is associated to the discovery of the electron antineutrino, as the Standard Model includes three active neutrino flavours (each flavour is associated to the corresponding charged lepton, electron, muon or tau) below $45\,\text{GeV}/c^2$. The discovery of the muon flavour occurred in 1962 at Brookhaven National Laboratory [7] by looking at the type of the final state charged lepton from interactions of neutrinos originating from the decays of charged pions. The fact that only muon-like tracks were observed implied the neutrino associated to a final state muon is of different type (flavour) from that emitted with an electron in beta decays. The observation of pion disintegration highlighted the emission of a neutrino associated to a final state muon and not of an electron as in beta decay. In 1975, the discovery of the τ lepton strongly suggested the existence of a third neutrino flavour. In 1990 the ALEPH experiment at LEP conducted a measurement consisting in comparing the Z boson decay width to the Standard Model predictions. The hypothesis of three neutrino flavours coupled to the Z boson was consistent with observations [8]. It should be noted, however, that this experiment could not exclude the existence of neutrino families with masses above half of the Z boson mass ($\approx 45\,\text{GeV}/c^2$), since the Z boson can not decay into such a pair of neutrino-antineutrino. In 2000, the DONUT collaboration directly observed the τ neutrino at Fermilab [9].

Neutrino history shows how important experiments were. It is a domain of physics where, because of the smallness of cross sections at play, the technology of the detectors must solve the challenge of reconciling precision measurements of the neutrino interactions with the deployment of very massive targets.



(a) neutrino (b) anti-neutrino

Figure 1.1: Charged currents vertices of neutrino and anti-neutrino. The flavour of the leptons is noted with the letter "l".

1.1.2 Neutrinos charged current interactions

In the Standard Model, neutrinos are leptons. They have spin 1/2, they are massless and by carrying no electric charge, the only interactions they are sensitive to are the weak interactions. They can interact by both neutral currents (exchange of a Z boson) and charged currents (exchange of a W boson, see the corresponding vertices in figures 1.1a and 1.1b) interactions. Three neutrino flavours exist: electron, muon and tau, each naturally associated to the corresponding lepton via the charged current interaction. For this reason, charged currents are pretty convenient for the experimental study of neutrino oscillations as they make it possible to tag the flavour of the impinging neutrinos. Experiments based on the detection of neutral current interactions can be performed, but they can not distinguish between the three neutrino flavours.

The most convenient charged lepton to identify is the muon, which has a long lifetime and low energy losses (via ionisation only at the energy considered) producing in the detectors a characteristic long, clean, and straight tracks. The electron, much lighter, loses energy by radiation and produces electromagnetic showers in the detectors. The tau lepton, at the considered energies (current neutrino beams used are typically of order 1 GeV), is hard to identify because of its short distance of flight before decaying (sub-millimetric), reaching the limits of detectors spatial resolution.

1.1.3 The discovery of neutrino oscillations

In the 70s, the Homestake experiment conducted by Raymond Davis and John N. Bahcall, first measured the electron neutrinos produced by the sun and found a deficit with respect to predictions. They used the reaction $\nu_e + {}^{37}\mathrm{Cl} \rightarrow e^- + {}^{37}\mathrm{Ar}$, and regularly counted the resultant quantity of argon produced in the detector by extracting it from the detector, and measuring its radioactivity in a proportional counter. Over the years, the Homestake collaboration measured a solar neutrino flux roughly three times below the predictions [10] (2.32 SNU [1], which corresponded to 0.437 argon atoms produced per day). The large discrepancy between these observations and the predicted solar neutrino flux triggered and maintained the so-called solar neutrino problem from the end of the 70's to the end of the 90's. To explain this anomaly, many alternative calculations of the solar flux were proposed. None, however, were able to resolve the discrepancy. Bruno Pontecorvo, believing that both the experiment and the solar flux calculations were correct, proposed to interpret this result as evidence of a new physical phenomenon: neutrino oscillations. Actually his proposal occurred in 1967 even before the first result of the

[1] The SNU is the Solar Neutrino Unit. It corresponds to 10^{-36} neutrino capture/atom/s.

Davis experiment [11]. Though, this hypothesis did not immediately convince the community, mainly because it required to extend the Standard Model to account for this phenomenon.

In the 1980's, in Japan, the Kamiokande water-Cherenkov experiment was started. Among its collaborators, Takaaki Kajita (Nobel Prize for physics in 2015 for the discovery of neutrino oscillations) had to evaluate the background from atmospheric neutrinos for the search of the proton decay (one decay channel at stake is the proton decaying into an electron and a π_0). Surprisingly, data were not in agreement with the predictions for the ratio of populations of electron and muon atmospheric neutrinos [12, 13]. This observation led to another anomaly named the "atmospheric neutrinos anomaly". Subsequent studies showed that this deficit depended on the angular direction considered (that is to say, if the neutrino is produced up in the atmosphere above the detector or on the one on the other side of the Earth and in that case it seems to "come from below"). Super-Kamiokande, the successor of the Kamiokande detector which started to operate in 1996, observed that the flux of atmospheric electron neutrinos was in agreement with predictions, as well as the flux of muon neutrinos produced in the atmosphere above the detector. There was a clear deficit, however, of muon neutrinos which was increasing when going at larger zenith angles pointing to the other side of the earth [14]. This angular (baseline) dependence of the deficit provided the first evidence of atmospheric neutrino oscillations.

Neutrino oscillations were demonstrated as well in 2001 by the SNO (Subdury Neutrino Observatory) experiment, led by Arthur B. McDonald (physics Nobel Prize 2015). The team measured simultaneously the ν_e flux from the sun (using charged current reactions, that only the electron flavour can yield given the neutrino low energy) and the total neutrino flux from the sun regardless of the flavour (using neutral current reactions). It was found that, while the ν_e flux had a deficit reproducing the data of the solar neutrino problem, the total neutrino flux was in agreement with the predictions [15]. This was a clear evidence of ν_e oscillation into the two other flavours. Within three years, both the solar neutrino anomaly and the atmospheric neutrino anomaly were solved by the neutrino oscillations discovery.

1.1.4 A list of key neutrino experiments

In this section we will go through a non-exhaustive list of neutrino experiments, which established some milestones in neutrino detection and in the establishment of neutrino oscillations. We will discuss their design, their objectives and their results. The reader should be aware that the proposed list is not complete, and it is mostly an orientated choice of the author.

Hanford and Savannah River

The history reminds of the Savannah River as the key experiment that first detected the neutrino. But three years earlier, another experiment, conducted by the same scientists Reines and Cowan, had set up a technique to detect the neutrino, next to the nuclear reactor of Hanford (Washington) [16]. They used a 300 l detector, filled with various scintillators and protected by shielding (mainly to stop background neutrons produced by the reactor). The objective of this experiment was to isolate the signal of the inverse beta decay $\bar{\nu}_e + p \rightarrow e^+ + n$, choosing scintillators containing a large fraction of hydrogen. The positron, once emitted, would produce an annihilation signal in two photons and photomultipliers in the detector would be able to detect this signal. The neutron, however, is much harder to detect. The team added cadmium salt in the scintillator, increasing the absorption cross section of the neutron by the medium. This absorption was accompanied by a light emission as well. A few microseconds delay was

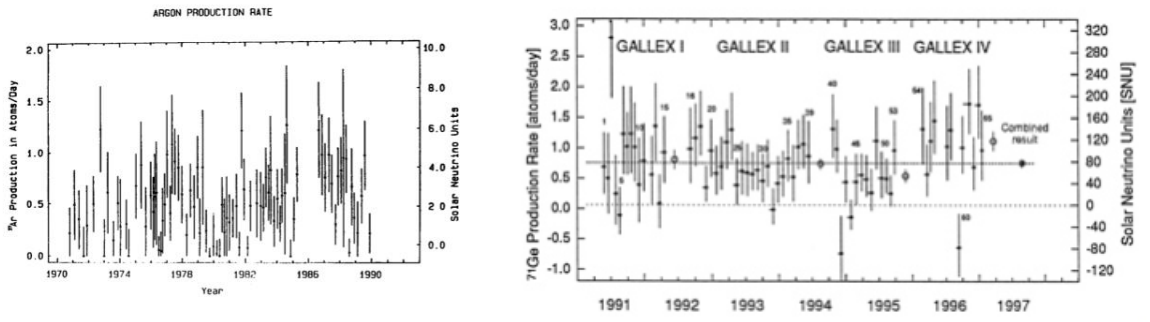


Figure 1.2: The argon production rate (left y axis) and solar neutrino flux (in SNU units) of the Homestake experiment between 1970 and 1990 [10], and from the GALLEX experiment between 1991 and 1997 [17].

expected between the positron signal and the neutron one. Thus, the team was looking for delayed coincidences. Actually, the idea to be able to capture the neutron was the argument that made it possible to transit toward a nuclear reactor source of neutrino. Originally, the experiment planned to use an atomic bomb as the neutrino source, because the flux would have been much higher.

The Hanford experiment could not conclude with definitive evidence of the (anti)neutrino signal. The signal to background ratio was too low (mainly because of cosmic rays), so that the neutrino signal was drowned. Still, while turning off and on the reactor, the team observed a variation in the number of delayed coincidences in agreement with the expected signal, which was 0.1 to 0.3 count per minute. That was a serious hint, but could not be considered as the evidence of the neutrino discovery. Three years later, the team designed a new experiment at the Savannah River nuclear plant, taking advantage of the previous inconclusive experiment, and finally made it to isolate the neutrino signal [6].

The Homestake Experiment

This experiment took place in the Homestake Gold Mine in South Dakota, and remained famous for being at the origin of the solar neutrino problem. Its purpose was indeed to measure the electron neutrino flux emitted by the sun, via the nuclear reaction $\nu_e + {}^{37}\mathrm{Cl} \rightarrow e^- + {}^{37}\mathrm{Ar}$, and then counting the number of argon atom produced (the isotope produced is radioactive and has a half-life of 35 days). A summary of the results published by the team experiment is shown in figure 1.2. The mean capture rate was 2.32 SNU $\pm$ 0.22, which was below any theoretical expectations by a factor of 3, approximately (see [10, Table 2]).

GALLEX

Standing for "Gallium Experiment", GALLEX was hosted in the underground laboratory of Gran Sasso in Italy in the 1990's. It aimed at measuring the solar electron neutrino flux at lower energy than Homestake, corresponding to the majority of neutrinos produced in the sun by the proton-proton cycle. The detector measured the electron neutrino reaction on Gallium $\nu_e + {}^{71}\mathrm{Ga} \rightarrow e^- + {}^{71}\mathrm{Ge}$, by regularly capturing the germanium produced. The observed solar neutrino flux was found to be $77.5 \pm 6.2^{4.3}_{4.7}(1\sigma)$SNU [17], whereas the theoretical predictions were, again, much higher than this measurement (see for instance [18]).

Kamiokande/Super-Kamiokande

Kamiokande and its successor Super-Kamiokande perhaps are the world's most famous neutrino experiments. The detection of neutrinos is based on the Cherenkov effect: charged particles, produced by neutrinos interacting in the water contained in the detector, propagating faster than the speed of light in water emit Cherenkov radiation. This radiation is detected with photomultipliers tubes that are installed all along the walls of the detector. The Cherenkov light is emitted on a cone around the trajectory of the charged particle. The projection of this moving cone on the walls of the detector results in rings which have sharp contours in the case of muons, just loosing energy in water, and fuzzier contours for electrons which also shower. The shape of the rings provides very good identification capabilities of muons and electrons. The experiment is then basically relying on the identification and the measurement of the charged lepton from the interactions of muonic and electronic neutrinos.

Super-Kamiokande is in fact the evolution of a first generation massive water detector installed in Kamiokande, Japan. Kamiokande, containing 3000 t of water, was installed in the 1980s. It observed in 1987 the Supernova 1987A in the Large Magellanic Cloud, capturing 12 neutrinos from this explosion [19]. As mentioned previously, this experiment was also at the origin of the atmospheric neutrino problem.

Super-Kamiokande has a detector filled with 50 000 t of ultra pure water. It is a 40 m large (diameter) and 40 m high cylinder, instrumented with 11 146 photomultipliers. The main result associated with the experiment is the neutrino oscillations discovery [14], proving the disappearance effect of muonic atmospheric neutrinos as a function of the zenith angle.

In addition to atmospheric neutrinos, the detector was also used by the long-baseline experiment K2K (1999-2004), and later by T2K which started to operate in 2010. The latter uses a human made beam of neutrinos produced at Tokaï using the J-PARC (Japan Proton Accelerator Research Complex). Neutrinos propagate over a distance of 295 km in the Earth before reaching the Super-Kamiokande detectors. The beam was optimized with an off-axis technique which maximizes the flux in the region around the first oscillation maximum. T2K is currently operating and is projected to run until the end of 2026.

The next evolution of the detectors at Kamioka is Hyper-Kamiokande, approved for construction in 2019. The detector will consist of two gigantic cylinders of ultra pure water, 74 m large and 60 m high, though only one is foreseen to be installed at Kamioka. The fiducial volume of the detector aims to be 10 times larger than the one of Super-Kamiokande, in order to search for CP violation using the neutrino beam from Tokai.

ν_μ discovery

It is worth saying a word about the muon neutrino discovery (Jack Steinberger, Melvin Schwartz and Leon Lederman were rewarded in 1988 by the Nobel Prize). Not only the team proved the existence of a neutrino associated to the production of a muon (and not an electron) [7], but they were the first to employ a proton accelerator to create a neutrino beam. They used the AGS (Alternative Gradient Synchroton) built at Brookaven National Laboratory to send 15 GeV protons on a beryllium target. The decay in flight of produced charged pions allowed for the production of the first artificial neutrino beam. This method is still employed today by all neutrino beam experiments.

NOMAD.

NOMAD [20, 21] was a short (< 1 km) baseline experiment which operated at CERN between 1994 and 1998 with the SPS (CERN) based neutrino beam. The beam was produced with the SPS 450 GeV protons and it had a mean energy of 24 GeV. The primary goal of this experiment

was to search for $\nu_\mu \rightarrow \nu_\tau$ oscillation appearance at small distances (the detector was located 820 m away from the beryllium target used to produce the neutrino beam). The detector was composed of a set of sub detectors devoted to the measurement of the neutrino interactions and in particular providing very good electron identification capabilities in order to exploit low background in the electronic τ decay channel. The biggest one was the neutrino target made of a set of drift chambers, which had a fiducial surface area of $2.6 \times 2.6\,\mathrm{m}^2$ (perpendicular to the neutrino beam) and a fiducial mass target of 2.7 t. The detector acted like an electronic bubble chamber. It was combined with a magnetic field, which allowed for an excellent momentum resolution of charged particles passing through the drift chambers. One important feature of the detector was the combination of a transition radiation detector (which allowed for electron identification and pion rejection) and an electromagnetic calorimeter with a pre-shower detector to precisely measure the energy of electrons and photons. All-together these detector provided electron identification with 90% efficiency and rejection of pions at the level of 10^6.

NOMAD was the first neutrino experiment dedicated to τ neutrino search based only on kinematical criteria, as it was suggested in 1979 by Robert Shrock and Carl Albright [22, 23] with what they had called the muon trigger test. At the time, the electron identification in the neutrino detectors was not good enough to consider a search via the $\tau \rightarrow e$ decay channel. The claim for ν_τ discovery was not the only motivation of NOMAD. In fact, the "to be discovered" τ neutrino was thought to, perhaps, have a mass in the range 1-10 eV/c^2. Combining this hypothesis with an expected small neutrino mixing (similar to what had been measured in the quark sector), the assumed large Δm^2 indeed motivated the search for τ neutrino at short baselines (the forthcoming equation 1.15 shall illustrate this reasoning). If NOMAD were to find signals consistent with ν_τ events, this would have corresponded not only to the first detection of tau neutrinos but also a discovery of a first serious hot dark matter candidate. Today we know that astrophysical observations favour the cold dark matter models (thus excluding neutrinos as dark mater candidates), and that cosmological arguments allow constraining the sum of all three neutrino mass states at the sub-eV level.

The collaboration reported a succession of negative results [24, 25, 26] with the final one reported in [27]. The data analysis covered 82.5% of the τ decay branching ratio and set a 90% CL upper limit of the probability oscillation $P(\nu_\mu \rightarrow \nu_\tau) \leq 1.63 \times 10^{-4}$. Considering the two neutrino families formalism, the resulting 90% CL interval can be reported in the plane $[\Delta m^2\,;\,\sin(2\theta_{\mu\tau})]$ (we refer to the coming equation 1.16 of the same chapter for further understanding). This constraint still holds today as the most stringent exclusion region.

Here we have decided to report a "key" experiment for which the final results were negative, in the sense that there were no discovery associated. Still, the work of this thesis will largely be inspired by NOMAD's methods, and we sketch that it was the first experiment looking for τ neutrino appearance based only on kinematical arguments (or indirect search). It is also worth to note that the combined negative reports of NOMAD and CHORUS (one competing experiment at CERN) participated in excluding the neutrinos as a hot dark matter candidate.

Recent reactor experiments

The double CHOOZ experiment was the extension of CHOOZ, a neutrino reactor experiment in Northern France. CHOOZ made use of a single detector located 1.05 km from the nuclear cores, while double CHOOZ used an additional and identical detector located 400 m from the cores. During the 2000's, the collaboration held the best constraint of the last measured mixing angle of the PMNS matrix: $\sin^2(2\theta_{13}) < 0.2$.

Neutrino interactions were recorded through the Inverse Beta Decay (IBD) process $\bar{\nu_e} + p \rightarrow e^+ + n$. The use of a $10^3 \mathrm{m}^3$ liquid scintillator enriched with Gadolinium allowed for the detection of both the positron and the neutron. The positron first emits light in the medium, and the delayed absorption of the neutron, produces a specific light signature. Coincidental records of the two can be attributed to a IBD process. PMTs outside the detection volume allowed for energy reconstruction of the events.

In 2012, the concurrent experiment Daya Bay published a measurement of the neutrino interactions spectrum at the far site, excluding the $\theta_{13} = 0$ hypothesis with more than 5σ [28]. They measured $\sin^2(2\theta_{13}) = 0.092 \pm 0.016(\mathrm{stat}) \pm 0.005(\mathrm{syst})$. θ_{13} is sometimes referred as the "reactor" angle for the reason that reactor experiments measured it thanks to $\bar{\nu}_e$ disappearance. In the years following the Daya Bay result, double CHOOZ confirmed similar results [29] with $\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029}$. This corresponds to a small mixing angle value $\theta_{13} = 8.6°^{+1.4}_{-1.5}$.

NOvA

NOvA and T2K are the two present competitor long-baseline beam neutrino experiments. We will mainly focus of NOvA since the DUNE experiment can be thought as its successor. NOvA [30] (NuMI Off-Axis ν_e Appearance) uses the NuMI (Neutrino at the Main Injector) neutrino beam produced at Fermilab, itself taking advantage of the 120 GeV protons of the Main Injector Fermilab accelerator. The experiment deployed a 0.3 kt near detector at Fermilab and a 14 kt far detector 810 km away from the production point at the Ash River site. Both detectors make use of segmented cells filled with liquid scintillator which allows for particle's track resolution with a precision of few centimeters.

NOvA recently published a combined ν and $\bar{\nu}$ modes analysis which tends to favour the neutrino mass normal hierarchy by 1.9σ and the θ_{23} upper octant by 1.6σ [31]. The vocabulary just used has not been explained yet, and it is the purpose of the following sections to introduce the neutrino oscillations formalism which will allow us to properly present the DUNE experiment and its main scientific program.

1.2 Neutrino oscillations

1.2.1 How could neutrinos have mass ?

Neutrinos are described in the Standard Model as left-handed Dirac fields ν_e, ν_μ and ν_τ, which are the ones that appear in the leptonic weak charged current and neutral current of the weak Lagrangian. The purpose of these fields is to account for the weak behavior of neutrinos. For instance, the β decay $n \rightarrow p + e^- + \bar{\nu_e}$ can be described by the electroweak Lagrangian, more specifically the charged current component of that Lagrangian.

Since non-massive neutrinos can not oscillate, a non-null mass must be added to the Dirac Lagrangian. In the Standard Model, the mass of the fermions is generated by the Higgs mechanism as a result of a Yukawa coupling with the Higgs scalar boson. In general, a mass term requires a coupling between the left-handed and right-handed fields of a fermion.

It is convenient to work at first with the case of quarks to understand the framework. Let $Q_{\alpha L}(\alpha = 1, 2, 3)$ be the left-handed doublets of the 3 generations of quarks ($Q_{1L} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$) , and $q^U_{\alpha R}$ and $q^D_{\alpha R}$ the corresponding right-handed fields (they're singlet under the electroweak gauge

group, U refers to the quarks for which the corresponding left-handed fields have $I_3 = 1/2$ and D to the quarks for which the corresponding left-handed fields have $I_3 = -1/2$). In the unitary gauge, the Yukawa coupling of such fermions reaches:

$$\mathscr{L} = - \left(\frac{v+H}{\sqrt{2}}\right) \sum_{\alpha,\beta=d,s,b} (Y^D_{\alpha\beta} \bar{q}^D_{\alpha L} q^D_{\beta R} + h.c) \tag{1.1}$$

$$- \left(\frac{v+H}{\sqrt{2}}\right) \sum_{\alpha,\beta=u,c,t} (Y^U_{\alpha\beta} \bar{q}^U_{\alpha L} q^U_{\beta R} + h.c). \tag{1.2}$$

The fermion fields which appear in the Lagrangian are the weak fields, that is to say the ones used in the electro-weak currents. However, there's no reason to assume that these fields are the ones diagonalising the matrices Y^U and Y^D. Thus, there exist four matrices, often noted V^D_L, V^D_R, V^U_L, V^U_R such that the two matrices $Y^D_{mass} = V^{D\dagger}_L \cdot Y^D \cdot V^D_R$ and $Y^U_{mass} = V^{U\dagger}_L \cdot Y^U \cdot V^U_R$ are diagonal, allowing to define respectively the mass fields of the quarks d,s,b and u,c,t as:

$$\boldsymbol{q}^U_{L,mass} = V^{U\dagger}_L \boldsymbol{q}^U_L, \tag{1.3}$$

$$\boldsymbol{q}^D_{L,mass} = V^{D\dagger}_L \boldsymbol{q}^D_L, \tag{1.4}$$

$$\boldsymbol{q}^U_{R,mass} = V^{U\dagger}_R \boldsymbol{q}^U_R, \tag{1.5}$$

$$\boldsymbol{q}^D_{R,mass} = V^{D\dagger}_R \boldsymbol{q}^D_R. \tag{1.6}$$

Now, one can re-inject these mass fields in the weak currents of the Standard Model. Doing so, one would account for the description of the quark mixing, which is described by the product of the matrix V^D_L and V^U_L only. This is the so-called CKM matrix (Cabibbo-Kobayashi-Maskawa) V, which has to be measured experimentally:

$$V = V^{U\dagger}_L \cdot V^D_L. \tag{1.7}$$

It should be noted that the two matrices with index R do not appear in the quark mixing description, since the weak currents only couples the left-handed Dirac fields.

The same process can be applied to the leptonic sector, defining the weak doublets and singlets as:

$$L_{\alpha L} = \begin{pmatrix} \nu_{\alpha L} \\ \alpha_L \end{pmatrix}, \nu_{\alpha R}, \alpha_R, \quad (\alpha = e, \mu, \tau). \tag{1.8}$$

The three neutrino flavour fields play the role of the quarks u,c,t (earlier indexed U because they have a weak isospin $I_3 = 1/2$), while the three charged lepton fields play the role of the quarks d,s,b (earlier indexed D) because they have $I_3 = -1/2$. The mixing matrix appearing is called the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrix, and it is the one arising from the massive neutrino fields definition:

$$\nu_i = \sum_{\alpha=e,\mu,\tau} U^{\star}_{\alpha i} \nu_{\alpha}. \tag{1.9}$$

This process we went quickly through is called the *minimally extended Standard Model*, and it allows for the description of massive Dirac neutrinos. There has been an extension from the Standard Model, because we injected three right-handed neutrino fields, one for each leptonic flavour. Indeed, the generation of the mass of a fermion requires a Lagrangian term proportional to the product of the left-handed and right-handed fields. In the Standard Model, neutrinos are only described by a left-handed Dirac field, thus they are massless. This is the reason why neutrinos are said to be an evidence of physics beyond the Standard Model. If neutrinos are Majorana fermions, this mechanism presented does not stand anymore. Reader can find a nice discussion of this point on [32, p.188]. These are the two ways of generating massive neutrinos.

1.2.2 Neutrino oscillations mechanism

The PMNS matrix exposed in the previous section is parametrized by four parameters. Indeed, a general result is that an unitary matrix of size N depends on N^2 independent real parameters, which can be divided into $N(N-1)/2$ mixing angles and $N(N+1)/2$ phases. For $N = 3$, we end up with three mixing angles and six phases. The only place of the Standard Model where these phases appear is in the weak charged current:

$$j^{\rho}_{W,L} = 2\,\bar{\mathbf{n}}_L\, U^{\dagger} \gamma^{\rho} \mathbf{l_L} = 2 \sum_{\alpha=e,\mu,\tau} \sum_{i=1}^{3} \bar{\nu}_{iL} U^{\star}_{\alpha i} \gamma^{\rho} l_{\alpha L}, \tag{1.10}$$

where $\bar{\mathbf{n}}_L$ denotes the triplet of chiral massive neutrino fields $n_{i=1,2,3}$. In this expression, we can choose to redefine each fermion field with a global phase without changing the physics predicted by the charged current. Only five of them can be redefined with a global phase independently, since the addition of a global phase to the current can always be chosen so that one of the fermion phase is absorbed. The five fermion phases can be chosen to absorb five of the six phases of the PMNS matrix, which leaves the number of independent parameters to 3 mixing angles and one real phase. The PMNS matrix is conventionally parametrized as:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_{23}) & \sin(\theta_{23}) \\ 0 & -\sin(\theta_{23}) & \cos(\theta_{23}) \end{pmatrix} . \begin{pmatrix} \cos(\theta_{13}) & 0 & \sin(\theta_{13}) e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -\sin(\theta_{13}) e^{i\delta_{CP}} & 0 & \cos(\theta_{13}) \end{pmatrix} . \begin{pmatrix} \cos(\theta_{12}) & \sin(\theta_{12}) & 0 \\ -\sin(\theta_{12}) & \cos(\theta_{12}) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where θ_{12}, θ_{13} and θ_{23} are the three mixing angles. The phase parameter δ_{CP} is indexed with the letters CP because of its close connection to charged-parity (CP) symmetry violation. Wu and her team discovered in 1957 parity symmetry (P) violation observing in β-decaying ^{60}Co in a magnetic field [33]. This was a huge breakthrough at the time ! The Fermi theory was modified into the $V \pm A$ theory in which the vector current was added an axial-vector component [34, 35]. The sign of the A component was determined by experiments, see for instance the one from Goldhaber in 1958 [36] in which the team determined the neutrino helicity by measuring the photon polarization of the de-exciting Sm following the electron capture of Eu (spin values are indicated in brackets):

$$e^-(1/2) + {}^{152}\mathrm{Eu}(0) \rightarrow \nu(1/2) + {}^{152}\mathrm{Sm}^\star(1) \rightarrow \nu(1/2) + \gamma(1) + {}^{152}\mathrm{Sm}^\star(0). \tag{1.11}$$

In this particular configuration, the final state photon and neutrino must have opposite helicities in order to conserve angular momentum, so a measure of the photon polarization indeed allows to measure the polarization of the neutrino. This measurement determined the sign of the axial-vector component to be negative: the $V - A$ theory. It predicted that only left-handed neutrinos and right-handed anti-neutrinos took part in weak interactions. But because spin is invariant under charge (C) transformation, the $V - A$ theory predicted C symmetry violation as well. The combined transformation CP, however, preserves the weak current, and it was thought to be the true symmetry between matter and antimatter. The study of neutral kaons by Fitch and Cronin in 1964 disproved this statement [37]. The small CP symmetry violation is carried by the complex phase parameter of the CKM matrix, the equivalent of the PMNS matrix of the quark sector. We shall see later in this chapter (1.3.1) that the δ_{CP} parameter indeed carries a possible CP symmetry violation in the neutrino oscillations phenomenon.

We now suggest to discuss what one will find in the literature as the so-called standard neutrino oscillations derivation. It is the simplest, though not very correct (see for instance [32, p.253]), manner of deriving the oscillation probabilities between neutrino flavours. The starting point of this derivation is to start from a neutrino of flavour α:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^\star_{\alpha i} |\nu_i\rangle. \tag{1.12}$$

It should be noted, however, that the transition from 1.9 to 1.12 is not trivial. A discussion on the justification of this writing is well exposed in [32, p.246]. Once equation 1.12 written, the initial state of the neutrino will evolve through time, and this can be done easily in the Schrödinger representation, where a massive neutrino state $|\nu_k\rangle$ of mass m_k evolves as $|\nu_k(t)\rangle = \exp^{-iE_k t/\hbar} |\nu_k(t=0)\rangle$ (a time plane-wave), E_k being the eigenvalue of the Hamiltonian (with the usual relation $E_k = \sqrt{p_k^2 + m_k^2}$). The oscillation probability of the flavour α into the flavour β is then given by $p(\nu_\alpha \rightarrow \nu_\beta) = |\langle \nu_\beta | \psi(t) \rangle|^2$.

In the computation, it is assumed that the neutrino is ultra-relativistic, which means one can approximate $t \approx L$, sometimes called the light-ray approximation (L being the distance of flight of the neutrino, very relevant for baseline experiments). Also we can approximate E_k as:

$$E_k \approx E(1 + \frac{m_k^2}{2E^2}). \tag{1.13}$$

If we state the neutrino of flavour α is created in a weak process, it is assigned, by energy-momentum conservation, a definite energy E and definite momentum $\overrightarrow{p}$ (such that $||\overrightarrow{p}|| = E$). This is one place where the computation is not correct. The emitted neutrino can not be attributed a given energy E and momentum $\overrightarrow{p}$ in the creation process and at the same time be written as a superposition of the massive states, each having its own mass m_k. Energy-momentum conservation can not be satisfied for all three massive neutrino states, since they have different masses. To solve this problem, it is necessary to describe the emitted neutrino as a wave-packet in energy, thus allowing the superposition of different energy states. This point is discussed in [32, p.283]. The exposed formalism of standard neutrino oscillations, based on incorrect assumptions, leads however to the relevant oscillation probabilities for all baseline experiments. Probabilities of oscillation are often driven by terms like $\sin\left(\frac{\Delta m^2 L}{4E}\right)$, where Δm^2

is the difference of square neutrino masses at play, L the length of the baseline and E the neutrino energy. The final result of the standard derivation of neutrino oscillations reaches in vacuum:

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta)(L,E) &= \delta_{\alpha\beta} - 4\sum_{k>j} \Re\left[U^\star_{\alpha k} U_{\beta k} U_{\alpha j} U^\star_{\beta j}\right] \sin^2\left(\frac{\Delta m^2_{kj} L}{4E}\right) \\ &+ 2\sum_{k>j} \Im\left[U^\star_{\alpha k} U_{\beta k} U_{\alpha j} U^\star_{\beta j}\right] \sin\left(\frac{\Delta m^2_{kj} L}{2E}\right), \end{aligned} \tag{1.14}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta, and $\Re$ and $\Im$ are the real and imaginary functions. In the case of antineutrinos, one only needs to change the sign of the last line of the equation. These previous computations are made using the natural units of particle physics, but it is often convenient to write this result with relevant units of baseline experiments: kilometers for L, GeV for E and eV^2/c^4 for Δm^2. One needs to multiply by $c^3/\hbar$ to recover the international system of units. The relevant term then reaches:

$$\frac{\Delta m^2 L}{4E}\,[\text{natural units}] = \frac{1.27 \Delta m^2[\mathrm{eV}^2/c^4] L[\mathrm{km}]}{E[\mathrm{GeV}]}. \tag{1.15}$$

In the two neutrino approximation, where the PMNS matrix reduces to a real rotation matrix with a single mixing angle θ and single $\Delta m^2 = m_2^2 - m_1^2$, equation 1.14 simplifies to:

$$P(\nu_\alpha \to \nu_\beta) = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \quad \text{for } \alpha \neq \beta \tag{1.16}$$

In long-baseline experiments, such as DUNE, neutrinos travel a large distance through the Earth. In this case the presence of ordinary matter can affect the oscillation probabilities. We give the main features of this phenomenon in the next section.

1.2.3 Matter effects

The expressions for the neutrino oscillations probability given in the previous section correspond the case where neutrinos propagate in vacuum. Wolfenstein discovered in 1978 [38] that neutrinos propagating in matter are affected by a potential which modifies the flavour evolution of the neutrinos, *i.e* neutrino oscillations. He took into account the forward coherent scattering of neutrinos on matter, for which the neutrino interacts with a particle of the medium (electron, neutron or proton) without changing the quantum state of the medium. All neutrino flavours can interact in such a manner via the neutral currents. The electron neutrino (and antineutrino) can in addition make a charged current interaction on an electron $\nu_e + e^- \to e^- + \nu_e$, requiring the electron of the final state has the same 4-momentum and helicity as the electron of the initial state. For a medium of low temperature and density, the effective potential seen by a neutrino of flavour α reaches [32, Chap. 9]:

$$V_\alpha = \sqrt{2} G_F \left(N_e \delta_{\alpha e} - \frac{1}{2} N_n\right), \tag{1.17}$$

where G_F is the Fermi constant, N_e the electron density, N_n the neutron density and $\delta_{\alpha e}$ the Kronecker symbol. The first term is the charged current component (non-null only for the electron flavour), the second is the neutral current contribution, in which the electron and

proton contributions canceled as a result of electrical neutrality of the medium. In fact, it can be shown for ultrarelativistic neutrinos that V_α is the potential felt by the left-handed neutrinos, while the right handed neutrinos feels a potential proportional to $\propto V_\alpha m_\alpha/E^2$, where m_α can be though as an effective mass for the neutrino of flavour α and E is its energy.

Let's use the notation $\psi_{\alpha e} = \langle \nu_e | \nu_\alpha \rangle$, where $\psi_{\alpha e}$ designates the amplitude of probability that the evolving neutrino state $|\nu_\alpha\rangle$ is measured at a given position in space as an electron neutrino. Let us define in the same way $\psi_{\alpha\mu}$ and $\psi_{\alpha\tau}$. In the Schrödinger picture, the evolution of the probability amplitudes is driven by:

$$i\frac{d\Psi_\alpha}{dx} = \frac{1}{2E}(\, U\, \mathbb{M}^2\, U^\dagger + 2E\, \mathbb{V}\,)\, \Psi_\alpha \tag{1.18}$$

$$\Psi_\alpha = \begin{pmatrix} \psi_{\alpha e} \\ \psi_{\alpha\mu} \\ \psi_{\alpha\tau} \end{pmatrix}, \qquad \mathbb{M}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Delta m^2_{21} & 0 \\ 0 & 0 & \Delta m^2_{31} \end{pmatrix}, \qquad \mathbb{V} = \begin{pmatrix} V_{CC} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.19}$$

where $V_{CC} = \sqrt{2} G_F N_e$ is the charged current term of the V_α potential defined in 1.17. When $V_{CC} \to 0$, one recovers the vacuum equation that leads to vacuum neutrino oscillations. It should be noted that the evolution is driven with the spatial variable x and not the time t because of the ultra-relativistic assumption which allows to write $t \approx x$.

The two neutrino case got particular interest since Wolfenstein shed light on the fact that it allowed for large neutrino mixing in matter even with small neutrino mixing in vacuum. Indeed, in the two neutrino case equation 1.19 equation simplifies (getting rid of terms leading to global phase factors) to a vacuum-like equation with an effective mixing matrix:

$$U_M = \begin{pmatrix} \cos(\theta_M) & \sin(\theta_M) \\ -\sin(\theta_M) & \cos(\theta_M) \end{pmatrix}, \quad \tan(2\theta_M) = \frac{\tan(2\theta)}{1 - \dfrac{N_e}{N_e^R}}, \tag{1.20}$$

where θ is the mixing angle between the two neutrino mass and flavour states and Δm^2 the square mass differences between the two mass states. U_M is the effective mixing matrix and θ_M the effective mixing angle in matter. The equation above has an very interesting feature: a resonance (*i.e* maximal mixing $\theta_M = \pi/4$) if the electron density reaches $N_e = N_e^R = \dfrac{\Delta m^2 \cos(2\theta)}{2\sqrt{2} E G_F}$.

In a medium with constant density, the $\nu_e\nu_\mu$ transition probability is given by:

$$P(\nu_\mu \to \nu_e)(L, E) = |\psi_{e\mu}|^2 = \sin^2(2\theta_M) \sin^2\left(\frac{\Delta m^2_M L}{4E}\right). \tag{1.21}$$

We indeed recover the formula 1.16 where the vacuum mixing angle and mass square difference have been replaced by their effective equivalent in matter.

This theoretical observation led to the so-called MSW (Mikheïev-Smirnov-Wolfenstein) effect, and got special interest because of its possibility to explain the solar neutrino problem [39]. The key point is that if the Solar core has a large enough density, the emitted solar electron neutrinos will travel as the effective neutrino mass state $\approx \nu_2^M$, assuming the two neutrino approximation. One can show that if the neutrino propagates adiabatically (*i.e* the density variations are small

along the trajectory path), there is no transition possible between the two effective mass states, and thus the neutrino will escape the Sun as ν_2. This mechanism can thus explain the electron neutrino depletion observed on Earth, as for a low vacuum mixing angle ν_2 contains a large fraction of ν_μ.

From the discussion we had above, one can see that only the charged current term of V_α (see equation 1.17) contributed to the effective oscillation, because the neutral current term only resulted in a global phase to the three amplitudes of probability $\psi_{\alpha\beta}$ with $\beta = e$, μ, τ, and thus did not impact the probabilities themselves.

Matter effects for antineutrinos can be obtained by simply switching the sign of V_{CC} [40]). This can be understood in a rather simple picture. The charged current diagram contributing to the coherent scattering used to compute the effective potential felt by the neutrino in matter is given by the left diagram of the following picture.

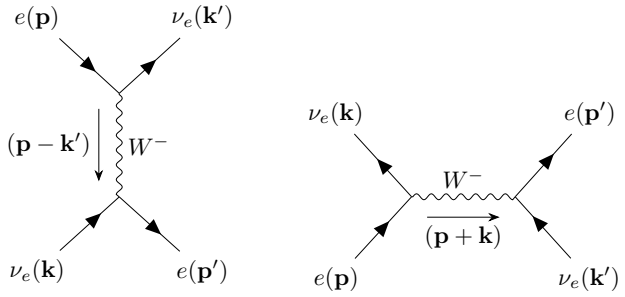


Figure 1.3: Charged current tree diagram of electron (anti)neutrino on the left (right), showing the coherent forward scattering of electron neutrinos on matter, used to account for matter effects in oscillations.

It describes the electron neutrino scattering on an electron (t channel). We can derive the s diagram of this t diagram via an exchange of the two neutrino lines (in order to keep momenta $\mathbf{k}$ and $\mathbf{p}$ as the initial momenta). We then obtain the right diagram of figure 1.3, which describes the annihilation of an electron antineutrino with an electron. The coherent scattering further implies $\mathbf{p}' = \mathbf{p}$ and $\mathbf{k}' = \mathbf{k}$. The matrix element calculation of the s diagram is the same as the t diagram, only requiring the change $\mathbf{k} \to -\mathbf{k}$ (which does not impact the potential derivation in our case). However since we exchanged two fermions building the s diagram from the t diagram, we need to add a -1 factor, which results in the change $V_{CC} \to -V_{CC}$.

1.3 The DUNE experiment

DUNE (Deep Underground Neutrino Experiment) is a major forthcoming long-baseline neutrino experiment, planning to start data taking by the end of this decade. The experiment will take place on two main sites in the USA: the Fermi National Accelerator Laboratory (Fermilab, Illinois) and the Sanford Underground Research Facility (SURF, South Dakota). We will discuss the purpose of the experiment and describe its design.

1.3.1 Scientific goals

Neutrino oscillations is the core subject which motivates the DUNE experiment [41, 42, 43]. The main goal of the experiment is the establishment of the possible charge-parity (CP) symmetry violation occurring in the lepton sector that can be probed via the neutrino oscillations phenomenon. Given the current knowledge of the PMNS (1.2.2) parameters (all but δ_{CP} have already been measured), the non-violation hypothesis dwells on the value of the parameter $\delta_{CP} = 0 \,\text{or}\, \pi$. Any other value for this parameter implies a CP symmetry violation in the neutrino oscillations. T2K recently published that a maximal violation at $\delta_{CP} \approx -\pi/2$ is favoured [44] while NOvA favours the region of $\delta_{CP} = \pi/2$ assuming a normal ordering [45]. The CP-violation in the leptonic sector is of great interest in cosmology. Neutrinos could be the key to understanding the underlying mechanism for the creation of the matter-dominated universe.

One can prove starting from equation 1.14 that the difference in the oscillation behaviour in vacuum between neutrinos and antineutrinos depends on $\sin(\delta_{CP})$. Defining $\Delta P_{\alpha\beta}$ as the difference of oscillation probabilities from flavour α to β between neutrinos and antineutrinos, we have:

$$\begin{aligned} \Delta P_{\alpha\beta}(L,E) &= P(\nu_\alpha \to \nu_\beta) - P(\bar{\nu}_\alpha \to \bar{\nu}_\beta) && (1.22) \\ &= -4J\left[\sin\left(\frac{\Delta m^2_{21} L}{2E}\right) + \sin\left(\frac{\Delta m^2_{32} L}{2E}\right) + \sin\left(\frac{\Delta m^2_{13} L}{2E}\right)\right] && (1.23) \\ J &= c_{12}c^2_{13}c_{23}s_{12}s_{13}s_{23}\sin(\delta_{CP}), && (1.24) \end{aligned}$$

where we have used the notation $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The quantity J, sometimes referred as the Jarlskog invariant, depends only on the parameters of the PMNS mixing matrix. The mixing angles have been measured to be not 0 mod π and not $\pi/2$ mod π, so that none of the c_{ij} and s_{ij} vanishes. The only way J can be 0 is if $\delta_{CP} = 0$ or π. This is why excluding these values would rime with discovering the violation of CP symmetry in the neutrino sector. DUNE will use the oscillation channel $\nu_\mu \to \nu_e$ to measure the δ_{CP} parameter by comparing the oscillations of neutrinos and antineutrinos.

In addition to the CP violation research program, DUNE will allow studying several other aspects of fundamental physics:

- As a large scale experiment with a total detector mass of $\approx 40\,000\,\text{t}$, the far detectors will act as a neutrino telescope in case of supernova burst inside the Milky Way. This would provide unique information on the collapse of the supernova. More discussion can be found in [46].
- The proton decay is not predicted by the Standard Model but only by theories beyond the Standard Model such as Grand Unification theories. DUNE will allow for the search of events like $p^+ \to K^+ + \bar{\nu}$.
- DUNE will aim at constraining the mixing angles of the PMNS matrix, especially the angle θ_{23}. Current fits allow θ_{23} to vary in a region below and above 45° [47], a problem referred as the octant problem or octant degeneracy. The reason is that the octant indeterminacy leads the $\nu_\mu \to \nu_e$ oscillations to be insensitive to some region of other oscillation parameters like δ_{CP}. See for instance a discussion in [48].

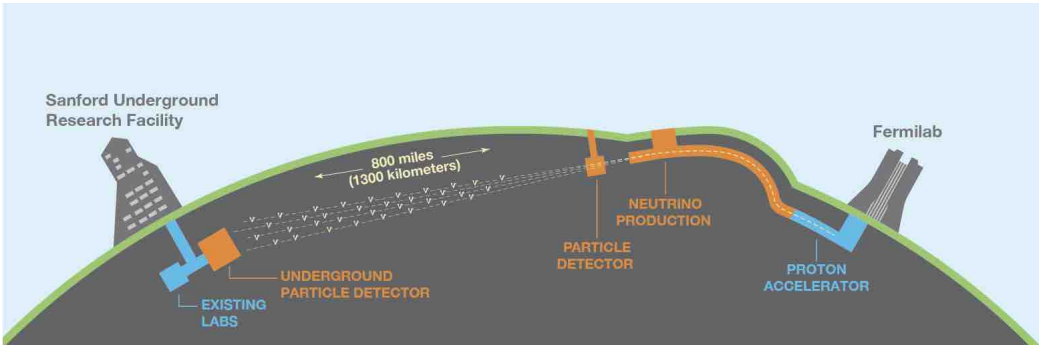


Figure 1.4: A schematic view of the long baseline neutrino facility [43]. From right to left: the Fermilab (Illinois) producing the neutrino beam and hosting the near detector (ND), and the Sanford Underground Research Facility (South Dakota) hosting the far detectors (FD).

- More generally, DUNE will bring information on the consistency of the 3-flavours oscillation mechanism. A significant number of τ neutrino events are also expected to occur during the experiment via the $\nu_\mu \to \nu_\tau$ oscillation channel. This is the main topic of this thesis and we shall review this aspect more in details later in this chapter (see 1.4).

- An other issue at play about neutrinos is the so-called mass ordering or mass hierarchy. The observations of neutrino oscillations implied that neutrinos were massive and that at least two of them were almost degenerate masses (in fact, all three are). The ordering $m_1 < m_2$ has already been established, but the total ordering could either be $m_1 < m_2 < m_3$ (Normal Ordering "NO") or $m_3 < m_1 < m_2$ (Inverted Ordering "IO"). DUNE expects to measure the ordering at 5σ confidence level after few years of data taking. A global analysis of 2017 favoured normal ordering at 2σ [49] and a more recent one favours NO at 1.6σ (which increases to 2.7σ if atmospheric data from Super-Kamiokande are included) [50].

1.3.2 Conceptual design of the DUNE experiment

As other long-baseline experiments, the DUNE experiment will rely on three main points:

- The production of a high intensity neutrino beam, hosted at Fermilab. The power of the proton beam used to produce the pions, which then decay into neutrinos plus charged leptons, is expected to be 1.2 MW. A power upgrade up to 2.4 MW is considered after several years of running. The energy of the neutrinos will cover a band from hundreds of MeV to several GeV. In connection with the flight distance (1285 km), the values of the L/E parameter presented before will allow studying precisely neutrino oscillations, close to the first probability maximum of oscillations in atmospheric neutrinos. This neutrino beam will be the world's most intense one.

- The design of a near detector, close to the production point (Fermilab) in order to characterise the neutrino beam composition (the beam will be essentially made of muon neutrinos with some electron neutrinos contamination), necessary to reduce the systematics in relation to the neutrino flux and neutrino cross sections.

- The far detectors design, ≈ 1300 km far from the neutrino production point. It will consist of four 10 kt liquid argon detectors located at the Sanford research facility. A single module is shown in figure 1.5.

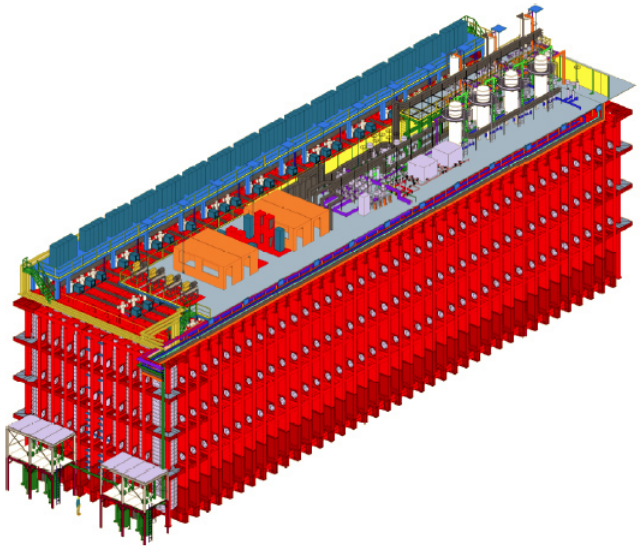


Figure 1.5: View of one DUNE far detector module. A human person is standing in front of it to show the scaling.

A schematic view of these three points is proposed on picture 1.4. It should be noted that there is a vocabulary distinction one has to make. DUNE will consist in both the far and the near detectors, which will make it possible to study several aspects of physics: neutrino oscillations, burst of neutrinos emitted by supernovae, proton decay. One will employ the term LBNF (Long Baseline Neutrino Facility) to refer to all the facilities at stake to get the detectors operating (the cavern at Homestake mine, the cryogenic technology to cool down liquid argon, the neutrino beam itself...).

The correct and successful deployment of the three points listed above is crucial for the success of the experiment itself. However, not much will be discussed about the neutrino beam facility and the near detectors. The purpose of this thesis mostly focuses on the physics at the far detectors. The technology chosen for the far detectors is the Liquid Argon Time Projection Chamber (LArTPC). More details on the basic detection principles of this technology are provided in the next section.

1.3.3 The Liquid Argon Time Projection Chamber (LArTPC)

The idea of using the LArTPC technology in neutrino physics was originally mentioned by Carlo Rubbia in a CERN internal report in 1977 [51] (note also the work of William Willis and Veljko Radeka on liquid argon calorimeter [52]). The technology was thought to allow combining large target mass with a precise calorimetric performance, (individual tracking and energy reconstruction of particles passing through the detector) at the time insured by bubble chamber experiments. Liquid argon has the advantage of being cheap, and presents similar properties as the freon CF_3Br, already used by the Gargamelle experiment in the 70's, allowing in addition for a completely electronic readout.

Intensive R&D work have been performed to characterize and get familiar with the technology. Elena Aprile, Karl-Ludwig Giboni and Carlo Rubbia studied a 2 l prototype in 1985 [53]. The ICARUS collaboration especially considerably worked at building and operating ever growing detectors (3 t [54], 14 t [55] and 600 t [56]). The collaboration also exposed a 50 l prototype to the WANF (West Area Neutrino Facility) neutrino beam at CERN in 1997 [57] to demonstrate

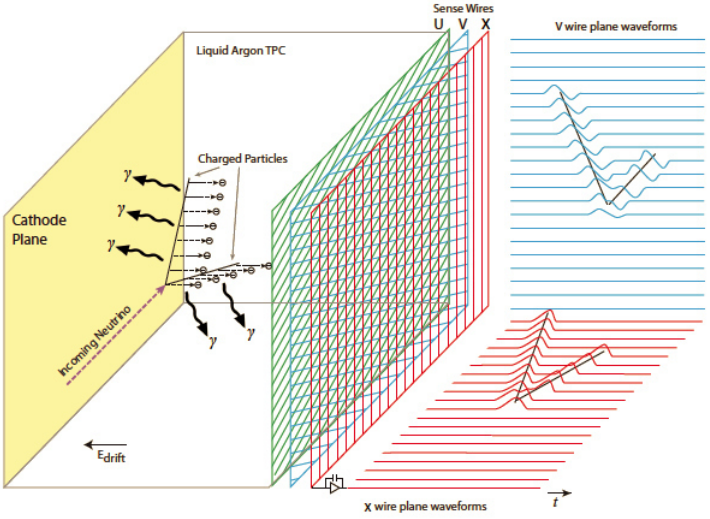


Figure 1.6: Schematic view of a single-phase LArTPC. The sense of drift of electrons is shown (left to right). Electrons resulting from the ionisation process are drifted toward the anode and induce signals on a set of wire planes making it possible to visualise the charged particle that propagated within the detector.

the feasibility of doing neutrino physics with such detectors. They gathered a sample of 86 clean ν_μ interaction for which they measured the event kinematics. The detector also proved to achieve particle identification via the energy loss behavior as a function of the range.

LArTPC working principle

Liquid Argon Time Projection Chambers (LArTPCs) can be thought as electronic bubble chambers. The idea is to fill a cryogenic tank with liquid argon to be used as the neutrino target. Very massive detectors are needed in order to record enough neutrino interactions for neutrino oscillations studies. Liquid argon has a density of $1.4\,\mathrm{g\,cm^{-3}}$, and it can be used in order to achieve target masses of thousands of tons. Furthermore, liquid argon allows detecting very efficiently the ionisation of charged particles produced in the final state of neutrino interactions. A charged particle will indeed leave on its trajectory in the liquid argon volume a series of ion argon-electron Ar^+-e^- pairs. The work function of ionization is about 20 eV and particles at play in the detector have typically energy losses of $2\,\mathrm{MeV\,cm^{-1}}$ of crossed path, producing several tens of thousands pairs over such a distance. This ionization process produces light, and some of the Ar^+-e^- pairs quickly recombine, producing an additional prompt emission of light (128 nm) detected by photo-multipliers located inside the detector, that provides information of the referent time of the interaction.

The objective is to retrieve the cloud of free electrons, making it possible to picture the trajectory of the charged particle. The ionization electrons are drifted towards the anode plane by a strong electric field (typically $0.5\,\mathrm{kV\,cm^{-1}}$) until they reach the anode. Figure 1.6 shows a schematic view of a single LArTPC.

The ionization electrons are detected via the currents they induce in the anode. The anode consists in two planes of oblique induction wires, with a readout pitch of 4.7 mm, and one plane of collection wires. The induction wires are orientated at $\pm 35.7°$ with respect to the vertical collection wires. The signal given by electrons on a single wire at a given time is reconstructed in a so-called hit: this information is the basis of any signal. Note that such detectors do not see an event in three dimensions, but three two-dimensional projections of the same event for

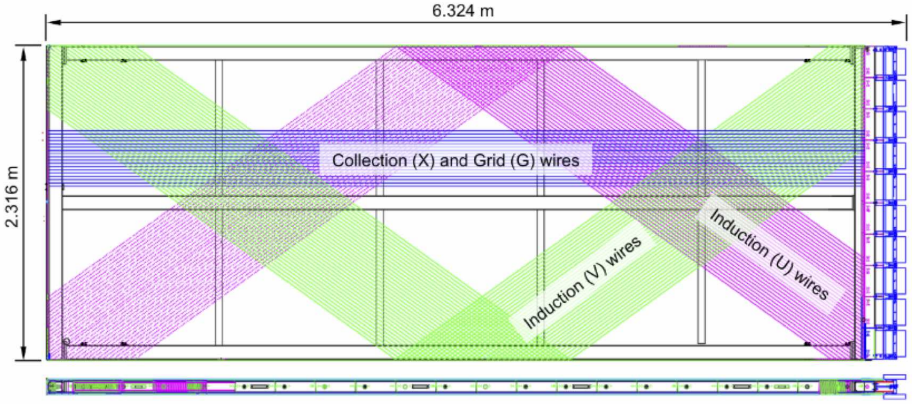


Figure 1.7: APA (Anode Plane Assembly) view [41, Vol. 2 p2-12]. The plane is composed of one horizontal (on the picture, in real life it is vertically orientated) collection plane and two oblique induction planes (green and pink).

the single phase case. By using the time of the drift which is equivalent to the last spatial coordinate (electrons drift velocity in argon is known), it is possible to reconstruct a three-dimensional event.

The acquisition of an event starts right after the ionisation created by a particle going through the detector. The time of the acquisition is defined as the time needed by an electron close to the cathode (the liquid argon beyond the cathode is not taken into account in the detection volume of the detector) to drift to the anode. This time depends on the electric field and properties of liquid argon. For instance, this time window acquisition is 4 ms for a drift of 6 m. An event is then defined as a set of hits collected by the wires in the different views.

1.3.4 LArTPC calorimetry performance

The strength of a neutrino detector based on the LArTPC technology is its ability to give a detailed view of the final state of a neutrino event, that is to say the particles emerging from the neutrino interaction with an argon atom. The detector provides proper spatial (few millimeters) and energy resolution in order to make it possible to reconstruct the kinematics of the reaction (including the impinging neutrino energy), a key element to measure the oscillation probabilities, and allow the identification of the particles in the final state. The final state may consist in a tree of secondary particles (called energy flow) which have to be individually identified, measured and connected to each other. The energy flow measurement allows reconstructing the energy of the impinging neutrino.

The word hermiticity is often used to describe detectors in particle physics. An ideal detector has a perfect hermiticity to particles, and can reconstruct/identify every particle propagating in it. Liquid argon was indeed remarked as a good detector medium by Willis and Radeka [52] because of its ability to absorb particle's energy. We should here refer to the intense work of the ICARUS collaboration which demonstrated the feasibility to reach large scale LArTPC detectors, more specifically to the T600 prototype for which they demonstrated the ability to measure precisely long-ionizing tracks (typically muons) [58] and electromagnetic showers [59] (see also the work of MicroBooNE with neutral pion decays into two photons [60]). ICARUS presented results on the muon decay spectrum [61] and MicroBooNE presented the corresponding Michele electron reconstruction [62]. Note also the work of the ArgoNeuT collaboration which demonstrated

the feasibility to extend the energy range of LArTPC detectors down to the MeV scale [63], studying low energy photons produced in neutrino interactions.

These results indicate that the LArTPC technology is mature and ready for larger scale neutrino experiments like DUNE. However achieving the final design and calibration of a 10 kt argon detector is challenging: the size of a single module detector will be 18 m high, 19 m wide and 66 m long. Moreover there exist two different designs of the LArTPC technology that were proposed for the DUNE far detector modules.

1.3.5 Single-Phase and Dual-Phase LArTPC

Two different versions of LArTPCs were originally proposed for DUNE: single-phase and dual-phase. In the traditional single-phase (which is the technology developed since 1977 and the one we previously exposed), the ionization electrons drift horizontally and they are collected by wire-planes. The dual-phase technology takes advantage of the argon gas phase in order to amplify the ionization signals and to build detectors with a longer vertical drift distance, up to 12 m. This configuration allows to extract the electrons from the liquid to the gas, where the signal losses related to the impurities in the liquid phase are compensated by the gain achieved in electron avalanches which are only possible in gas. The dual-phase, due to its geometrical layout and detector components to collect the electrons at the anode (printed circuit boards used for both the tasks of electrons amplification and collection) is cheaper and faster to build and install compared to the more traditional single-phase technology.

The DUNE collaboration has been recently focusing on an evolution of the dual-phase design, the vertical drift (VD) LArTPC detector. It takes advantage of the experience acquired with the dual-phase technology. In the Vertical Drift the anodes to collect the electrons are still made of printed circuit boards. These are however simpler than the ones designed in dual-phase to achieve both the amplification and the collection of the electrons in the gas layer above liquid argon. For the Vertical Drift the amplification stage in the gas is suppressed and the perforated anode printed circuit boards can directly operate immersed in liquid argon in order to collect the drifted electrons. This simplification is supported by the very good liquid purity records achieved in the prototypes, which decrease the need for amplification. The Vertical Drift exploits most of the detector elements developed for the dual-phase design and does not need to pursue further R&D activities on the amplification stage in gas. It then allows building on a faster time-scale and in a cheaper way the second DUNE far detector module.

Before reaching the final far detector scale, the DUNE collaboration decided to build several intermediate detector prototypes [64, 65, 66] of variable size in order to get more familiar with the technology at stake and prepare the analysis for the real experiment, mostly the systematics (neutrino event contamination, misidentification of particles, uncertainties on the reconstructed energy of the impinging neutrino, hadronic shower reconstruction...).

1.3.6 Development of the DUNE far detector design and prototypes

Intermediate detectors, progressively of larger dimensions, were built by the DUNE collaboration in order to get more familiar with the technology before reaching large scale volume of ≈ 10 kt. These detectors have been developed at the CERN Neutrino Platform to prototype key elements in the DUNE far detector design. Detectors are named by their effective size, that is to say the volume of liquid argon really used for measurements.

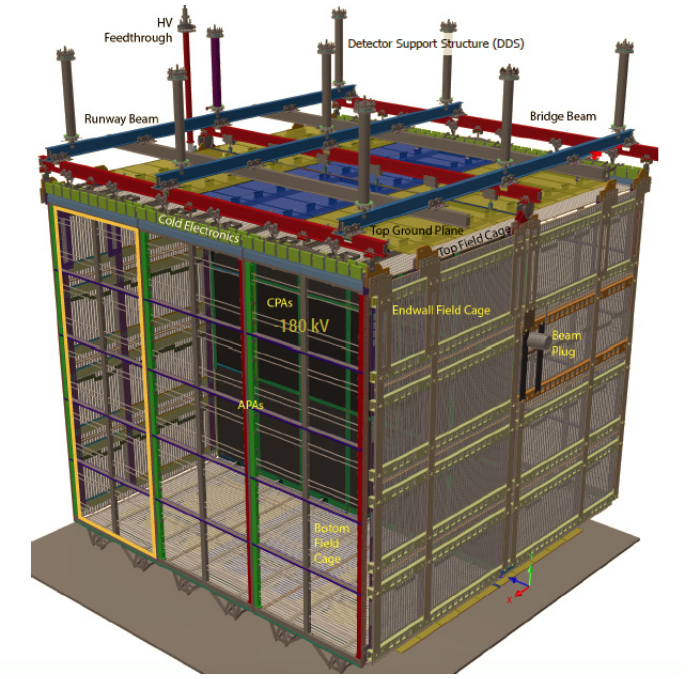


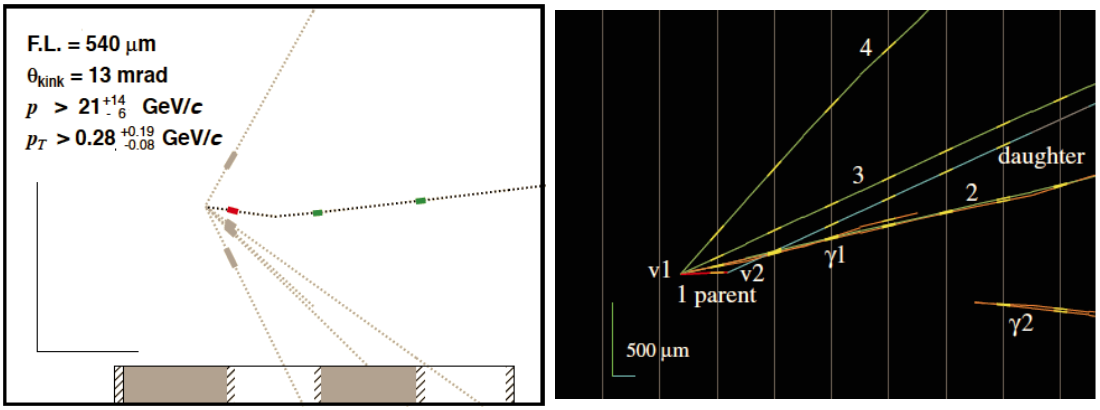
Figure 1.8: Schematic picture of the single-phase LArTPC detector installed at CERN. The volume of argon is actually the small square cage at the center of the whole structure.

A first dual-phase detector, the $3 \times 1 \times 1\,\mathrm{m}^3$, was built at CERN [66, 67]. As suggested by its name, its active volume is a parallelepiped of dimensions $3 \times 1 \times 1\,\mathrm{m}^3$. Note that this volume dedicated to the detection does not include the whole facility surrounding this volume detection of argon. Installed at CERN in 2016, on surface, the detector has been exposed to cosmic rays for less than a year. Then in 2018-9 followed two detector modules, the so-called protoDUNEs, one using the single-phase technology and one using the dual phase-technology on an active volume of $6 \times 6 \times 6\,\mathrm{m}^3$. A schematic view of the single-phase prototype is shown in figure 1.8. One purpose of these detectors was to learn about hadronic interactions in liquid argon and calorimetry performance of the technology in order to reduce the systematics for DUNE, a work which has already started [68].

The availability of these detectors started the experimental analysis for DUNE. At a prototyping scale, a lot of work has to be done to calibrate the detectors response, to get more familiar for the technology at play before the large scaling required at the DUNE FD, and to better understand the hadrons behavior in matter.

1.4 τ neutrino physics at DUNE and elsewhere

In 1.3.1, we have presented the main scientific program of the DUNE experiment, and then dedicated time to briefly present the state of the art of the LArTPC technology which will be used for the far detectors of the experiment. In the year 2020, the DUNE Technical Design Report was achieved in four volumes, demonstrating the capability of the experiment to reach its main objectives [69, 41, 70, 71]. Because the ν_τ study was not originally present in the first definition of the scientific program of DUNE, not much is said about it in the TDR. This thesis proposes to explore the capability of DUNE to study the $\nu_\mu \to \nu_\tau$ oscillation channel of the neutrino beam. We will briefly review the scientific interest of such a study.



(a) (b)

Figure 1.9: Left: one of the ν_τCC candidate published by the DONUT collaboration in [9]. Its scale has been exaggerated to better illustrate the visible kink. The vertical line represents 0.1 mm, while the horizontal line represents 1.0 mm. Right: one of the ν_τ CC candidate published by the OPERA collaboration in [78]. In both cases, the impinging τ neutrino comes from the left of the picture. The visible tracks represent the observed final state particles. One of them shows a kink, indicating a decay process compatible with a τ decay.

1.4.1 Current status of the τ neutrino measurements

The data available in the scientific literature on the τ neutrino are poor. The DoNuT experiment first published a total of 9 candidates in 2008 [72], followed by OPERA in 2018 (10 candidates) [73] using the $\nu_\mu \rightarrow \nu_\tau$ oscillation channel. Super-Kamiokande [74, 75, 76] proposed a method using a combination of neural network and azimutal angle information to test the ν_τ appearance hypothesis in the total sample of atmospheric neutrinos. They excluded the no-ν_τ appearance hypothesis with a significance of 4.6σ (assuming normal hierarchy, more than 5σ with inverted hierarchy). It was found to correspond to a total 338.1 ± 72.7 atmospheric τ neutrino events over the runtime of the experiment. IceCUBE also proposed a similar analysis and rejected the no-ν_τ appearance hypothesis at 3.2σ [77].

The τ neutrino is thus the least known particle of the Standard Model. DUNE will offer a unique opportunity to study this specific flavour, with an expected event rate of $\mathcal{O}(100-1000)$ $\nu_\mu \rightarrow \nu_\tau$ CC events during the run time of the experiment. However, this specific neutrino flavour is quite tricky to recognize. The identification of neutrino flavour is made possible thanks to weak charged current interactions with ordinary matter (see 1.1.2), in which the neutrino transforms into its corresponding charged lepton. Indeed, assume one observes a neutrino event with an electron in the final state, one deduces the neutrino flavour was "electron". In LArTPC, an electron will trigger an electromagnetic shower, quite easy to recognize. Muons leave long and straight tracks, also easy to recognize. Further details will be given on these points in the next chapter (see 2.5.2). The charged lepton τ, however, decays too promptly (we will detail this point in 2.1.6) to allow for a direct identification. OPERA and DONuT used detectors with very precise (≤ 1 mm) spatial resolution to be able to directly observe a kink due to the τ decay. Illustrations of published candidates are shown in figure 1.9 [9, 78].

1.4.2 Physics motivations for ν_τ search

The τ sector of the neutrino is still little explored, and a growing interest can be found in the literature concerning the physics to be explored with the τ flavour appearance [79, 80] in

DUNE. Up to now, these studies have been conducted assuming given signal selection efficiencies. However there is a lack of assessment of the capability of DUNE to identify the τ neutrino flavour. It is necessary to provide credentials to support these studies and drive future ones. While this is the purpose of this thesis to provide such data, we propose in a first place to review the main physics potentials associated to the τ neutrino appearance in DUNE.

Cross section measurement:

OPERA [73], Super-Kamiokande [76] and IceCube [77] reported the flux averaged ν_τ cross section measurement which appeared to be in agreement with the Standard Model expectations. It is worth to further explore this domain for a better understanding of neutrino interactions. The charged current double differential τ neutrino cross section, expressed in terms of the Bjorken variables x and y (see next chapter equation 2.5, for the present discussion it is not necessary to go through these concepts) typically reaches [81]:

$$\begin{aligned}\frac{d^2\sigma^{\nu(\bar{\nu})}}{dxdy} = \frac{G_F^2 M E_\nu}{\pi(1+Q^2/M_W^2)^2}&\left(y^2 x + \frac{m_\tau^2 y}{2E_\nu M}\right) F_1 \\ &+ \left[\left(1 - \frac{m_\tau^2}{4E_\nu^2}\right) - \left(1 + \frac{Mx}{2E_\nu} y\right)\right] F_2 \\ &\pm \left[xy\left(1 - \frac{y}{2}\right) - \frac{m_\tau^2 y}{4E_\nu M}\right] F_3 \\ &+ \frac{m_\tau^2(m_\tau^2 + Q^2)}{4E_\nu^2 M^2 x} F_4 - \frac{m_\tau^2}{E_\nu M} F_5.\end{aligned} \tag{1.25}$$

In this formula, M is the target mass, E_ν the impinging neutrino energy, m_τ is the mass of the charged τ lepton, M_W the mass of the W boson, G_F is the Fermi constant. x, y, Q^2 are standard kinematical variables commonly used to describe scattering processes and they will be defined in the next chapter (see 2.1.2). The F_i are structure functions. At this point, we emphasize that the last line of equation 1.25 contains factors m_τ^2/M and m_τ^2/M^2. For the ν_e (ν_μ) cross sections, the mass of the lepton τ is replaced by the one of the electron (muon). The factor m_l/M with $l = e, \mu$ reaches 5×10^{-4} for the electron and 0.1 for the muon. Thus for GeV scatterings these terms are systematically neglected, which means the structure functions F_4 and F_5 are omitted. However for the ν_τ case these structure function contribute, in a way that can bias the ν_τ expected number of events.

3-flavour phenomenology:

The $\nu_\mu \to \nu_\tau$ oscillation channel is sensitive to the θ_{23} mixing angle and to the atmospheric mass splitting Δm^2_{31}. In vacuum the oscillation probability can be approximated to:

$$P(\nu_\mu \to \nu_\tau) \approx 4|U_{\mu 3}|^2|U_{\tau 3}|^2 \sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right) = \cos^4(\theta_{13}) \sin^2(2\theta_{32}) \sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right), \tag{1.26}$$

where L and E are respectively the length of the baseline and the neutrino energy. However it has been pointed out that this oscillation channel would not help significantly in constraining the θ_{13} and θ_{23} mixing angles [79, 80], for which there exists a stringent complementarity

between the $\nu_\mu \to \nu_e$ appearance and $\nu_\mu \to \nu_\mu$ disappearance channel. However the ν_τ and $\bar{\nu}_\tau$ appearance channels were proposed [82] to help at resolving the mass hierarchy, and it is always worth to perform a complementary check on the oscillation parameters if only for reinforcing the robustness of the 3-flavour neutrino paradigm.

The 3+1 neutrino paradigm

The $\nu_\mu \to \nu_\tau$ oscillation channel shows interest in testing the so-called 3+1 neutrino paradigm, where a fourth massive neutrino state is added, and the corresponding fourth weak neutrino state is sterile with respect to bosons Z and W. This process also requires to add 3 new mixing angles θ_{14}, θ_{24}, θ_{34} as well as two more phases. The extended PMNS matrix now is of size 4.

This paradigm is currently under investigation by short-baseline experiments, typically taking advantage of nuclear reactor as sources of electron antineutrinos. They study the $\bar{\nu}_e$ disappearance as a function of the distance to the detector and the neutrino energy. As for illustration, a recent publication by the Neutrino-4 collaboration can be found here [83], in which they observed the oscillation phenomenon (*i.e* $\bar{\nu}_e$ disappearance) at 2.4σ significance with oscillation parameters $\Delta m^2_{14} = 7.20 \pm 1.13\,\mathrm{eV}^2/c^4$ and $\sin^2 2\theta_{14} = 0.29 \pm 0.12$. Global analysis on these type of experiments show a deficit in electron antineutrinos not enough to be called "reactor anomaly", and more investigation is ongoing to elucidate this current discrepancy.

The use of the $\nu_\mu \to \nu_\tau$ appearance channel in addition to the two standard $\nu_\mu \to \nu_e$ and $\nu_\mu \to \nu_\mu$ oscillation channels is expected to help at constraining the 3+1 scenario. For more detailed studies, we again refer to [79, 80]. Note that for values of Δm^2_{14} greater than $0.1\,\mathrm{eV}^2/c^4$, ν_τ appearance at the DUNE near detector (*i.e* short baseline) is expected to bring a significant contribution as well.

PMNS matrix unitarity

The measurement of the unitarity of the CKM quark mixing matrix has been an intense and successful part of the scientific program of accelerator based particle physics. As far as the PMNS matrix is concerned, the unitarity test remains poor. Unitarity implies the relations:

$$\sum_{k=1,2,3} U_{\alpha i} U^\star_{\beta i} = \delta_{\alpha\beta} \qquad \text{and} \qquad \sum_{\rho=e,\mu,\tau} U^\star_{\rho k} U_{\rho l} = \delta_{kl}. \tag{1.27}$$

Each of the six vanishing relations can be represented in the complex plane and visualized as a so-called unitary triangle. Ref [79] has reported that DUNE will help constraining the relation $|U_{e1}|^2 + |U_{e2}|^2 + |U_{e3}|^2 = 1$. The non-unitarity hypothesis is conventionally expressed as an additional triangle matrix to re-parametrize the PMNS matrix:

$$U \to VU \qquad \text{with } V = \begin{pmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}. \tag{1.28}$$

The constraints on the α_{ij} are a direct measure of the PMNS unitarity. Here again, the combination of the three channels $\nu_\mu \to \nu_{\alpha=e,\mu,\tau}$ all together are expected to bring substantial constraints.

Non-standard interactions

Beyond the Standard Model (BSM) scenarii propose non-standard interactions (NSI) in the neutrino sector. They are typically described in the low energy region as a 4-fermion interaction Lagrangians which include non-diagonal terms in the leptonic current. Such non-standard processes are expected to be observable in the propagation of neutrinos in matter, and they require a modification of the potential matrix which we used in equation 1.19:

$$\mathbb{V} \rightarrow \mathbb{V}_{NSI} = V_{CC} \begin{pmatrix} 1+\epsilon_{ee} & \epsilon_{e\mu} & \epsilon_{e\tau} \\ \epsilon^{\star}_{e\mu} & \epsilon_{\mu\mu} & \epsilon_{\mu\tau} \\ \epsilon^{\star}_{e\tau} & \epsilon^{\star}_{\mu\tau} & \epsilon_{\tau\tau}, \end{pmatrix} \tag{1.29}$$

where the various $\epsilon_{\alpha\beta}$ parametrize the NSI Lagrangian. The $\nu_\mu \rightarrow \nu_\tau$ is expected to display a maximal sensitivity to $\epsilon_{\mu\tau}$. It is observed that the three oscillation channels have a stringent complementarity, though most of the time the $\nu_\mu \rightarrow \nu_\tau$ channel be less sensitive than the two others.

1.4.3 Beam ν_τ search at DUNE far detectors

The DUNE spatial resolution will not allow for a direct visual identification of the charged lepton τ (further details in next chapter, see 2.1.6). In this case, one needs to rely on kinematic criteria to distinguish the τ flavour from the others. This idea dates back to Albright and Shrock in 1979 [22, 23], and was for instance deployed by the NOMAD collaboration at the end of 90's, for τ neutrino discovery, without positive signal [27]. The basic idea is to take advantage of the presence of two undetected neutrinos in the final state if the τ decays leptonically. Indeed, if the τ decays into an electron, then the charged current interaction looks like $\nu_\tau + X \rightarrow \tau^- + Y \rightarrow e^- + \nu_\tau + \bar{\nu}_e + Y$, where X and Y stands for the initial target and final state hadronic part. If the initial neutrino direction is known like in beam experiments, then there should be a substantial invisible energy in the transverse plane of the ν_τCC interactions, which would not be observed for the corresponding ν_eCC interactions. In addition there should be correlations between the directions of the leptonic, hadronic and missing momenta in the transverse plane.

Additional kinematical variables may also play a non-negligible role for identifying τ neutrino interactions, but the key point is that kinematics in the transverse plane carry substantial information on the interacting neutrino flavour. We also emphasize that semi-leptonic τ decays represent about 64% of the total branching ratio, and that adapted analysis must be deployed for these decays. Indeed, the main background component are going to be the neutral current interactions in which there is one neutrino in the final state, and the semi-leptonic decays of the τ only bring one neutrino in the final state as well. As a consequence the transverse missing energy is expected to play a lesser role than for leptonic decays.

It has been long suggested that massive LArTPC based neutrino experiments like DUNE would have an unprecedented sensitivity to the τ flavour appearance both for beam [84, 85] (and atmospheric [86] sources). Some analysis must however still be deployed to exploit this ν_τ appearance sensitivity. Correctly identifying the ν_τ component of the DUNE neutrino beam is a non-trivial task and is the subject of this thesis. We shall follow the philosophy of the NOMAD experiment "1 τ decay mode = 1 dedicated analysis". It should be noted that the study of atmospheric τ neutrinos is beyond the scope of this thesis.

$\star$ $\star$ $\star$

Chapter 2

Tools and methods

Contents

The purpose of this chapter is to introduce the main key features to understand the neutrino interactions at the simulation level for DUNE. The collaboration relied on a neutrino event generator called GENIE [87] (GEnerate Neutrino Interaction Event), a Monte Carlo neutrino event generator. The generator can calculate the neutrino cross sections on a various set of target atoms, and then uses the incident neutrino energy distributions (and their flavours) to predict the neutrino events of an experiment at the individual level. This chapter first consists in an introduction to the concept of neutrino cross sections and presents the beam neutrino flux planned for DUNE, which are the two ingredients needed by GENIE. The default version of GENIE used in the following is v2.12.00 (except otherwise specified) which is the same as the one used in the DUNE Technical Design Report. Then a section is dedicated to the calculation of neutrino oscillations in matter. Combining the neutrino cross sections, the beam neutrino flux and the oscillation calculations allows predicting the neutrino event rates expected for the DUNE experiment. A dedicated section discusses this point.

A special care is given to the understanding of general detection techniques and concepts to describe them. The key features are presented as well as a method, called smearing, to rapidly take into account the detector energy reconstruction effects at the single particle level. Finally the presentation of the Convolutional Visual Network (CVN), that the DUNE collaboration developed for neutrino flavour identification to assess its sensitivity to its main scientific program, closes this chapter. The CVN is a powerful neutrino flavour classification tool and will lately be used in combination with the kinematic analysis deployed in this thesis.

2.1 Neutrino interactions

2.1.1 Cross sections

In particle physics, the connection between theory and experiments stands in the cross sections. The theory typically predicts a rate at which a given elementary process occurs and experiments perform event counting measurements. The most general way of writing a cross section σ for a scattering process $a + b \to ...$ (where b is the target and a the probe) is:

$$\Gamma = N_b \times \phi_a \times \sigma, \tag{2.1}$$

where Γ is the event rate (s^{-1}), N_b is the number of targets, ϕ_a the flux of incident particles a per unit of time and surface. The cross section σ has the dimensions of an area and its physical magnitude represents how likely a scattering process is going to occur. Experiments will typically measure the event rate of a given process and compare it to the predicted cross section. For a scattering process $a + b \to ...$, the theoretical differential cross section reaches, in the case where there is only one target ($N_b = 1$):

$$d\sigma = \underbrace{\frac{1}{2E_a 2E_b |\overrightarrow{v}_a - \overrightarrow{v}_b|}}_{\text{Flux factor}} (2\pi)^4 |\mathscr{M}_{fi}|^2 \delta^{(4)}(\mathbf{p}_a + \mathbf{p}_b - \mathbf{p}_f) \prod_{f=1}^{N_f} \frac{d^3 p_f}{(2\pi)^3}. \tag{2.2}$$

The notation $\mathbf{p}$ designates a 4-momentum. The subscript f refers to particles in the final state. The Lorentz invariant matrix $\mathscr{M}_{fi}$ contains the Standard Model predictions, and is calculated by taking into account the relevant scattering processes and their corresponding Feynman diagrams. Simulations in particle physics act like a bridge between the theory and

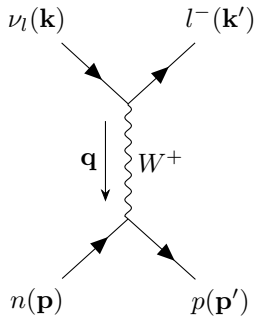


Figure 2.1: Tree Feynman diagrams of a quasi-elastic charged current interaction between a neutrino of flavour l and a free neutron. the particles 4-momenta are indicated between brackets.

experiments which means that the correct calculation of $\mathscr{M}_{fi}$ is crucial in order to conduct an experiment.

For neutrino-nucleus interactions, the total cross section includes three main components, evolving progressively from exclusive final states to more inclusive ones, which are the Quasi-Elastic (QEL), the Resonant (RES) and the Deep Inelastic (DIS) scattering processes. We will briefly overview them, after we introduce the relevant kinematic description of neutrino scattering.

2.1.2 Charged current (CC) neutrino interactions kinematics

We discussed in 1.1.2 that the study of charged current interactions is necessary to oscillation experiments since they rely on the identification of the neutrino flavour. We give in figure 2.1 the tree Feynman diagram of a neutrino of flavour l scattering on a free neutron (the antineutrino case is obtained switching the roles of the proton and the neutron, and switching the W boson electrical charge). This process is called a quasi-elastic scattering, however in the general case the outgoing hadronic part can be more complex. It can involve resonant hadron productions (such as Δ^{++} if the hit nucleon is a proton), see discussion in 2.1.3. Similarly as in figure 2.1, one can define as $\mathbf{k}$ ($\mathbf{k}$') the 4-momentum of the incoming neutrino (outgoing lepton) and $\mathbf{p}$ ($\mathbf{p}$') the 4-momentum of the hit nucleon (outgoing hadronic system), with $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ the 4-momentum transfer. Then the description of the kinematics of these scatterings often involves the following variables:

$$\begin{aligned} s &= (\mathbf{p} + \mathbf{k})^2 = m_n^2 + 2m_n E_\nu \\ \nu &= \frac{\mathbf{p} \cdot \mathbf{q}}{m_n} = E_\nu - E_l \\ Q^2 &= -\mathbf{q}^2 = -m_l^2 + 2E_\nu E_l(1 - \beta_l \cos\theta) \\ W^2 &= m_n^2 + 2\nu m_n - Q^2, \end{aligned} \tag{2.3}$$

where we expressed the calculations in the laboratory frame assuming the neutron at rest. E_l is the energy of the charged lepton, E_ν the energy of the neutrino, m_l is the mass of the charged lepton, m_n is the mass of the neutron, θ is the angle between the neutrino momentum and charged lepton momentum, and $\beta_l = v_l/c$ is the laboratory speed of the charged lepton. s is the Mandelstam variable which corresponds to the energy in the center of mass. Note the variable ν which corresponds to the energy transferred to the hadronic part must not be confused with

the neutrino letter. W^2 is the hadronic final state invariant mass. For the Feynman diagram we sketched in figure 2.1, we merely have $W^2 = m_p^2$, the mass of the proton. This variable W^2 is fixed when the hadronic final state is exclusive, regimes that are described by the quasi-elastic and resonant scatterings.

However, when the transferred momentum Q^2 reaches higher values, one enters the so-called deep inelastic scattering mode, in which the high momentum transfer allows for probing the structure of the proton itself, and a new formalism needs to be developed (see for instance [88, 8.3] [32, 5.3.3]). The deep inelastic regime is defined by the conditions:

$$Q^2 \gg m_n^2 \text{ and } \mathbf{p}\cdot\mathbf{q} \gg m_n^2. \tag{2.4}$$

This scattering regime allowed for the discovery of the proton's structure (quark-partons model) by the end of the 60's at SLAC, where physicists used electrons to probe the internal structure of the proton. The surprising feature was to recover point-like scatterings at high Q^2, where one would have expected the electrons to scatter on a diffuse electrical charge. The DIS formalism makes use of the dimensionless Bjorken variables:

$$\begin{aligned} x &= \frac{Q^2}{2\mathbf{p}\cdot\mathbf{q}} = \frac{Q^2}{2m_n(E_\nu - E_l)} \\ y &= \frac{\mathbf{p}\cdot\mathbf{q}}{\mathbf{p}\cdot\mathbf{k}} = 1 - \frac{E_l}{E_\nu}, \end{aligned} \tag{2.5}$$

where we again expressed the calculations of x and y in the laboratory frame assuming the hit neutron at rest. x represents the fraction of nucleon momentum carried by the interacting valence quark, and y represents the fraction of energy transferred by the neutrino to the hadronic system, in the laboratory frame. Both variables vary between 0 and 1.

The various kinematical variables just defined allows to have a simple view of the three scattering processes QEL, RES and DIS.

2.1.3 Quasi-elastic (QEL), Resonant (RES) and Deep Inelastic (DIS) scatterings

Quasi-Elastic:

An elastic scattering is by definition a scattering in which the kinetic energy of the total system is conserved. One can think classically of two billiards balls colliding. The energy dissipation due to the inelasticity of the balls is negligible, and the kinetic energy of the two balls is conserved throughout the scattering process. Neutrinos rarely interact elastically. A counter example is $\nu_e + e^- \to e^- + \nu_e$, but cross section on electrons is several orders of magnitude smaller than the cross section on nuclei.

Quasi-elastic scattering is the simplest way a neutrino can interact with a nucleon. The reaction for flavour α is $\nu_\alpha + n \to \alpha^- + p$ for neutrinos and $\bar{\nu}_\alpha + p \to \alpha^+ + n$ for antineutrinos. $\alpha^\pm$ is the charged lepton coupled to the neutrino of flavour α. The 1^{st} order Feynman diagram of the neutrino case is given in figure 2.1. Since the neutrino interacts with a nucleon bound in the argon nucleus, additional nuclear effects have to be included in the cross section calculations. One of them is the Fermi momentum, for which we consider the target nucleon to have a non-null momentum inside the nucleus. This momentum is typically at the level of 200 MeV.

The scattering is not entirely elastic for three reasons: there is a small mass difference between the proton ($938.27\,\mathrm{MeV}/c^2$) and the neutron ($939.57\,\mathrm{MeV}/c^2$), there is the creation of the massive charged lepton, and because the target nucleon is bound in an argon nucleus, the emitted nucleon may re-scatter inside the nucleus (affecting its energy) before leaving the nucleus and becoming observable. Given the fact that the primary interaction is a two body scattering process where both the nucleon and the lepton change their charge, but the hit nucleon remains a single nucleon, the process is called quasi-elastic scattering. The kinematics at the lepton vertex is approximately the one to scatter elastically on a free nucleon and it is characterized by a small momentum transfer Q^2, with an associated wavelength of the order to the nucleon size. A form factor disfavours higher transferred momenta Q^2.

Since $W^2 = m_p^2$, one can prove that the Bjorken variable $x \simeq 1$, which indeed indicates that the neutrino interacts with the nucleon as a whole. Moreover in this particular case the neutrino energy can be deduced from the outgoing charged lepton kinematics only:

$$E_\nu = \frac{2m_n E_l + m_p^2 - m_n^2 - m_l^2}{2m_n - 2E_l(1 - \beta_l \cos\theta)}, \tag{2.6}$$

if we neglect the nucleon's binding energy and assume it is at rest in the laboratory frame.

For completeness, it should be noted that quasi-elastic scatterings can involve other final states baryons than neutron/proton. If the neutrino brings sufficient energy, one can observe the creation of the charged baryons $\Lambda_c^+ = (udc)$ (isospin 0), $\Sigma_c^+ = (udc)$ (isospin 1) and $\Sigma_c^{++} = (uuc)$ (isospin 1). Past experiments like CHORUS [89] have measured the occurrence of such processes for muon neutrinos of mean energy 27 GeV, and found 1.7‰ for the Λ_c^+ and 0.7‰ for the Σ_c^{++}, normalized to the total charged current cross section of muon neutrino.

Using the GENIE v3.00.02 muon neutrino cross sections on argon, we found that for DUNE we should expect a fraction of 0.2% for the Λ_c^+, 0.05% for the Σ_c^+ and 0.09% for the Σ_c^{++}, also normalized to the total charged current cross section. These percentages are flux averaged by the oscillated muon neutrino spectrum at the far detector. We find the fractions to be comparable with CHORUS measurements. Since both the total charged current and charmed QEL cross sections are disfavoured at small energies, a direct comparison between CHORUS and the expectation for DUNE is delicate.

Resonances:

Evolving in complexity from the situation described above for the quasi-elastic scattering, the interaction of a neutrino impinging on a nucleon can involve the creation of hadronic resonant states typically decaying into a nucleon plus a pion. These processes occur at higher energy than quasi-elastic scattering since the creation of massive resonant states requires sufficient momentum transfer. In GENIE, the resonant scattering reactions are based on the Rein-Sehgal model [90].

As for illustration, we give the example of single pion resonant production by charged current interactions of a neutrino on a nucleon: $\nu_\mu + p \rightarrow \mu^- + p + \pi^+$, $\nu_\mu + n \rightarrow \mu^- + p + \pi_0$, $\nu_\mu + n \rightarrow \mu^- + n + \pi^+$. The same reactions are of course possible for the other two neutrino flavours.

The exclusive final states at the hadronic vertex fixes the value of W^2, thus the description of resonant and quasi-elastic scatterings is commonly described by one kinematical variable, either Q^2 or ν.

Deep Inelastic Scattering (DIS):

At higher neutrino energies (typically several GeV), the DIS contribution becomes dominant with respect to the two latter since its cross section increases linearly with the neutrino energy while the QEL and RES contributions tend to keep constant. DIS involves higher transferred momenta corresponding to a De Broglie wavelength smaller than the size of a nucleon ($\simeq 10^{-15}$m). The neutrino can resolve the structure of the nucleons and interact with the quarks themselves.

The final hadronic state is more inclusive than in quasi-elastic and resonant regimes, because the proton will not survive the high transferred energy. Thus, W^2 is not fixed anymore, and the scattering is described by two independent variables, such as W^2 and Q^2. That said, the Bjorken variables are more often used for their practical meaning: y represents the energy conceded by the neutrino to the hadronic system, so it informs on the inelasticity of the event. An elastic scattering on a free nucleon would result in a low energy transferred, so $y \to 0$. x represents the fraction of nucleon momentum carried by the interacting valence quark.

Other interaction processes:

There are also other scattering modes of neutrino on matter, with smaller contributions than the three main QEL, RES and DIS. We mention non-exhaustively in the following two of them:

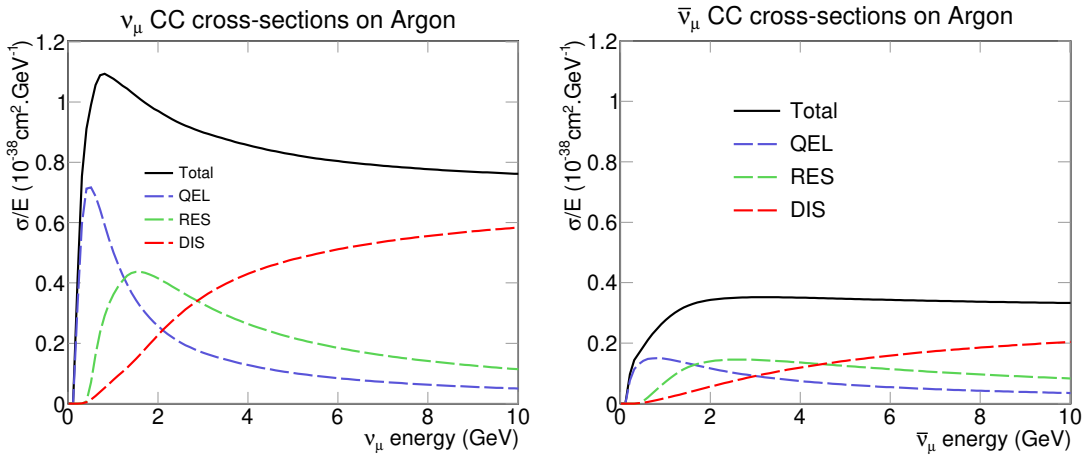
- IMD (inverse muon decay): neutrinos can interact, in a negligible way, with electrons present in the detector, via for example an inverse muon decay $\nu_\mu + e^- \to \nu_e + \mu^-$.
- COH: this term literally means "coherent". It refers to the a process in which the neutrino reacts with the whole nucleus A and produces only a pion $\nu_\mu + A \to \mu^- + \pi^+ + A$.

2.1.4 Total neutrino cross section in GENIE

The GENIE cross sections are given in terms of splines. Each of them displays the energy dependence of a given scattering process. In figure 2.2, we show the total cross section of charged current interactions of muon neutrinos (left) and muon antineutrinos (right), divided by neutrino energy, on argon. We display as well the contributions of QEL, RES, and DIS scatterings. These splines were obtained with GENIE v3.00.02, and are normalized to the number of nucleons in argon (40).

The cross sections are divided by the neutrino energy in order to highlight the linear behaviour of the DIS cross section which becomes dominant at high energies. A striking feature of the cross sections is the difference between neutrinos and antineutrinos. At high energy, where the total spline (dominated by DIS scattering) flattens, one can see that $\sigma_{\nu_\mu}/E \simeq 2 \times \sigma_{\bar{\nu}_\mu}/E$. The antineutrino-quark scattering is disfavoured at high y because of angular momentum conservation along the interaction axis. It can indeed be showed that a high y corresponds to a backward scattering in the center of mass frame, and the latter is disfavoured as the right-handed antineutrino and left-handed quark helicities are parallel. In the neutrino-quark scattering, the spin projections are antiparallel and thus angular momentum conservation does not favour any direction for the scattered particles.

The Particle Data Group regularly updates the global knowledge in particle physics and cosmology. One can find all the current muon neutrino and antineutrino cross section measurements in [91, chap.51 p.692]. This plot gathers all the available cross section data with respect to



(a) (b)

Figure 2.2: Muon neutrino (left) and antineutrino (right) charged current cross sections on argon divided by the neutrino energy and the atomic number of argon (40). The black spline is the total cross section, and we also show the QEL (blue), RES (green) and DIS (red) contributions.

neutrino energy on a various set of targets. The cross sections are normalized to the number of nucleons in each target. We have displayed this plot in figure 2.3 as well as the total charged current cross section of muon neutrino on argon (blue and red) as predicted by GENIE v2.12.10 and v2.08.06, normalized to the number of nucleons. Even if the blue and red splines "fit" in the overall data, reader should be careful when comparing cross sections measurements obtained with different targets. The nuclear effects are not a linear function of the number of nucleons, and it is hard to measure an uncorrelated cross section from the target used. In addition the isoscalarity (equal fraction of neutrons and protons in the nucleus) of the targets used must be taken into account for a proper comparison. This asymmetry is taken into account in the GENIE splines on argon, and the normalization to the total number number of nucleon is simply a convenient way of looking at things together.

2.1.5 Kinematic suppression factor of ν_τCC

The charged current interaction of a neutrino implies the creation of the charged lepton of the same flavour. As a consequence, and in contrast to neutral current interactions, charged current interactions have an energy threshold since the neutrino must bring sufficient energy to allow for the massive lepton creation. The beam neutrinos in DUNE will have energies of typically $\mathcal{O}(4\,\mathrm{GeV})$ and the mass of the $\tau^\pm$ is $1.777\,\mathrm{GeV}/c^2$, which means the τ neutrino CC cross section is kinematically disfavoured in a non-negligible way in comparison to ν_μ and ν_e CC cross sections.

In order to calculate the charged current energy threshold, we consider a QEL scattering of a neutrino ν_α on a free neutron n (see figure 2.1 for the notation conventions), *i.e* the reaction $\nu_\alpha(\mathbf{k})+n(\mathbf{p}) \rightarrow \alpha^-(\mathbf{k'})+p(\mathbf{p'})$, where the 4-momenta of the particles are written in the laboratory frame as:

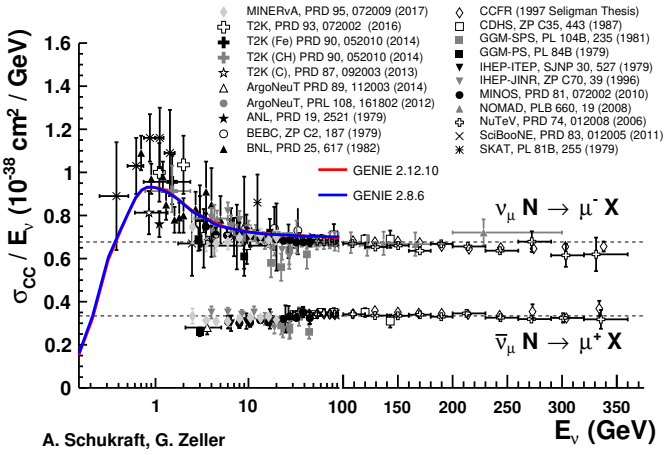


Figure 2.3: Total ν_μ cross section on argon as predicted by GENIE (v2.12.00 in red and v2.08.06 in blue) divided by the number of nucleons (40) and by neutrino energy in blue and red. The plot also gathers all current ν_μ and $\bar{\nu}_\mu$ neutrino cross section measurements on a various set of targets.

$$\mathbf{k} = \begin{pmatrix} E_\nu \\ 0 \\ 0 \\ E_\nu \end{pmatrix}, \mathbf{p} = \begin{pmatrix} m_n \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{k'} = \begin{pmatrix} E_\alpha \\ \vec{p}_\alpha \end{pmatrix}, \mathbf{p'} = \begin{pmatrix} E_\mathrm{p} \\ \vec{p}_\mathrm{p} \end{pmatrix}. \tag{2.7}$$

We have assumed the neutrino to be massless and the target neutron to be at rest. In the center of mass frame, the energy available $E^\star$ is straightforward and is connected to the Mandelstam variable as:

$$E^\star = \sqrt{s}, \text{ and } s = (\mathbf{k} + \mathbf{p})^2 = 2E_\nu m_n + m_n^2. \tag{2.8}$$

The energy must allow for the creation of the massive particles (proton and charged lepton), which implies:

$$E^\star \geqslant (M_\alpha + m_p) \iff E_\nu \geqslant \frac{(m_p^2 - m_n^2) + 2m_p M_\alpha + M_\alpha^2}{2m_n}. \tag{2.9}$$

The antineutrino case is given by inverting the roles of the proton and the neutron. Numerical values of these thresholds are: 0 for ν_e (the result is actually lower than 0), 0.11 GeV for ν_μ and finally 3.45 GeV for ν_τ. At DUNE considered energies, this means that there will be a significant kinematic suppression factor between ν_τ and ν_e/ν_μ CC coming from the cross sections. We illustrate this point by defining the factor $\Gamma^{CC} = \dfrac{\sigma^{CC}(\nu_\tau)}{\sigma^{CC}(\nu_\mu)}$. We define as well the same factor for the anti neutrino case: $\overline{\Gamma}^{CC} = \dfrac{\sigma^{CC}(\overline{\nu}_\tau)}{\sigma^{CC}(\overline{\nu}_\mu)}$. The evolution of these factors with respect to the neutrino energy are given in figure 2.4. The suppression factors comparing the ν_τ and ν_e CC cross section give very similar results. Mind the threshold at 3.45 GeV indicated with the arrow. At 10 GeV, the ν_τ cross section reaches only 25% of the ν_μ cross section. The suppression factors of ν_τ and $\bar{\nu}_\tau$ are similar.

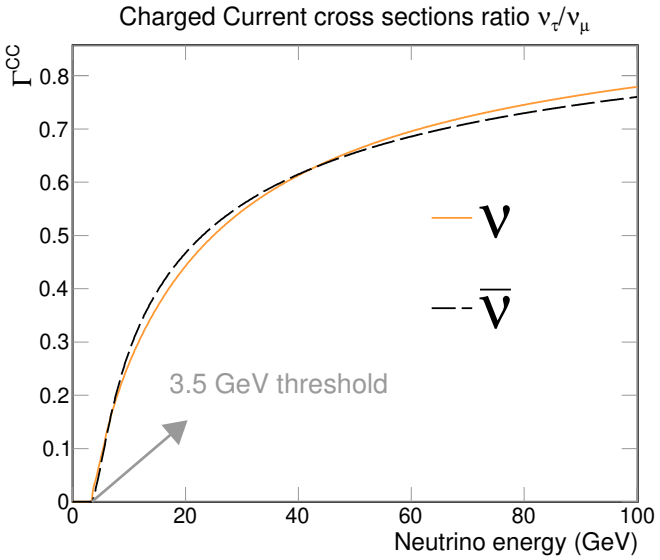


Figure 2.4: Suppression factors Γ^{CC} (neutrino) and $\bar{\Gamma}^{CC}$ (antineutrino) on argon with respect to the neutrino energy, showing the kinematic suppression factor for the τ flavour. The suppression factor is defined as the ratio between the ν_τ and ν_μ CC cross sections on argon. This figure was obtained using the splines of GENIE v3.00.02.

For completeness, it should be noted that in the energy threshold calculation the hit neutron was assumed to be at rest. Nucleons inside the nucleus have a Fermi momentum at the level of $p_n = 200\,\mathrm{MeV}/c$. In the most favourable case where the neutron Fermi momentum is antiparallel to the neutrino direction, a collider effect is observed and it lowers the neutrino energy threshold following:

$$E_\nu \geqslant \frac{(m_p^2 - m_n^2) + 2m_p M_\alpha + M_\alpha^2 + p_n^2}{2(m_n + p_n)}. \tag{2.10}$$

Numerically this gives $E_\nu \geqslant 2.86\,\mathrm{GeV}$. It should be noted that if the neutrino Fermi momentum is parallel to the neutrino direction ($p_n \to -p_n$), then the energy available in the center of mass is lowered (unfavourable case) and the neutrino energy threshold increases to 4.41 GeV.

2.1.6 DUNE will not see the τ !

The basic assumption to identify the impinging neutrino flavour in a neutrino detector is to tag the outgoing lepton (electron, muon or tau) thanks to its signature (here in the LArTPC). At DUNE considered energies, both the electron and the muon will leave observable signatures. Though the muon is unstable, it will propagate over long distances depositing ionization along its trajectory. The electron, due to its mass much smaller than the one of the muon, will mainly loose energy by Bremsstrahlung radiation and produce an electromagnetic shower. The tau, contrary to the muon, will not travel over macroscopic distances before decaying, and we suggest to illustrate this point with a numerical example.

Let us assume a 3 GeV τ lepton created at a neutrino vertex. Its rest frame lifetime is $\tau_{rf} = 3 \times 10^{-13}$s. The typical distance covered by the τ before decaying in the laboratory frame L_{lab} obeys (the relativistic γ factor is $\gamma = E_\tau/m_\tau$ in natural units):

$$L_{lab} \leq \tau_{rf} \times c \times \frac{E_\tau}{m_\tau} \simeq 150\,\mu\text{m}. \tag{2.11}$$

150 µm is not a distance that the DUNE LArTPC will be able to resolve. Even a 10 GeV τ would propagate over half a millimeter only. We conclude that tagging the τ lepton requires more sophisticated tools, DUNE will not be able to rely on a visual kink as exposed in figure 1.9.

2.2 Beam fluxes

2.2.1 Producing the DUNE neutrino flux

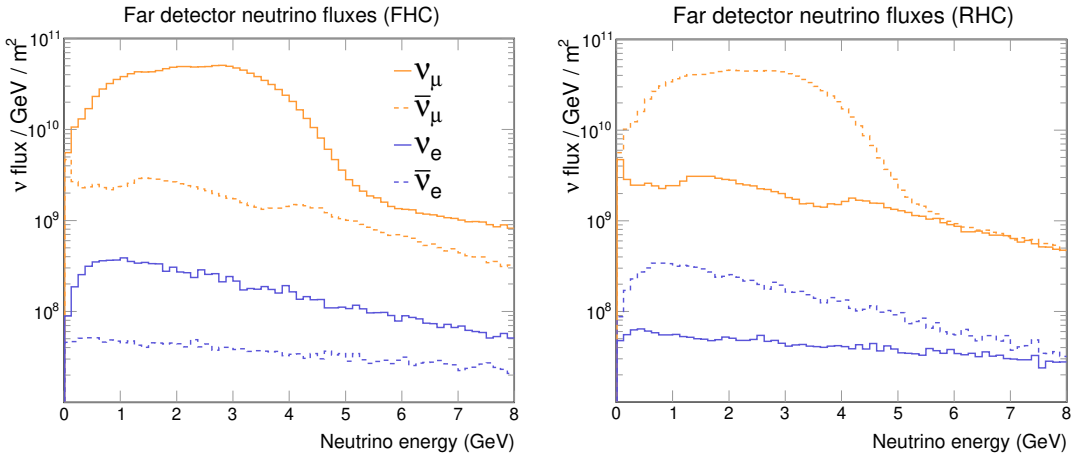
The neutrino flux at the near detector and the far detector sites was simulated by the DUNE collaboration using G4LBNF, a simulation package based GEANT-4, adapted to LBNF (Long Baseline Neutrino Facility). The data are available on this webpage [92]. The DUNE collaboration used the so-called CP optimized beam for the Technical Design Report. This beam was designed in order to maximize the sensitivity of the experiment to the CP violation measurement.

The accelerator complex of the Fermilab will be used to provide 120 GeV protons (using the Main Injector ring) with a power of 1.2 MW. The number of protons delivered in one year $n_p^{(1\,yr)}$ follows $n_p^{(1\,yr)} \times 120\,[\text{GeV}] = 1.2\,[\text{MW}] \times t^{(1\,yr)}$ (noting $t^{(1yr)}$ the time corresponding to one year). Numerically this gives $n_p^{(1\,yr)} \simeq 1.97 \times 10^{21}$ POT (Protons On Target), higher than the usual value 1.1×10^{21}POT used by the collaboration, because we need to take into account the efficiency and the actual runtime of the Fermilab accelerator complex, which is about 57%. In this thesis, we will use the reference protons on target value of 1.1×10^{21}POT.

The accelerated protons will hit a cylindrical graphite target which will be 16 mm large (diameter) and 2.2 m long, producing a bunch of secondary hadrons, mostly pions. The latters will be focused using a series of three magnetic horns operating at 300 kA, and sent to a 200 m long decay pipe. The charged pions decay about 99.9% of the time as $\pi^+(\pi^-) \to \nu_\mu(\bar{\nu}_\mu) + \mu^+(\mu^-)$, producing the muon (anti)neutrino beam. The experiment will be able to run in the so-called Forward Horn Current (FHC) and Reverse Horn Current (RHC) modes, selecting respectively positively and negatively charged pions. This means that DUNE will have the possibility to run with a beam of muon neutrinos or of muon antineutrinos. This is a basic requirement for a beam neutrino experiments which aims at establishing a matter/antimmater asymmetry.

The muon neutrino flavour is expected to have a 1% contamination of electron flavour coming from kaon (charged and neutral) and muon decays, for both the FHC and RHC running modes. However the dominant contamination is expected to be the wrong sign background, that is to say π^- contributing to the FHC beam and π^+ to the RHC beam. This background is expected to contribute at the level of 10%. The simulated fluxes at the far detector site are given in figure 2.5 for both the FHC (left) and RHC (right) running modes. In this plot we have not taken into account the oscillations, so we actually show the unoscillated fluxes. Though this does not reflect the actual flavour proportion expected at the far detector site, it is a convenient way to master the flux and the oscillation calculations separately.

This method to produce artificial beams of neutrinos, using a proton accelerator complex and a hit target, is the common way of doing beam neutrino oscillations physics. It was actually first used at Brookhaven National Laboratory and allowed for the ν_μ discovery in 1962 [7].



(a) (b)

Figure 2.5: Neutrino unoscillated fluxes expected at the DUNE far detector in FHC mode (*i.e* neutrino mode, left) and RHC mode (*i.e* antineutrino mode, right). The plots show the contributions of the ν_μ, $\bar{\nu}_\mu$, ν_e, $\bar{\nu}_e$ which were created at the beam site (Fermilab). Note that fluxes are given per unit of energy with an energy bin size of 0.125 GeV.

Unfortunately, this method comes with quite high uncertainties at the level of 10%. The DUNE collaboration assessed these uncertainties, and we give the result obtained for the ν_μ flavour in both FHC and RHC modes in figure 2.6, taken from [41, fig.4-5]. There exist two main sources of uncertainties. The dominant one concerns the hadron production at the target, more specifically what hadrons leave the target (hadrons produced within the target may re-interact, be deflected or absorbed). These uncertainties are labeled as "hadron production". The second source is referred to as "focusing": it gathers the uncertainties associated to the beam parameters, such as the proton on target counting, the horn currents, the target and horn positioning... Finally we see from figure 2.6 that ν_μ uncertainties in FHC mode and $\bar{\nu}_\mu$ uncertainties in RHC mode are quite similar.

The unoscillated flux is expected to be $\mathcal{O}(10^6)$ times greater at the ND than at the FD. The near detector will have a key role on constraining the neutrino flux systematic uncertainties, which are correlated between the two sites.

This beam method production will be deployed by the DUNE collaboration. The standard LBNF beam design will allow for the production of a muon neutrino beam optimized for the search of CP violation in $\nu_\mu \to \nu_e$ oscillations. An alternative beam design has also been proposed in DUNE internal reports, studied to improve the DUNE sensitivity to the ν_τ appearance.

2.2.2 τ optimized beam flux

The DUNE scientific program aims primarily at measuring the CP violation in the leptonic sector, the neutrino mass ordering and the octant of the mixing angle θ_{23}. Thus, the neutrino flux which will be used for this program is designed in order to achieve an optimized sensitivity to the δ_{CP} parameter via the study of the $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations. The $\nu_\mu \to \nu_e$ oscillation channel has several maxima in terms of L/E_ν (L being the length of the baseline). The typical energy E_ν is chosen so that the range of L/E_ν covers the so-called second oscillation maximum in addition to covering the first oscillation maximum (see the blue curve in figure

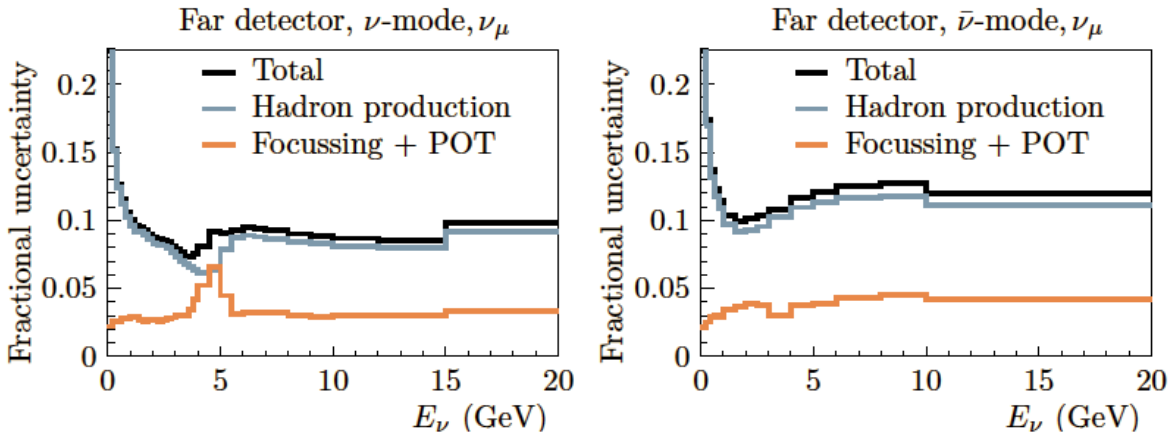


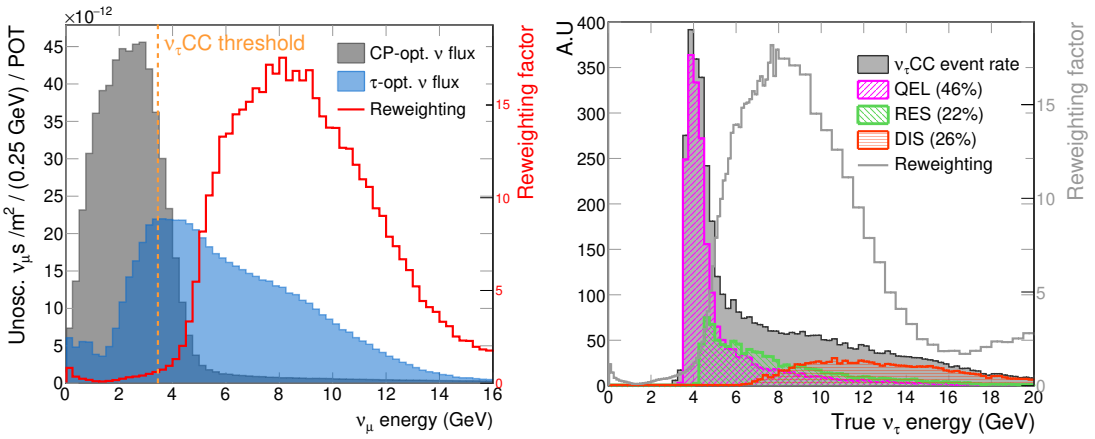
Figure 2.6: Neutrino flux uncertainties at the FD related to the ν_μ flavour production in FHC mode (left) and RHC mode (right) as can be found in [41, fig.4-5]. Uncertainties are divided in two components: hadron production (what hadrons actually leave the neutrino target) and focusing (related to beam parameters, such as horn current) uncertainties. The first category dominates the total uncertainties.

2.8 where the second maximum is at 1 GeV for and the first one at 2.5 GeV) occurring in this oscillation channel, because of its particular sensitivity to the δ_{CP} parameter.

The CP optimized neutrino flux will be used all along this thesis as a matter of coherence with the collaboration, and it can be found on [92]. Yet, if the δ_{CP} true value happened to be not too close from CP conservative values (0 and π), DUNE would conclude on the symmetry violation within the running time of the experiment [42] (10 years for a 5σ confidence). The DUNE collaboration may thus plan to tackle other scientific programs. One possibility discussed would be to use a higher energy neutrino flux, called the τ optimized flux for its ability to boost the DUNE sensitivity to the beam ν_τ appearance. Indeed, as discussed in 2.1.5, the ν_τ charged current cross section has a threshold at 3.45 GeV, which means that most of the τ neutrinos of the CP optimized flux will not be able to produce a charged current interaction. In figure 2.7(a), we show the superposition of the CP optimized unoscillated muon neutrino flux (black filled) and τ optimized unoscillated muon neutrino flux (blue filled) at the far detector site for neutrino energies in [0 ; 15]GeV, assuming the FHC running mode. For each energy bin, the red histogram shows the ratio of the τ optimized flux on the CP optimized flux (with its own axis on the right of the plot). This red histogram will lately be used in case we want to reweight some results obtained with the CP optimized flux in the light of the τ optimized flux. We assess that 79% of the CP optimized neutrino flux stands below the 3.45 GeV ν_τCC threshold (shown in orange), this number falls to 23% for the τ optimized flux.

An other interesting feature of the τ optimized flux is the relative proportions of the QEL, RES, DIS scattering types of the ν_τCC interactions. We show this feature on 2.7(b). Using the CP optimized neutrino flux, about half (46%) of the ν_τCC events are quasi-elastic, 23% are resonant and 26% are deep inelastic. We assessed that the τ optimized reweigthing would dwindle this quasi-elastic population to 32%. The RES and DIS fractions would respectively reach 36% and 27%. The τ optimized flux will favour RES scattering and disfavour QEL scattering, but surprisingly will not affect the DIS fraction.

In this thesis the τ optimized beam will assume a beam power of 1.2 MW. However it should be noted that a power upgrade to 2.4 MW is planned for DUNE after six years of running which would foster the neutrino event statistics [41, Section 5.2]. Indeed the beam power is



(a) (b)

Figure 2.7: Left: CP optimized unoscillated neutrino flux (black filled) and τ optimized flux (blue filled), as well as their ratio for each energy bin (red histogram with its own axis on the right), which will be used for reweigthing results from CP optimized to τ optimized flux. The ν_τCC threshold of 3.45 GeV is shown in orange. One can see that most of the neutrinos (79%) of the CP optimized flux have energy below the threshold. This fraction falls to 23% for the τ optimized flux. Right: relative abundance of QEL/RES/DIS in the ν_τCC far detector event rate for the CP optimized flux. In gray is also displayed the reweigthing histogram which was shown in red on the left plot. This reweigthing fosters the RES scatterings and disfavours the QEL scatterings.

related to the number of accelerated protons hitting the target, a magnitude referred as the number of protons on target (POT). However it does not affect the shape of the neutrino energy distributions shown in figure 2.5.

DUNE is planning a beam power upgrade from 1.2 MW to 2.4 MW after several years of running. This timeline suggests that it is unlikely that the τ optimized beam configuration (if used) is deployed with a 1.2 MW power. Nevertheless the 1.2 MW beam power will be assumed both for the CP optimized and τ optimized neutrino beam configurations to allow for a direct comparison.

2.3 Oscillations probabilities calculations

The far detector neutrino fluxes (see figure 2.5) are unoscillated, that is to say everything goes like there is no flavour change along the baseline. To oscillate the fluxes, we use GLoBES [93, 94] (General Long Baseline Experiment Simulator), a software package designed to simulate baseline neutrino experiments and predict for instance event rates and $\Delta\chi^2$. At our level, we will use a much lower information level since we will only compute oscillation probabilities in matter to weight each neutrino event by the corresponding probability. This procedure allows to make oscillate the fluxes of figure 2.5 in a convenient way. All along the thesis, we will use the oscillation parameters (PMNS and Δm^2) given in table 2.1. It corresponds to the configuration used in the DUNE TDR [95] assuming the normal hierarchy of the neutrino mass states, and with the difference that a CP conservative value $\delta_{CP} = 0$ was chosen for this thesis. The baseline length is set to 1285 km and earth density, when assumed constant, reaches 2.8 g cm^{-3}.

We show the oscillation probabilities $\nu_\mu \to \nu_{\alpha=e,\mu,\tau}$ in vacuum (full line) and in matter (dashed lines, assuming a constant earth density) calculated with GLoBES, as a function of neutrino

Mixing angles and phase (rad.)				Masses (eV^2)		Baseline	
θ_{12}	θ_{23}	θ_{13}	δ_{CP}	Δm^2_{21}	Δm^2_{32}	L (km)	ρ ($g\,cm^{-3}$)
0.5903	0.866	0.150	0.	7.39×10^{-5}	2.451×10^{-3}	1285	2.8

Table 2.1: Oscillation parameters used in the DUNE TDR [95] except for δ_{CP}. This choice assumes the Normal Ordering (NO) hypothesis and uses a CP conservative value $\delta_{CP} = 0$ (which is not the case in the DUNE TDR).

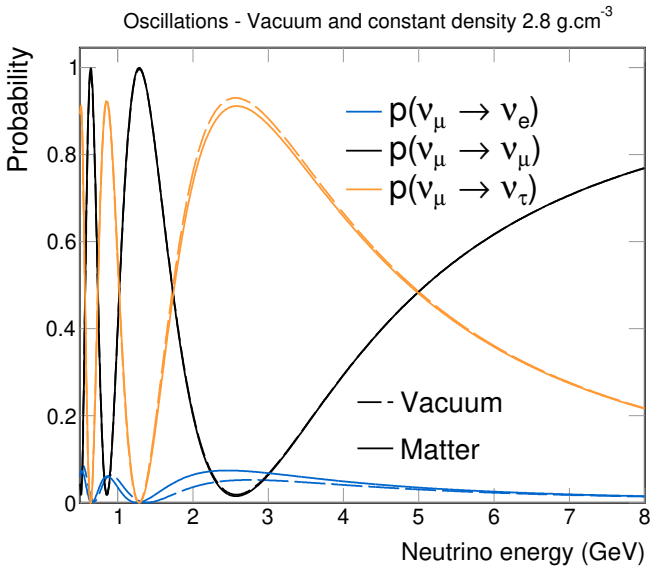
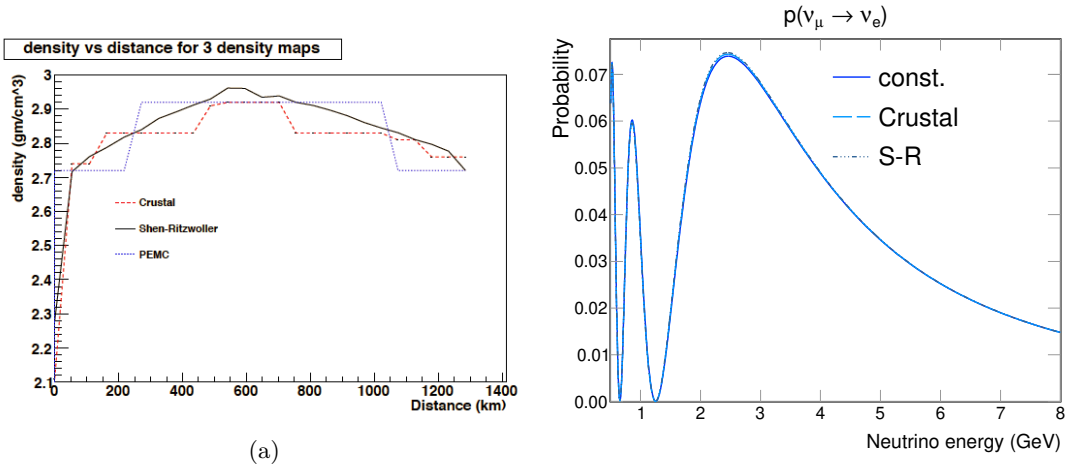


Figure 2.8: Oscillation probabilities of $\nu_\mu \to \nu_e$(blue), $\nu_\mu \to \nu_\mu$(black) and $\nu_\mu \to \nu_\tau$(orange) in vacuum (dashed line) and in matter (full line) assuming a constant earth density of $2.8\,gcm^{-3}$. The dominant oscillation channels are $\nu_\mu \to \nu_\mu$ and $\nu_\mu \to \nu_\tau$, while the $\nu_\mu \to \nu_e$ oscillations are only at the level of 10%. The electron flavour is the most sensitive to matter effects.

energy, in figure 2.8. The dominant oscillation channels are $\nu_\mu \to \nu_\tau$ and $\nu_\mu \to \nu_\mu$, and the $\nu_\mu \to \nu_e$ oscillation is sub-leading at the level of 10%. One can observe the two $\nu_\mu \to \nu_e$ oscillation maxima at 0.8 GeV and 2.5 GeV. We note that the electron flavour is quite sensitive to matter effects, while the muon neutrino survival probability is barely affected.

To account for matter effects more deeply, we refer to [96], which compares various earth density models and discusses their relative effects on oscillation probabilities. We report in figure 2.9a the earth density evolution along the neutrino baseline for the so-called Crustal, Shen-Ritzwoller and PEMC models. The impact of the earth Density model choice is evaluated on the $\nu_\mu \to \nu_e$ oscillation channel (because of its particular sensitivity to matter effects) in figure 2.9b. We compared the Shen-Ritzwoller and Crustal models to the basic assumption of a constant matter density ($2.8\,g\,cm^{-3}$). We omitted the PEMC model. In agreement with [96], effects are relatively small. As for sake of simplicity, we will use in this thesis the constant density model.



(a) (b)

Figure 2.9: Left: various earth density models along DUNE baseline described in [96]. Here is shown the evolution of earth density as a function of the baseline position. We then compare on the right three earth density models: constant density ($2.8\,\mathrm{g\,cm^{-3}}$), the Crustal and the Shen-Ritzwoller ones. Effects are very small, and are only slightly visible on the maximum around 2.5 GeV.

2.4 Neutrino event rates

2.4.1 Calculation of the expected event rates

The combination of the neutrino flux, the neutrino cross sections and the calculation of oscillation probabilities presented in the previous sections allow us predicting the neutrino event rates expected at the DUNE far detector site. In equation 2.1 we wrote that a general event rate in particle physics is given by the product of the incident particles flux, the cross section at play and the number of targets. Oscillation probabilities must be added to this general formula in order to make the incident neutrino flux oscillate into the desired flavour. In DUNE the neutrino flux is given in units of POT (protons on target) which is an LBNF characteristic. For a given physical process described by a cross section σ (which can describe the charged/neutral current interactions of a neutrino flavour, or a single specific scattering type...), the event rate reaches:

$$N_{\text{event}}(E_\nu) = \underbrace{\phi(E_\nu)}_{\nu.\mathrm{m}^{-2}.\mathrm{POT}^{-1}} \times \underbrace{\sigma(E_\nu)}_{10^{-38}\mathrm{cm}^2} \times p_\alpha \times POT \times \underbrace{\mathscr{N}_A \frac{m_{FD}}{M(Ar)}}_{\text{Number of argon targets}}, \tag{2.12}$$

where $\mathscr{N}_A$ is the Avogadro number, m_{FD} the fiducial mass (*i.e* the mass used for the detection) of the far detector, $M(Ar)$ the mass of one mole of liquid argon, ϕ is the unoscillated incident neutrino flux, σ the cross section at play, p_α is the relevant oscillation probability. As the flux, the cross sections and the oscillation probabilities are energy dependent, so the event rates are often given as a function of neutrino energy. The factor POT is proportional to the time of exposure chosen: for DUNE one year corresponds to 1.1×10^{21} protons on target.

The integration of equation 2.12 with respect to neutrino energy allows to predict the total number of neutrino events of a certain flavour and for a certain interaction type. The energy interval used in the DUNE TDR [41, Tables 1.1 & 1.2] for this integration is [0.5 ; 8.0]GeV. It is

however not too relevant for the study of ν_τCC interactions, since 8 GeV is approximately the expected mean energy of interacting tau neutrinos in DUNE (without considering the atmospheric neutrinos). We choose a larger integration range [0.5 ; 50.0]GeV to compute the expected number of neutrino events at the DUNE far detector. These numbers are important, since they will be used as a reference for the normalization of the efficiencies obtained with the analysis later developed in this thesis.

2.4.2 The 3.5 years staged deployment plan

The exposure time used to calculate the neutrino event rates is specified in this section. We use the so-called 3.5 years staged hypothesis which takes into account a delay between the installation of the four DUNE far detector modules, each accounting for 10 kt of fiducial mass. This timeline assumes one year of running with 20 kt (two modules), then two years with 30 kt detectors (three modules), and finally half a year with a full 40 kt detector configuration. The beam power during all this period is assumed to be 1.2 MW. DUNE sensitivities (mass hierarchy, CP violation..) were estimated using this limited schedule [42], and we will also refer to this deployment plan when it comes to calculate neutrino event rates.

We can express the 3.5 years stage event rate $N^{(3.5)}_{events}$ with the event rate corresponding to one year and one far detector module $N^{10\,\mathrm{kt}}_{1year}$:

$$N^{(3.5)}_{events} = \underbrace{2\ N^{10\,\mathrm{kt}}_{1year}}_{1^{\mathrm{st}}\ \mathrm{year}} + \underbrace{2\times 3\ N^{10\,\mathrm{kt}}_{1year}}_{2^{\mathrm{nd}}\ \mathrm{and}\ 3^{\mathrm{rd}}\ \mathrm{years}} + \underbrace{\frac{1}{2}\times 4\ N^{10\,\mathrm{kt}}_{1year}}_{\mathrm{half\ of}\ 4^{\mathrm{th}}\ \mathrm{year}} \tag{2.13}$$

$$= 10\ N^{10\,\mathrm{kt}}_{1year}. \tag{2.14}$$

The expected number of events for various neutrino flavours and interaction types are summarized in table 2.2, for both the neutrino and antineutrino beam configurations. We have distinguished three models of earth density to account for matter effects: a constant density ($2.8\,\mathrm{g/cm^3}$), the Shen-Ritzwoller and the Crustal models which we exposed in 2.3. The ν_e event rate due to $\nu_\mu \to \nu_e$ oscillations is the most sensitive to the model of earth density used. Other flavours as well as the ν_e due to the neutrino beam contamination are not much affected.

It should be noted that we have not written uncertainties to go with the calculated event rates. As for illustration, the Poisson 1σ statistical fluctuation of ν_e due to $\nu_\mu \to \nu_e$ oscillations assuming the constant density model and a neutrino beam running mode (1197 events) is $\sqrt{1197} \approx 35$ (which represents a relative fluctuation of 2.9%). We also showed the neutrino flux systematic uncertainties which are estimated by the DUNE collaboration to be at the level of 10% (see figure 2.6), and one could in addition include uncertainties related to the neutrino cross sections and oscillation parameters. Such considerations are however beyond the scope of this thesis, which is to deploy signal/background analysis in order to assess the DUNE sensitivity to the $\nu_\mu \to \nu_\tau$ oscillation channel. We will at first order stick to the events rates exposed in table 2.2 whenever we wish to normalize the forthcoming signal selection (ν_τ events) and corresponding background rejection efficiencies to the DUNE expected event rates.

The exposed neutrino event rates are reasonable when compared to the ones which can be found in [42, Tables 6, 7]. The number of beam ν_e is smaller than ours because we used a larger integration interval. It might be surprising, though, that in antineutrino mode, one expects

	Constant density	**Shen-Ritzwoller**	**Crustal**
ν mode			
ν_e from osc.	1197 (564)	1207 (555)	1202 (559)
$\bar{\nu}_e$ from osc.	18 (29)	18 (30)	18 (30)
ν_e from beam cont.	365 (371)	365 (372)	365 (371)
$\bar{\nu}_e$ from beam cont.	57 (56)	57 (56)	57 (56)
ν_μ	9660 (9674)	9659 (9675)	9660 (9675)
$\bar{\nu}_\mu$	741 (732)	741 (732)	741 (732)
ν_τ from oscillation	270 (290)	270 (290)	270 (290)
$\bar{\nu}_\tau$ from oscillation	25 (26)	25 (26)	25 (261)
NC	8832 (8832)	8832 (8832)	8832 (8832)
$\bar{\nu}$ mode			
ν_e from osc.	92 (47)	92 (47)	92 (47)
$\bar{\nu}_e$ from osc.	206 (416)	203 (420)	204 (418)
ν_e from beam cont.	184 (186)	184 (186)	184 (186)
$\bar{\nu}_e$ from beam cont.	109 (107)	109 (107)	109 (107)
ν_μ	3096 (3078)	3096 (3078)	3096 (3078)
$\bar{\nu}_\mu$	2653 (2596)	2654 (2595)	2653 (2595)
ν_τ from oscillation	92 (99)	92 (99)	92 (99)
$\bar{\nu}_\tau$ from oscillation	61 (63)	61 (63)	61 (63)
NC	4796 (4796)	4796 (4796)	4796 (4796)

Table 2.2: Number of events expected at the DUNE FD (located 1285 km away from the beam production point) corresponding to the 3.5 years staged deployment plan. The PMNS parameters used are recalled in table 2.1. Both the neutrino (top) and antineutrino (bottom) beam configurations are given. The number of events corresponding to the two mass hierarchies hypothesis and are specified using the format NO (IO). The constant density used reaches 2.8 g/cm^3 (second column). Event rates are also given for the Shen-Ritzwoller (third column) and Crustal (fourth column) earth density models. The abbreviation "cont." stands for "contamination".

	CP optimized flux	**τ optimized flux**
ν mode		
ν_e from osc.	1197	1199
$\bar{\nu}_e$ from osc.	18	11
ν_e from beam cont.	365	543
$\bar{\nu}_e$ from beam cont.	57	56
ν_μ	9660	37673
$\bar{\nu}_\mu$	741	683
ν_τ from oscillation (QEL/RES/DIS)	270 (124/62/70)	1658 (531/597/448)
$\bar{\nu}_\tau$ from oscillation	25	22
NC	8832	18126

Table 2.3: Comparison of the number of events expected at the DUNE FD for the 3.5 years staged deployment plan and neutrino beam configuration between both the CP optimized neutrino flux and the τ optimized neutrino flux. The PMNS parameters used are recalled in table 2.1. The constant earth density model and the NO hypothesis are assumed.

more ν_μ than $\bar{\nu}_\mu$, more ν_e for beam contamination than $\bar{\nu}_e$ from beam contamination and more ν_τ than $\bar{\nu}_\tau$. In [42] the DUNE collaboration used a reconstructed energy interval of [0.5 ; 10] GeV to calculate the integrated event rates. This is justified by the fact that the oscillations patterns associated to the electron neutrino appearance and muon neutrino disappearance occur in this energy window (see [42, Figures 10, 11]). However at higher energies, one would find that the wrong sign contamination (*i.e* ν_μ or ν_τ) dominates the $\bar{\nu}_\mu$ or $\bar{\nu}_\tau$ signal, which explains why in this thesis some integrated event rates are reversed between neutrinos and antineutrinos in RHC mode.

We also give in table 2.3 the expected number of events corresponding to the FHC (*i.e* neutrino) beam configuration for the 3.5 years staged hypothesis, comparing the CP optimized flux and τ optimized flux. In this table we use a constant density value of $2.8\,\mathrm{g\,cm^{-3}}$. These expected number of events associated to the τ optimized flux will later be used in the thesis whenever we suggest results based on the alternative τ optimized flux instead of the standard LBNF neutrino flux, optimized for the CP violation study. The main feature is that the use of the τ optimized neutrino beam allows to multiply the ν_τCC statistics by a factor of 6.

2.5 Detector effects

2.5.1 The smearing process

The full simulation chain used by the DUNE collaboration to produce the Technical Design Report is described in [41, Section 5.6]. Here we only summarize the relevant information to us. This simulation proceeds in three main steps:

1. Simulating neutrino interactions on argon taking into account the nuclear effects. This first step is performed by GENIE [87] v2.12.10. It takes into account the incident unoscillated neutrino flux and the neutrino cross sections of various scattering types described in 2.1 (such as quasi-elastic, resonant and deep inelastic scatterings) to predict a neutrino event rate. For each neutrino event, the generators predicts among others the particles composing the final state and their 4-momentum.

2. The final state particles of the neutrino interactions are then propagated in liquid argon using GEANT4 [97]. This steps predicts the behavior of each particle in argon, and more specifically the energy deposited in the argon medium via ionization and showers.

3. Finally the detector response must be simulated from the drift of the freed electrons (taking into account electron lifetime due to impurities and recombination) to the electronic response and the reconstruction algorithms.

The TDR used a NTuple called CAFAna to store the relevant calorimetric information used in oscillations analysis. Unfortunately to us, some kinematics we required for our analysis, such as the total momentum of the hadronic system, were not included and we could not benefit from this full simulation chain. Instead we propose to deploy a method called smearing, which starts from step 1, and accounts for steps 2 and 3 in a faster manner. Assuming some calorimetric performance of the LArTPC detectors for each type of particle (pions, protons, neutrons...) that we will describe later in this chapter, we attribute to each of them an energy resolution $\sigma(K_{true})$, where K_{true} is the true kinetic energy of the particle at the generator level.

K_{true} is biased manually to take into account the energy misreconstruction due to the detector performance following:

$$K_{smr} = K_{true}\left(1 + \text{Gaus}(0, \sigma(K_{true}))\right). \tag{2.15}$$

The smeared kinetic energy K_{smr} is calculated randomly following a normal distribution centered on K_{true} and a variance $\sigma(K_{true})$. This new kinetic energy is taken as the measured kinetic energy of the particle. Assuming a 100% efficiency on particle identification (additional smearing effects can take into account particle misidentification, which we do not include here), we deduce the smeared 3-momentum $\overrightarrow{p}_{smr}$ from the true 3-momentum $\overrightarrow{p}_{true}$ of the particle as:

$$||\overrightarrow{p}_{smr}|| = \sqrt{(K_{smr}+m)^2 - m^2} \;\Rightarrow\; \overrightarrow{p}_{smr} = \frac{||\overrightarrow{p}_{smr}||}{||\overrightarrow{p}_{true}||}\overrightarrow{p}_{true}, \tag{2.16}$$

where m is the mass of the particle. It should thus be noted that this steps assumes that the particle is correctly identified. The last step of the smearing takes into account the misreconstruction of the direction of the particle. We associate to each type of particle a direction resolution σ_{ang}. Given its true direction $\overrightarrow{p}_{true}$, we randomly generate two angles θ and ϕ in the rotated frame where the particle propagates along the z-axis:

$$\theta = \text{Gaus}(0, \sigma_{ang}) \text{ and } \phi = \text{Uniform}(0, 2\pi), \tag{2.17}$$

where $\text{Uniform}(0, 2\pi)$ generates a random angle following a flat distribution and $\text{Gaus}(0, \sigma_{ang})$ generates a random angle following a normal distribution centered on 0 with a variance σ_{ang}. The smeared direction of the spatial momentum $\hat{p}_{smr}$ in this rotated frame is:

$$\hat{p}_{sm} = \begin{pmatrix} \cos(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) \\ \cos(\theta) \end{pmatrix}, \tag{2.18}$$

and the smeared momentum direction in the laboratory frame is deduced thanks to a reverse frame rotation.

2.5.2 Particles energy losses

When a particle enters a given medium, typically a detector, it will transfer energy to that medium in a way which depends on its momentum and its electrical charge. Charged particles mainly lose energy via radiation and ionization. The first one, which dominates at higher energy, is known as the Bremsstrahlung radiation that occurs as the charged particle decelerates in the medium because of the atomic collisions. At lower energy, the ionization loss (atomic collisions) dominates. The charged particles excites atomic electrons energy states which results in the creation of pairs of electrons/ions. As far as LArTPC detectors are concerned, we mainly find three signatures of particles:

- **Tracks**. This will be the case of charged particles that live long enough to propagate on distances greater than the pitch of the detector resolution, and that leave a clear ionization signal. The particle needs to be heavy enough not to suffer from the multiple Coulomb scattering with electrons of the medium, so its mass needs to be greater than the electron one ($511\,\text{keV}/c^2$). In this case, the ionization process is the main way the particle releases

energy to the medium. Typically, for a particle with a mass equal or greater than the one of the muon ($105.6\,\mathrm{MeV}/c^2$), we can apply the Bethe-Bloch formula that describes the mean energy loss per unit of length:

$$-\langle\frac{dE}{dx}\rangle = 4\pi m_e^2 c^2 \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{nz^2}{\beta^2}\left[\ln\left(\frac{m_e\beta^2\gamma^2c^2}{I^2}\right) - \beta^2\right], \tag{2.19}$$

where m_e is the mass of the electron, e the elementary electrical charge, ϵ_0 the vacuum permittivity, n the electron density in the medium where the particle propagates, z the charge of the particle (in units of e), and I the mean excitation potential. β and γ refer to the propagating particle relativistic kinematics. It should be noted that this formula does not depend on the mass of the propagating particle, but only on its speed (β).

- **Electromagnetic showers**. At higher momentum, the energy loss by radiation becomes dominant. For the DUNE experiment, that will be the case for electrons/positrons with energies greater than 50 MeV. They start loosing energy by atomic collisions, but because of their small mass, they are easily deflected and thus emit photons (Bremsstrahlung) which have enough energy to create pairs of electron/positron. The created pair behaves in the same way and the resulting cascade is called an electromagnetic shower. It should be noted that a photon will also trigger an electromagnetic shower in LArTPC if it has sufficient energy, as well as neutral pions π_0 (since they decay into two photons about 99% of the time).
- **Hadronic showers**. An hadronic shower is merely a handful of hadrons emitted locally in space, producing a combination of tracks and electromagnetic showers (when π_0 or photons are involved). The sometimes complex topology of these showers makes them hard to reconstruct precisely.

2.5.3 Range

The treatment of heavy charged particles (charged hadrons and muons) which lose energy according to the Bethe-Bloch formula (see equation 2.19) requires that we first introduce the notion of range. The length of the track is related to the total energy deposited by the particle and for this reason it can be used to reconstruct the particle energy. For this reason, one often defines the range R of a particle as (assuming the particles has an initial kinetic energy E_0):

$$R = \rho_{Ar}\int_0^{E_0} -\frac{1}{dE/dx_t}dE, \tag{2.20}$$

where x_t is the coordinate along the particle track and ρ_{Ar} the liquid argon density. R/ρ_{Ar} is the length traveled by the particle in the medium. The integral is not trivial since dE/dx_t depends on β which is the particle speed. In the end, one finds that the ratio R/M, M being the mass of the particle, only depends on the product $\beta\gamma$, which is a kinematic term independent of the particle at play.

The NIST website [98] proposes a platform to compute the range of protons, electrons, and helium ions in various medium, including argon. We downloaded the data for protons in liquid argon, which we show in figure 2.10. The x-axis is then converted into the product $\beta\gamma = \sqrt{\left(\frac{K}{M_p}+1\right)^2 - 1}$ (noting K the kinetic energy and M_p the mass of the proton). The y-axis is

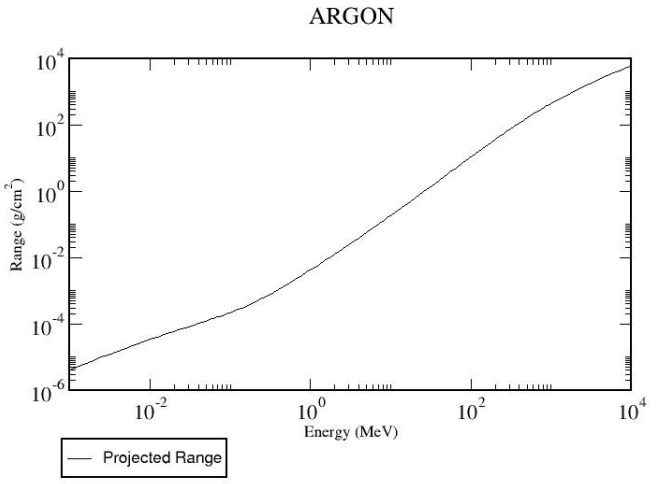


Figure 2.10: Proton range in liquid argon as a function of the kinetic energy of the proton. Numerically, a 100 MeV proton has a range of 3.614 g/cm^2. It can be converted into a length by dividing by the argon density 1.4 g/cm^3, which gives 2.6 cm.

scaled by the inverse proton mass so that the resulting curve only depends on β and γ. As a consequence, the re-scaled plot obtained for the proton can be used to deduce the range of other charged particles (such as pions) given their $\beta\gamma$ and their mass, since the ratio R/M with R defined by equation 2.20 only depends on the relativistic kinematical variables γ and β.

2.5.4 Interaction Length

Another key point about charged hadrons (not muons) is the interaction range or interaction length. It corresponds to the range for which the hadron interacts in a destructive way with the medium, producing often an hadronic shower. In other words, the charged hadron stops propagating and interacts, and reconstructing its kinetic energy relies on reconstructing the energies of the secondary particles produced by the interaction. In this case the energy resolution goes bad, but it should be noted that in LArTPC it is the only way to reconstruct the energy of neutral particles like neutrons. The interaction length of protons in liquid argon is 85.7 cm. The conversion if this length into a range allows to define the critical $\beta\gamma$ using the NIST data on proton. Protons below this value will most likely stop in the medium because of energy losses before having the chance to produce hadronic interactions. Numerically this gives $\beta\gamma = 1.037$ and it corresponds to a proton kinetic energy of 413 MeV. This interaction range applied to charged pions gives an interaction length of 18 cm (corresponding to a kinetic energy of 62 MeV). In other words, charged pions are much more likely to produce secondary interactions in liquid argon than protons. This motivates the deployment of a survival method in our smearing framework.

Survival Method: Given a charged hadron (proton, pion, kaon...) with a kinetic energy K, we compute its $\beta\gamma$ according to $\beta\gamma = \sqrt{\left(\frac{K}{M}+1\right)^2 - 1}$. We then extrapolate its predicted range R using NIST data on proton assuming this range to be particle independent. We then compute $p_{surv} = \exp\left(-R/R_{int}\right)$, where R_{int} is the interaction range of proton in liquid argon divided by its mass. As suggested by the notation, this number is interpreted as a survival probability of the particle. The greater the range, the smaller the probability the particle

survives ($p_{surv} \to 0$). Then, a random number is generated in $[0\,;\,1]$. If this number is smaller or equal to the survival probability, we consider that the particle survives and continues ionizing the medium until it stops. In that case, we apply an energy smearing in relation to the hadronic calorimetry performance of the detector on this type of particle. If the number is greater than the survival probability, we assume that the particle interacts destructively and produces an hadronic shower, and we apply a corresponding reconstruction efficiency which is much less precise.

2.5.5 Energy resolution

We have reviewed enough key ideas to define the individual particle smearing that we will use throughout the thesis. Previous experiments already assessed the LArTPC performance on reconstructing various particles energy [57, 59, 62, 60]. We will rely as much as possible on such results for the energy resolutions. If no data is available, we will refer to the DUNE Conceptual Design Report [99, Table 3.3]. Note also that the protoDUNE single phase published an overall performance assessment of the detector [68], in which a calorimetric energy response of the detector (dE/dx) was performed on pions, protons, electrons and muons, as well as an energy reconstruction evaluation of electrons using the photon detection system of the detector.

We will use an energy resolution of $\dfrac{\Delta E}{E} = \sqrt{(0.02)^2 + \dfrac{(0.15)^2}{E[\text{GeV}]}}$ for electromagnetic showers [59], which corresponds to the hypothesis used in the DUNE Conceptual Design Report for the long-baseline analysis. Charged hadrons producing secondary interactions are badly reconstructed with a resolution of $\dfrac{\Delta E}{E} = \sqrt{(0.05)^2 + \dfrac{(0.30)^2}{E[\text{GeV}]}}$. If a charged hadron does not produce a secondary interaction, its energy resolution is taken to be 10% for protons and 5% for pions. In this case the energy is measured according to the range of the particle and the resolution is limited by the spatial resolution of the detector. In [57] protons were found to be reconstructed with a much better precision of 3% for 50 MeV protons and 1% for protons with kinetic energy exceeding 200 MeV. The resolutions used in the analysis of this thesis are thus rather conservative and may underestimate the real capability of LArTPC detectors as far as protons are concerned. The different energy and angle smearing values are gathered in table 2.4.

Neutrons need a special care because they are neutral and they are detected only when they interact destructively in the medium with hadronic interactions. We assume a 10% chance that the neutron goes undetected regardless of its kinetic energy and are otherwise recontructed with an energy resolution of $\dfrac{\Delta E}{E} = \dfrac{0.40}{\sqrt{E[\text{GeV}]}}$.

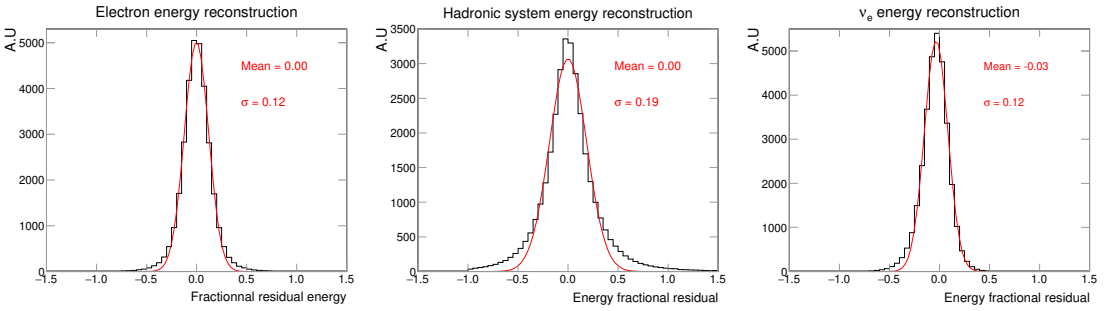
Finally, an energy threshold is applied to the kinetic energy of the particles. If the smeared energy of a particle (following equation 2.15) is below the corresponding threshold, the particle is undetected.

2.5.6 Smearing assessment

Because we use a simpler version of the reconstruction effects than the one used in the Technical Design Report of DUNE, we need a validation of the smearing method deployed. On page 5-152 of the second volume [41, p.5-152], one finds the energy fractional residual of the ν_μ and ν_e. The energy fractional residual is defined as $(E_{reco} - E_{true})/E_{true}$. We have used the same simulated

Particle	Detection Threshold (MeV)	σ	Angular Resolution (°)
$\mu^{\pm}$	30	5%	1
$e^{\pm}$, π^0, γ (electromagnetic showers)	10	$\sqrt{(0.02)^2 + \frac{(0.15)^2}{E[\text{GeV}]}}$	1
Protons	50	if survives: 10% if interacts: $\sqrt{(0.05)^2 + \frac{(0.30)^2}{E[\text{GeV}]}}$	5
$\pi^{\pm}$	20	if survives: 5% if interacts: $\sqrt{(0.05)^2 + \frac{(0.30)^2}{E[\text{GeV}]}}$	1
Neutrons	50	if detected: $\frac{0.4}{\sqrt{E[\text{GeV}]}}$	5
Others	50	$\sqrt{(0.05)^2 + \frac{(0.30)^2}{E[\text{GeV}]}}$	5

Table 2.4: Reference energy and angular smearing applied to particles.



(a) (b) (c)

Figure 2.11: The reference smearing method defined in table 2.4 applied to a sample of oscillated $\nu_\mu \to \nu_e$ events. We show the electron energy fractional residual (a), the hadronic system energy fractional residual (b) and the electron neutrino energy fractional residual (c). Oscillation parameters are taken from table 2.3. Histograms are fitted with a Gaussian function.

sample of oscillated $\nu_\mu \to \nu_e$ CC interactions as in the TDR, using the oscillation parameters in table 2.3. We assess the energy reconstruction performance on the emitted electron, the hadronic system and the electron neutrino. The reconstructed (or smeared) neutrino energy is merely taken as the sum of the smeared electron energy and the smeared hadronic energy.

We show the results of this crosscheck in figure 2.11, as well as Gaussian fits to assess the energy resolution. The electron neutrino energy resolution (right) is in agreement with [41, Fig. 5-11]. In the DUNE TDR it has been reported that the electron energy resolution is $4\% \bigoplus 9\%\sqrt{E[\text{GeV}]}$ (where $\bigoplus$ means quadratic sum), which reaches 8% for a 2 GeV electron (which is the expected typical energy of electrons produced by ν_eCC interactions in DUNE). It seems that we slightly underestimate the energy resolution of the electrons, as the Gaussian fit gives a variance of 0.12. We have however no element of comparison for the energy resolution of the hadronic system.

2.6 τ decays

GENIE uses the TAUOLA [100] library to decay the charged τ lepton. TAUOLA handles about 20 leptonic and semi-leptonic decay modes, taking into account the τ spin polarization effects. It was largely used by the so-called B-factories, experiments tuned to produce B mesons, which actually produced many τ leptons as well, to refine the generator predictions. Babar and Belle were the two main B factories operating at the beginning of the 2000s, using electron-positron collisions. An exhaustive discussion on the results established by the two experiments on the τ lepton is available in [101, chap.20]. TAUOLA met numerous corrections since its first release in 1991, see for instance [102].

The identification of ν_τ interactions in DUNE will rely on the search of the daughter particles of the lepton τ since the τ itself will decay in flight before leaving a visible track. A given τ decay mode will orientate a dedicated analysis to look for the τ signature. A first overview of the τ decay modes is required in order to understand how to look for τ decay signatures. We give in table 2.5 the main branching ratios taken from the particle data group (PDG).

	$\tau \to e$	$\tau \to \mu$	$\tau \to \rho \to \pi^{\pm}\pi_0$	$\tau \to \pi$	$\tau \to 3\pi$
Branching ratio (%)	17.83 ± 0.04	17.41 ± 0.04	25.52 ± 0.09	10.83 ± 0.06	9.31 ± 0.06

Table 2.5: PDG branching ratios of the main τ decay modes. 35.24% of the decays are leptonic and 64.76% of the decays are semi-leptonic. In the notations the final state $\nu_\tau/\bar{\nu}_\tau$ and the electrical charge are systematically omitted.

We mentioned in 1.4.3 that in this thesis we will deploy analysis to search for ν_τ charged current interactions in DUNE following the philosophy of "1 τ decay mode = 1 dedicated analysis". An interesting τ decay mode is the $\tau \to e$ that Albright and Shrock had already spotted in 1979 and which was extensively used in the NOMAD experiment. For DUNE, the $\nu_\mu \to \nu_e$ oscillation channel will be sub-leading, and this makes the search for ν_τ interactions via the $\tau \to e$ decay more favourable than for the $\tau \to \mu$ decay since the ν_e background will be smaller than the ν_μ background. The $\tau \to \rho$ decay mode presents the largest branching ratio and the possibility to exploit the kinematic signature of the ρ resonance with the invariant mass of the two final state pions, so this makes it an interesting decay mode to look at as well. We also show the $\tau \to 1\pi$ decay mode which has the advantage of having a rather exclusive final state. The $\tau \to 3\pi$ decay mode can also be interesting in the sense that the corresponding ν_τ events would have a rather high final state hadronic activity (3 charged pions in addition to the hadronic system).

2.7 Convolutional Visual Network (CVN)

The DUNE collaboration relied on a convolutional neural network (CVN) to tackle the neutrino flavour identification at the far detector simulated events [103]. The precise spatial resolution of LArTPC and the precise track reconstruction justifies the use of such tools to classify neutrino interactions. The CVN itself is a CNN (Convolutional Neural Network), and applies a series of filters and convolutions to the event pictures offered by the LArTPC technology to identify structures in neutrino interactions and finally to classify the neutrino events. The DUNE collaboration trained the CVN using in inputs sets of three pictures per neutrino event corresponding to the three views of the LArTPC single phase module. These images correspond to the hit level reconstruction of the simulated events which means that no pre-reconstructed structures (tracks, electromagnetic showers, vertices...) were given as input to the CVN. Only one output

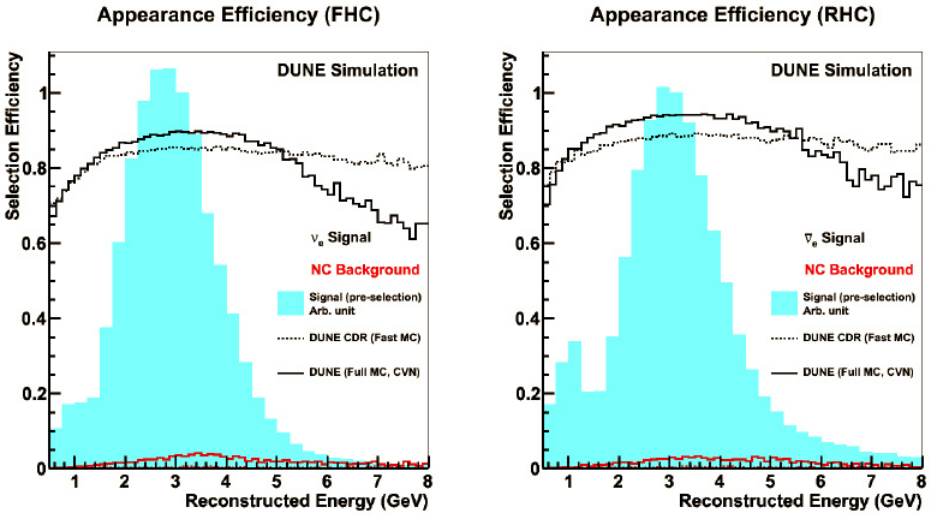


Figure 2.12: Selection efficiency of ν_eCC events in FHC running mode (left) and RHC running modes (right) as a function of reconstructed neutrino energy, as can be found in [103, fig.9]. The black histogram is the CVN response, the dashed histogram the predictions made in the DUNE CDR [99] and the blue histogram the $\nu_e/\bar{\nu}_e$ event rates.

layer of the CVN is used in this thesis. It contains four neurons associated to the four neutrino interactions (ν_eCC, ν_μCC, ν_τCC and NC). The neurons were used by the collaboration to perform the events classificationwhich allowed to produce the DUNE sensitivity results in [42]. Each of these four output neurons contains a number between 0 and 1 and all four sum to 1, so it would be tempting to call them probabilities. A more dedicated terminology is to call them scores. The collaboration scanned the ν_e score cut values to use in order to maximize the sensitivity on the PMNS parameter δ_{CP}. This value was found to be 0.85. The corresponding ν_eCC selection efficiency selection peaks at 90% close to the reconstructed neutrino energy where the ν_eCC event rate is maximal (about 3 GeV). The 90% selection efficiency peak was observed both for the FHC (forward horn current) and RHC (reverse horn current) beam configurations. We report the ν_eCC CVN efficiency as a function of the reconstructed neutrino energy as can be found in [103, fig.9] in figure 2.12 for a ν_eCC score of 0.85.

It should be noted that the CVN also has a one neuron output to assess if the event is more neutrino-like or antineutrino-like, but this score was not used. Thus, the ν_eCC classification gathers ν_eCC and $\bar{\nu}_e$CC events. Distinguishing between the two would require to be able to tag the charge of the outgoing lepton or the one of the hadronic system, and in LArTPCs this can not be done easily.

No significative effects on the DUNE scientific program was observed as far as the muon flavour identification is concerned, so the collaboration decided to use a score of 0.5 to classify an event as a ν_μCC. The τ flavour was not discussed, which sounds reasonable as ν_τCC will contribute in minority to the ν_eCC background (when the charged lepton τ decays electronically), the relevant flavour of CP violation studies. But as for curiosity at least, it is interesting to have a look at the CVN performance on the ν_τCC sample. We looked into the simulated beam ν_τCC events (FHC mode, excluding the $\bar{\nu}_\tau$CC) and their corresponding CVN classification scores. We gather the observed results in table 2.6 where we used a score of 0.5 for the ν_τCC identification. We observe that the global efficiency on the τ selection reaches 15%, and that most of the events (41%) are classified as neutral currents. This efficiency is quite low and indicates that tagging the τ flavour requires more sophisticated tools than mere visual characteristics.

CVN classification	p(ν_eCC)>0.85	p(ν_μCC)>0.5	p(ν_τCC)>0.5	p(NC)>0.5	else
Fraction (%)	9.9	16.0	15.4	41.0	17.7

Table 2.6: CVN classification of the sample of simulated beam ν_τCC events produced by the DUNE collaboration. Most of the events are classified as neutral currents (41%), as most of ν_τ do not have charged lepton in the final state. In about 15% of the cases, the CVN correctly identifies the τ flavour.

τ decay mode (BR %)	p(ν_eCC)>0.85	p(ν_μCC)>0.5	p(ν_τCC)>0.5	p(NC)>0.5	else
$\tau^- \to e^-$ (17.8)	45.7	0.7	10.4	16.7	26.5
$\tau^- \to \mu^-$ (17.5)	0.1	66.2	5.5	17.5	10.7
$\tau^- \to \rho^-$ (22.4) [1]	3.6	4.7	21.2	53.5	17.0
$\tau^- \to 2\pi^- 1\pi^+$ (9.3)	0.6	12.4	19.8	53.5	13.7
Others (33.0)	2.6	6.3	18.3	54.6	18.2

Table 2.7: CVN classification of the sample of simulated beam ν_τCC events produced by the DUNE collaboration where we split the CVN response with respect to four τ decay modes. The chart should be read as follow: one row corresponds to a given τ decay mode (the row "Others" contains several hadronic decay modes). The observed branching ratio (BR) of this decay mode is given in brackets. For this specific decay mode, the rest of the row indicates the CVN flavour classification percentages using the scores indicated in the first row. As for illustration, the first value to be read for the row $\tau^- \to e^-$ is 45.7. It means that 45.7% of ν_τCC (for which the τ^- decays into an electron) have a CVN score greater than 0.85.

A complementary approach is to distinguish the CVN response for various τ decay modes. Obviously, the CVN response is expected to be sensitive to the presence of a charged lepton in the final state. We gather our findings in table 2.7 looking at the following τ decay modes: $\tau^- \to e^-$, $\tau^- \to \mu^-$, $\tau^- \to \rho^-$, $\tau^- \to 2\pi^- 1\pi^+$. We also indicate for each τ decay mode the effective branching ratio found in the simulation sample. Mind the reading of the table which is not obvious at first sight, detailed explanations are given in the caption. One interesting feature we want to stress is that the correct τ identification occurs mainly for hadronic decays (about 20% each). We guess that in these cases the CVN noticed an "unusual" large hadronic activity as the decaying τ provides additional hadrons to the final state. One can note also that for the leptonic decays, the CVN tends to classify the ν_τCC interaction as the corresponding charged current flavour ν_eCC or ν_μCC.

⋆ ⋆ ⋆

[1] This number is in contradiction with the PDG BR which indicates $25.52\% \pm 0.09$. Statistical fluctuations (at the level of 0.1% for us) can not explain the discrepancy. This issue of GENIE v2.12.00 was reported to the GENIE collaborators.

Chapter 3

The $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ decay mode analysis

Contents

In this chapter we propose a quantitative analysis to identify τ neutrino charged current interactions, using the $\tau \to e$ decay channel at the level of the simulation files produced for the DUNE Technical Design Report.

The branching ratio of the charged lepton τ decaying into an electron is 17.83% so this channel deserves to be studied. In addition electrons have a rather identifiable final state signature in LArTPC detectors (electromagnetic showers). Conversely the muonic decay mode (though it has a similar branching ratio) is expected to have a much larger background due to the muon neutrino charged current interactions which makes less favourable to search for ν_τ events. The last section will still be dedicated to the muonic decay mode discussion.

The analysis suggested in this chapter will rely on kinematics in the transverse plane of the neutrino interactions to discriminate between ν_τCC($\tau \to e$) and the main associated backgrounds which are the ν_eCC interactions. A likelihood approach inspired by the pioneering work of the NOMAD experiment is developed. Machine learning techniques are also explored.

3.1 Motivations and kinematics of interactions

3.1.1 Motivations

The first proposal for identifying ν_τCC interactions in beam experiments was suggested by Albright and Shrock in 1977 [22], few years after the discovery of the charged τ lepton. The feasibility of the ν_τ discovery relied on a good knowledge of the relative beam neutrino flavour composition (ν_e, ν_μ, ν_τ) and on having detectors able to precisely reconstruct the directions of the final state particles in order to spot unusual large amounts of transverse missing momentum. It should be noted that that thanks to neutrino oscillations, one can completely neglect the very tiny ν_τ production in the beam and exploit the large $\nu_\mu \to \nu_\tau$ oscillation probabilities (see figure 2.8). The authors proposed to exploit the leptonic decays $\tau^- \to \alpha^- + \bar{\nu}_\alpha + \nu_\tau$, with $\alpha = e,\ \mu$. The ν_eCC and ν_μCC interactions would be the associated backgrounds.

These leptonic decays of the τ imply two undetected neutrinos in the final state (by the way that is a serious issue for ν_τ energy reconstruction, because there will be a large fraction of missing energy in the final state). The key concept is to exploit the kinematics in the transverse plane, because reconstructing the kinematics along the beam neutrino axis requires knowing the impinging neutrino energy. The transverse momentum of the initial state is instead zero if we assume that the nucleon hit by the neutrino is at rest. This is approximately true by neglecting the struck nucleon Fermi momentum inside the argon nucleus. Due to momentum conservation the transverse momentum of the final state should then be also null. But as the two final state neutrinos of the interacting ν_τ ($\tau \to e$) carry undetected energy with them, a large missing momentum is expected in the transverse plane. We divide the final state of the neutrino interaction into the leptonic (where the only visible particle is the electron in ν_τCC events) and the hadronic momenta noted $\overrightarrow{p}^{(tr)}_{lep}$ and $\overrightarrow{p}^{(tr)}_{had}$. Thus the transverse missing momentum $\overrightarrow{p}^{(tr)}_{miss}$ is defined as:

$$\overrightarrow{p}^{(tr)}_{miss} = -(\overrightarrow{p}^{(tr)}_{lep} + \overrightarrow{p}^{(tr)}_{had}). \tag{3.1}$$

(a) ν_τCC ($\tau \to e$)

(b) ν_eCC

Figure 3.1: Schematic views of ν_τCC ($\tau \to e$, left) and ν_eCC (right) interactions in the transverse plane. The leptonic system is shown in blue, while the hadronic system (composed of an arbitrary numbers of hadrons) in black. These two systems tend to go in opposite directions because of momentum conservation. Undetected neutrinos of ν_τCC strongly bias this configuration, resulting in a larger transverse missing momentum (green).

The subscript "lep" refers either to muon or electron, because both the $\tau \to e$ and $\tau \to \mu$ decay modes can be treated similarly. However we emphasize the electron case because of the smaller ν_eCC background in comparison to ν_μCC (see table 2.2). We draw schematic views of ν_eCC and ν_τCC ($\tau \to e$) in the transverse plane of the interaction in figure 3.1 to further illustrate the principle of kinematical identification. Energy momentum conservation implies the sum of momenta of final state particles to be 0, because the impinging neutrino did not bring energy in this plane. Thus the leptonic (blue) and hadronic (black) systems should be back to back in the transverse plane. However, in the ν_τCC case, the undetected neutrinos bias the observation of the leptonic system resulting in a large transverse missing momentum (green). ν_eCC interactions are expected to present smaller transverse missing momentum related to reconstruction errors or undetected particles such as neutrons. One should note that there exist irreducible sources of transverse missing momentum (such as Fermi momentum or intra-nuclear scatterings of emitted hadrons) thus limiting the power of this method.

In addition to the three transverse momenta moduli, we define as well the three angles $\phi_{hl}^{(tr)}$, $\phi_{hm}^{(tr)}$ and $\phi_{lm}^{(tr)}$. The first designating the angle between the transverse hadronic and the transverse leptonic momenta. The second designates the angle between the hadronic system transverse momentum and the transverse missing momentum. The third designates the angle between the transverse lepton momentum and the transverse missing momentum.

3.1.2 Kinematical variables at play

We studied height kinematical variables to help in discriminating the ν_τCC ($\tau \to e$) signal against ν_eCC background. Namely, they are:

- The electron kinetic energy K_{lep}.
- The electron transverse momentum modulus $p_{lep}^{(tr)}$.
- The hadronic transverse momentum modulus $p_{had}^{(tr)}$.

- The missing transverse momentum modulus $p_{miss}^{(tr)}$
- The transverse angle $\phi_{hl}^{(tr)} = (\overrightarrow{p}_{had}^{(tr)}; \overrightarrow{p}_{lep}^{(tr)})$.
- The transverse angle $\phi_{hm}^{(tr)} = (\overrightarrow{p}_{had}^{(tr)}; \overrightarrow{p}_{miss}^{(tr)})$.
- The transverse angle $\phi_{lm}^{(tr)} = (\overrightarrow{p}_{lep}^{(tr)}; \overrightarrow{p}_{miss}^{(tr)})$.
- The asymmetry ratio $p_{asym} = \dfrac{p_{lep}^{(tr)} - p_{had}^{(tr)}}{p_{lep}^{(tr)} + p_{had}^{(tr)}}$

The last variable p_{asym} was not found to be of substantial help, and will thus be omitted in the following of the chapter. The case of K_{lep} is interesting since it is the only variable not defined in the transverse plane, and a dedicated discussion shall be given.

3.2 Kinematical distributions

In this section we discuss the main features of the kinematical variables previously exposed. Our signal sample is composed of the oscillated beam $\nu_\mu \to \nu_\tau$ neutrinos, and our background of the oscillated $\nu_\mu \to \nu_e$ neutrinos. It should be noted that the beam muon neutrino of DUNE is expected to have an intrinsic ν_e contamination (see figure 2.5). A non negligible fraction of ν_eCC events will be associated to the survival of these electron neutrinos. However, in a first place we only study the oscillated $\nu_\mu \to \nu_e$ neutrinos, and add in a second place this second source of background. We will conventionally refer as "oscillated" electron neutrinos to the set of ν_eCC interactions due to the oscillations $\nu_\mu \to \nu_e$, and to "beam" electron neutrinos the ones associated with the survival of electron neutrinos contaminating the muon neutrino flux at the production point.

Finally, an important point to mention is that we assume a neutrino beam running mode, and use the Monte Carlo truth of simulated events to reject the charged current interactions of antineutrinos. We could argue that the LArTPC technology does not allow for a good distinction between ν and $\bar{\nu}$ charged current events, so we should rather consider our signal to be $\nu_\tau(\tau^- \to e^-) + \bar{\nu}_\tau(\tau^+ \to e^+)$ and the background composed of $(\nu_e + \bar{\nu}_e)$CC interactions. While this statement is true, we decide to restrict the analysis to ν events for the reason that the final state composition of ν and $\bar{\nu}$ charged current differ. For instance, a quasi-elastic charged current neutrino interaction produces a proton, while the antineutrino quasi-elastic produces a neutron, and both particles are not reconstructed with the same efficiency in the LArTPC detectors. For this reason, we find it rather instructive to separate ν and $\bar{\nu}$ in a first place, even if a complete analysis would require to take both contributions into account.

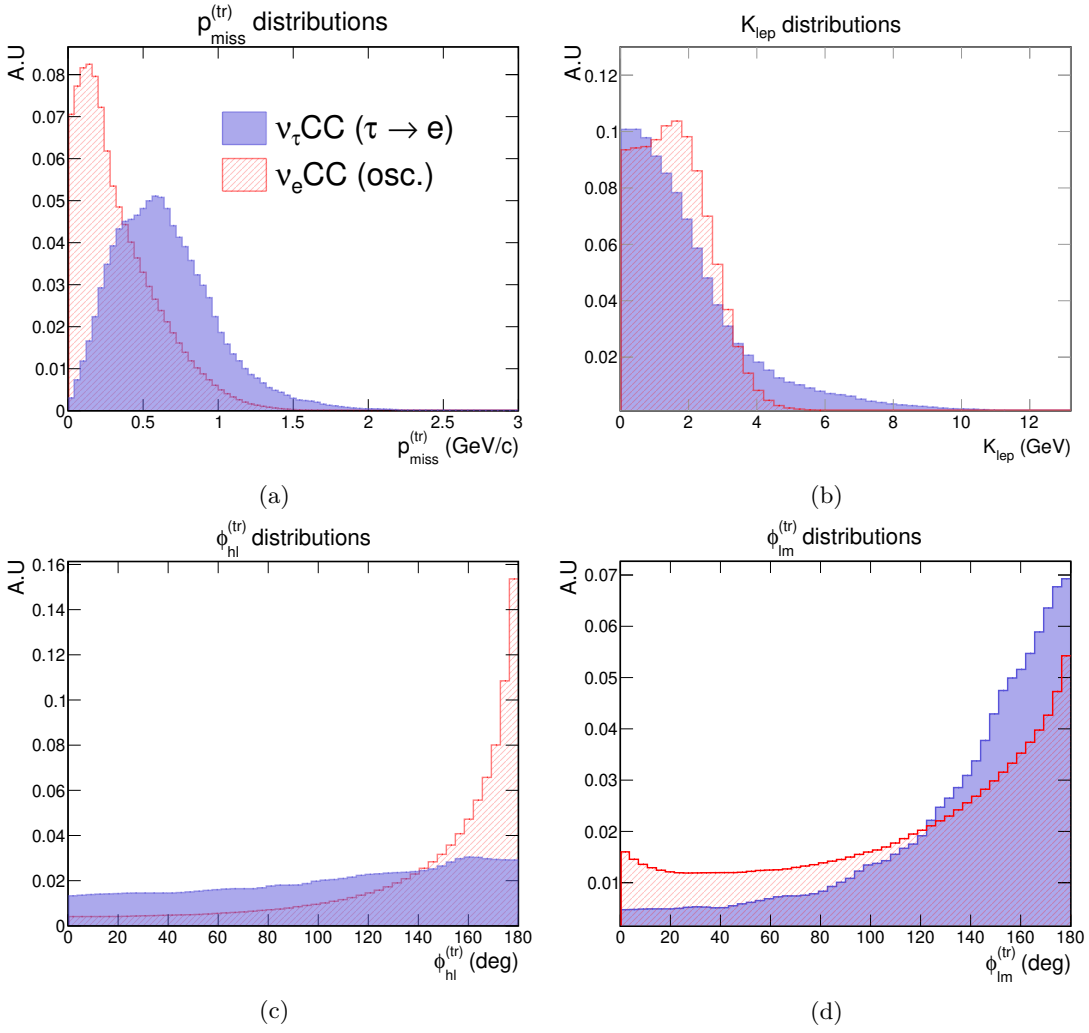
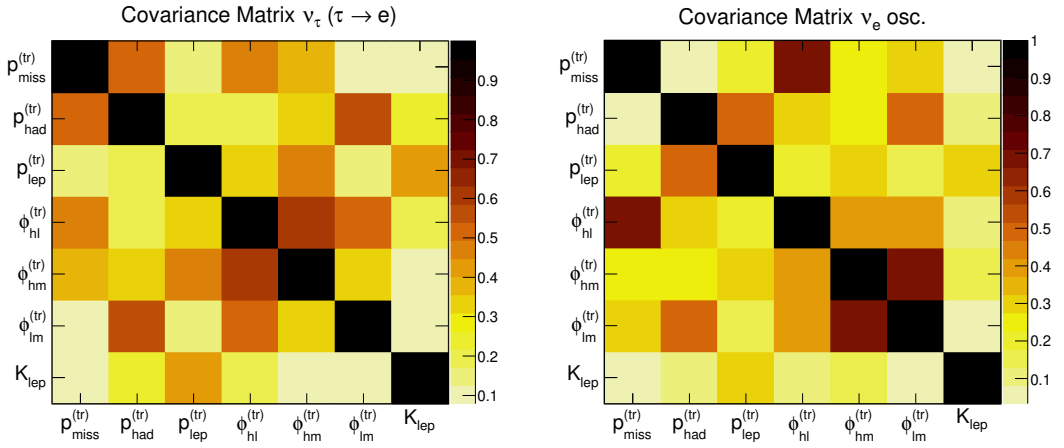


Figure 3.2: Kinematical distributions of ν_τCC ($\tau \to e$) events (blue) versus ν_eCC events (red) coming from the $\nu_\mu \to \nu_e$ oscillations. Top left: transverse missing momentum. Top right: kinetic energy of the electron. Bottom left: transverse angle $\phi_{hl}^{(tr)}$. Bottom right: transverse angle $\phi_{lm}^{(tr)}$.



(a) (b)

Figure 3.3: Correlation matrices of the kinematical variables defined in 3.1.2 for signal (left) and background (right).

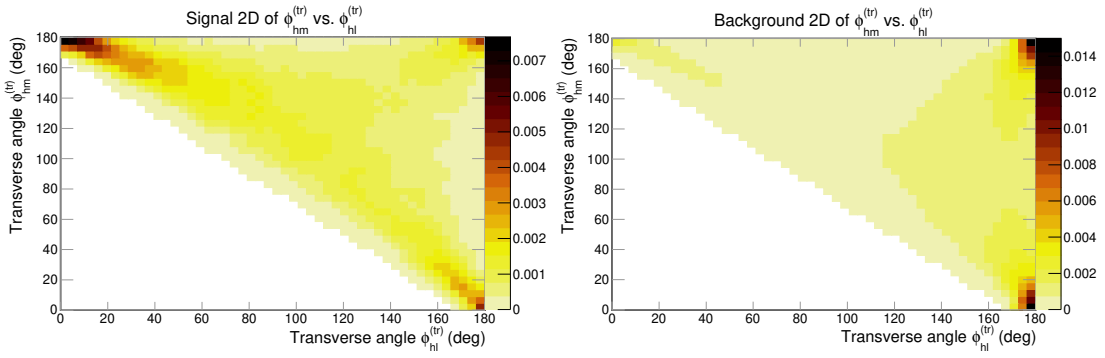
3.2.1 A first overview

We plot in figure 3.2 kinematical distributions associated to the ν_τCC interactions (blue) and oscillated ν_eCC interactions (red) at the MC truth level of the simulation files generated for the DUNE TDR (which used GENIE v2.12.00). We have omitted neutrons in the reconstruction of the final state to avoid simulating a too ideal detector. The distributions were obtained with a ν_τCC ($\tau \to e$) sample of $\mathscr{O}(30000)$ events and an oscillated ν_eCC sample of $\mathscr{O}(700000)$ events. The events are individually weighted by the corresponding oscillation probability calculated following the discussion in 2.3, assuming a constant matter density of $2.8\,\mathrm{g/cm^3}$. The oscillation parameters are taken from table 2.1, which assumes normal ordering. Once obtained, the distributions are smoothed and normalized to unity in order to be interpreted as (discrete) probability density functions.

As expected, the transverse missing momentum is larger for the signal than for the background. One can observe that the electrons coming from ν_τCC interactions have a slight higher energy. The $\phi^{(tr)}_{hl}$ distribution of ν_eCC peaks at 180°, indicating that in the transverse plane the leptonic and hadronic momenta are correlated and point towards opposite directions. The corresponding signal distribution is rather flat. For completeness we have shown the $\phi^{(tr)}_{lm}$ distributions, which do not show special discriminating power.

These distributions can now serve as individual probability density functions to decide whether a given event should be classified as a ν_τ or ν_e. A more powerful discrimination can be obtained by looking at the correlations between these variables. We sketch in figure 3.3 the correlation matrices (signal and background) for the set of kinematical variables used. The variables $p^{(tr)}_{had}$ and $p^{(tr)}_{miss}$ as well as $\phi^{(tr)}_{hl}$ and $\phi^{(tr)}_{hm}$ show strong correlations for the signal. The variables $p^{(tr)}_{miss}$ and $\phi^{(tr)}_{hl}$ as well as $\phi^{(tr)}_{hm}$ and $\phi^{(tr)}_{lm}$ show strong correlations for the background. These observations orientate the choice of 2-dimensional distributions to rely on in order to optimize the signal/background separation power.

As for illustration, we display in figure 3.4 the 2-dimensional distributions $[\phi^{(tr)}_{hl}\,;\,\phi^{(tr)}_{hm}]$, comparing the signal and background favoured regions. The background events populate mostly the region of $\phi^{(tr)}_{hl} \simeq 180°$, which was already observed in the 1-dimensional distribution. Now we add



(a) (b)

Figure 3.4: 2-dimensional distributions of the transverse angles $\phi_{hl}^{(tr)}$ and $\phi_{hm}^{(tr)}$ for the signal (left) and background (right). One can observe that only the signal populates the region $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}] \simeq [0\,;\,180]°$, which corresponds to hadronic and leptonic systems being parallel in the transverse plane, while the transverse missing momentum is anti-parallel. This specific configuration is made possible thanks to the two undetected neutrinos which carry a significant fraction of missing momentum.

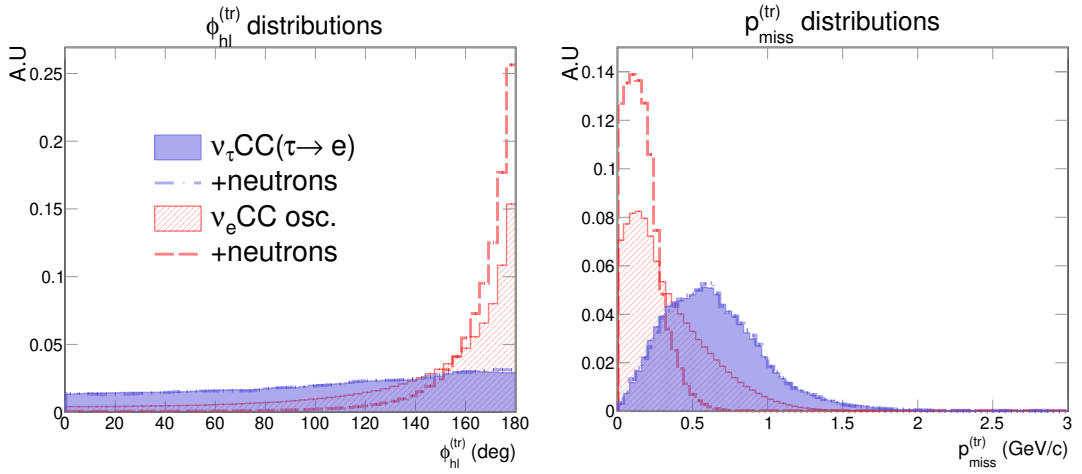
the information that background events favour the regions $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}] = [180\,;\,0]°$ and $[180\,;\,180]°$. The first corresponds to hadronic and leptonic systems going in opposite directions in the transverse plane, while the missing momentum is mostly carried by particles of the hadronic system, like neutrons, resulting in $\phi_{hm}^{(tr)} \simeq 0$. The second case is a bit more subtle. We find that these events are correlated with a low transverse missing momentum. In this case $\overrightarrow{p}_{lep}^{(tr)}$ and $\overrightarrow{p}_{had}^{(tr)}$ are in a back to back configuration and almost compensate. The small $\overrightarrow{p}_{miss}^{(tr)}$ is then either parallel (antiparallel) to $\overrightarrow{p}_{had}^{(tr)}$ resulting in $\phi_{hm}^{(tr)} = 0°$ ($\phi_{hm}^{(tr)} = 180°$).

Finally, one can observe that only the signal populates the region $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}] \simeq [0\,;\,180]°$. It is a configuration where the transverse leptonic and hadronic momenta are parallel, while the missing momentum points in opposite direction. Such a configuration is allowed thanks to the two final state neutrinos of the $\nu_\tau(\tau \to e)$ interaction. It makes it a powerful region to look for signal events with a low background contamination.

3.2.2 Impact of neutrons

The results discussed above were obtained by considering that all particles in the final state (except neutrons and neutrinos) could be ideally reconstructed. Such a situation is not realistic, but is useful, at some extent, in order to assess the best achievable performance. Still we have decided not to take into account the neutrons in order not to obtain too optimistic results. We propose in this section, as for curiosity, to include the neutrons and look at the effect on the previously exposed kinematical distributions. We expect them to display more discriminating power, since neutrons tend to add missing energy to the hadronic system, and thus compete with the missing energy carried away by the neutrinos of the $\nu_\tau(\tau \to e)$ interactions.

We present the results on the transverse missing momentum and the transverse angle $\phi_{hl}^{(tr)}$ in figure 3.5, which were the most affected variables. The first observation is that only the ν_eCC distributions are affected, which indicates that the two undetected neutrinos of the ν_τ events dominate in terms of transverse missing momentum, neutrons playing a secondary role.



(a) (b)

Figure 3.5: Same as figure 3.2, but for which the final state neutrons of the hadronic system were included. This gives an illustration of the impact of taking neutrons into account when building the kinematical variables. Associated distributions are dashed, while the filled ones are the previous ones (see figure 3.2), obtained without the neutrons. We chose to illustrate the transverse angle ϕ_{hl} and the transverse missing momentum, which were the most affected and it should be noted that only the ν_eCC (background) distributions are affected.

Among the ν_eCC kinematics, the $\phi_{hl}^{(tr)}$ distribution narrowed toward 180° and the $p_{miss}^{(tr)}$ narrowed toward 0, which exacerbated the observations made in figure 3.2. This would result in a easier background rejection. We stress the fact that there still is missing momentum even when taking into account the neutrons. The residual transverse missing momentum ($\mathcal{O}(0.2\,\mathrm{GeV}/c)$) is attributed to the Fermi momentum of the target nucleon and final state hadrons re-interactions inside the nucleus. As a result, their initial energy and direction might get biased.

We also observe (but we do not show it here) that when taking into account neutrons, the $\phi_{lm}^{(tr)}$ distribution of ν_eCC (red) shown in figure 3.2 (bottom right) flattens, while without the neutrons it tends to favour the region $\phi_{lm}^{(tr)} = 180°$. This is a configuration where the transverse missing and leptonic momenta are anti parallel. We interpret this result by inferring that without biasing the reconstructed hadronic system by omitting the neutrons, there is no more preferred orientation between the transverse leptonic momentum and the transverse missing momentum. Indeed in this case the Fermi momentum is the dominant transverse missing momentum contributor, and the initial direction of the hit nucleon is random with respect to the neutrino beam direction.

3.2.3 Two background sources

It is expected at the DUNE far detectors site that the ν_eCC interactions have two components. The main one, as suggested in table 2.2, is attributed to the $\nu_\mu \to \nu_e$ oscillation channel (peaking at about 10%) of the originally produced muon neutrinos. As the produced muon beam is expected to suffer from a 1% electron flavour contamination at the emission point, one should expect ν_eCC interactions due the survival of this contamination, which we refer to as "beam" ν_e. We found $\mathcal{O}(10\,000)$ such events in the simulation files of the DUNE TDR. Since this background component accounts for a smaller ν_e event rate than the oscillated ν_e (see table 2.2), we decided to perform a signal/background analysis using the oscillated ν_e sample as the only

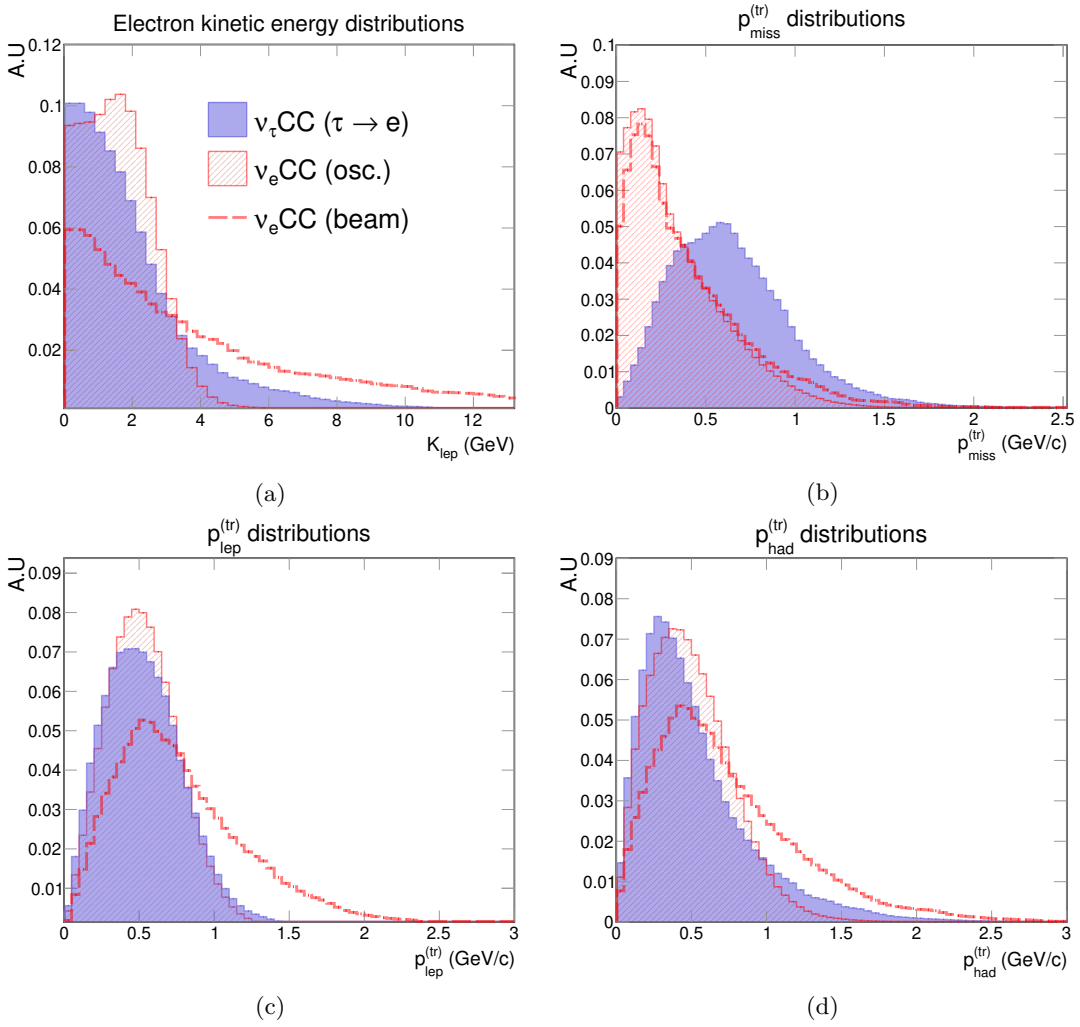


Figure 3.6: Comparison of kinematical distributions of oscillated ν_eCC (red filled) and beam ν_eCC (dashed line), the signal remaining in blue. The main difference is that the beam ν_e energy distribution stands at higher energy, and consequently so does the kinetic energy of the electron. We noted also that the transverse hadronic and leptonic momenta were shifted to higher values. However, the transverse missing momentum is not much affected.

source of background. The beam ν_e will later be tested using the hypothesis of ν_τ and oscillated ν_e. The survival probability of $\nu_e \to \nu_e$ oscillations is taken into account in the analysis.

Beam ν_e events are expected to have higher energy than ν_e events coming from the $\nu_\mu \to \nu_e$ oscillations, so a consequence the electrons produced will also have higher energy. We show this property in figure 3.6a, where the reader should focus on the red filled (oscillated ν_e) and red dashed line (beam ν_e) distributions, normalized to unity. As a consequence, the beam ν_e events will also bring more energy in the transverse plane with respect to the oscillated ν_e, which we show on 3.6c and 3.6d. However, the transverse missing momentum is not much affected (see 3.6b). We observed that the transverse angles were not much affected neither.

The electron kinetic energy distribution drastically changes between oscillated and prompt beam electron neutrinos, and it has an important drawback for the following. A careful look at figure 3.6a reveals that there is a fraction of ν_τCC events for which the electron has an energy greater than 5 GeV while the electrons of oscillated ν_e barely have any. This makes the region

above 5 GeV very signal like if we restrict ourselves to the oscillated ν_e as the only source of background. However, this same region is very populated by beam electron neutrino events.

Because we have decided to optimize the likelihood analysis on the basis of using only the oscillated ν_e as background (they indeed are the main component), we should avoid using the kinetic energy of the electron as a discriminating variable. It might be efficient at selecting the ν_τ events for which the electron kinetic energy is above 5 GeV, but afterward beam ν_eCC will be selected as signal as well, making the electron kinetic energy a counter-productive variable.

3.3 Likelihood analysis

3.3.1 Setting the framework

Once the 1-dimensional and 2-dimensional distributions of the kinematical variables defined in 3.1.2 built both for our signal sample and our background sample, we can then compare the likelihood of a given event of being signal or background. Given a reference kinematical variable, we can compute the likelihood L_S of the event to be signal, and the likelihood of the event L_B to be background. We then compute the log-likelihood ratio:

$$\mathscr{L} = \log\left(\frac{L_S}{L_B}\right). \tag{3.2}$$

Note that the background hypothesis is: "the event is a ν_eCC (oscillation) event". We presented that there exists another source of background in 3.2.3, namely the beam ν_eCC. From a picky point of view, we are not entirely rigorous here, the kinematical distributions of the background should take into account both sources and weight them according to their relative proportion expected at the far detector. However, we have chosen to optimize the likelihood analysis taking as the background hypothesis the oscillated ν_eCC only. The likelihood of the beam ν_eCC will be computed using the same distributions.

The log-likelihood ratios are calculated for both 1-dimensional and 2-dimensional variables. If the associated kinematical variable(s) to L_S or L_B were such that one of the two vanishes (or both), the minimal value of the corresponding distribution is used. Using the MC truth, we can display the log-likelihood ratio distributions of all 1-dimensional and 2-dimensional variables, and isolate the most powerful ones. We can them combine two kinematical variables (whether 1d or 2d) "i" and "j" to build an extended log-likelihood ratio as:

$$\mathscr{L}^{extended} = \log\left(\frac{L_S^{(i)} \times L_S^{(j)}}{L_B^{(i)} \times L_B^{(j)}}\right). \tag{3.3}$$

This definition is however not entirely rigorous. The likelihood of two variables is given by the product of the individual likelihoods if the two variables are mathematically independent. It is unlikely that this criteria is satisfied for variables emerging from the same neutrino event. Variables can at best be uncorrelated, but they describe the same physical process. We should thus employ the terminology "pseudo-likelihood", but we will not for sake of simplicity. The reader can think of this method as a more sophisticated way of applying kinematic cuts. We will merely avoid using several times the same information when combining different likelihood

functions. Since we restrict ourselves to transverse plane momenta (moduli and relative orientation), and that they are connected by the relation 3.1, it results we exploit five degrees of freedom. It implies we should avoid using more than 5 variables, otherwise we would only insist on some already used information without bringing new information to the problem.

We have found that the use of the correlation of $[p_{miss}^{(tr)}\,;\,p_{lep}^{(tr)}]$ combined with $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}]$ led to optimal separation results. It is to be noted, however, that other combinations let to very similar results. That is the case if we use $[p_{had}^{(tr)}\,;\,p_{miss}^{(tr)}]$ instead of $[p_{lep}^{(tr)}\,;\,p_{miss}^{(tr)}]$ and $[\phi_{hl}^{(tr)}\,;\,\phi_{lm}^{(tr)}]$ instead of $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}]$. For the reason exposed in 3.2.3, we will not use the kinetic energy of the electron in the likelihood analysis.

3.3.2 Efficiencies and CVN bias

In this section we present the log-likelihood ratio distributions of the signal ($\nu_\tau(\tau \to e)$) and background (oscillated ν_e) sets of events. We normalize the distributions to unity. It is expected that signal events have positive log-likelihood ratios while background events have negative ratios.

We also emphasize that conducting the signal/background analysis, we implicitly assume that we correctly reconstruct and identify the electron of the final state. We propose here to integrate the CVN response exposed in 2.7. When the simulation files of the DUNE TDR were produced, each neutrino event was stored and two NTuples were associated. One of them was a GENIE record of the simulated interaction, listing the particles of the final state, their 4-momentum, but also the features of the interaction like the Bjorken variables x and y, the transferred momentum Q^2... The other one is related to the reconstruction chain of the event and contains among other the CVN outputs associated to it. The output we are interested in is the classification between ν_eCC, ν_μCC, ν_τCC and NC. Each category is given a score between 0 and 1, and scores sum to 1. We refer to section 2.7 for more details.

In the following, when we explicitly mention the use of the CVN, we in fact require that the CVN score of a given event (signal or background) for the category ν_eCC be greater than 0.85 (the score used to evaluate the DUNE sensitivities [42]) in order to be integrated in the analysis. As far as we are concerned, we only use the CVN as a "pre-classifier" to account for the efficiency of correctly identifying the electron of the final state. Quantitatively, we found that 45.7% of the signal events (ν_τCC with $\tau \to e$), 72.3% of the oscillated ν_eCC and 58.3% of the beam ν_eCC were classified by the CVN as ν_eCC, using the score of 0.85.

We show in figure 3.7 the signal (blue) and background (red) log-likelihood ratio distributions associated to the combined variables $[p_{miss}^{(tr)}\,;\,p_{lep}^{(tr)}]$ and $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}]$. Filled (dashed line) distributions were obtained without (with) the CVN pre-selection, and all distributions are normalized to unity to allow for a direct shape comparison. One can observe that the signal distribution obtained with the CVN is slightly shifted to the left. This indicates that signal events misclassified as ν_e by the CVN are a little harder to identify as ν_τ based on likelihood criteria. There is a small CVN bias, in the sense that the CVN "prefers" signal events that are easier to identify as ν_τ.

We illustrate more precisely this feature in figure 3.8, where on the left we display the signal and background efficiencies as a function of the log-likelihood ratio cut. CVN biased samples are shown in a filmy style, while unbiased samples are shown in a bright color style. The CVN bias effect is more important for the signal efficiency, where it can result in an efficiency drop at

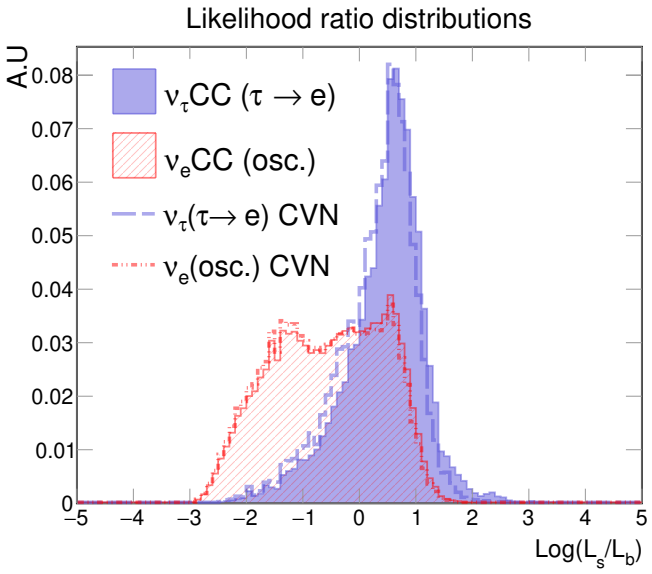
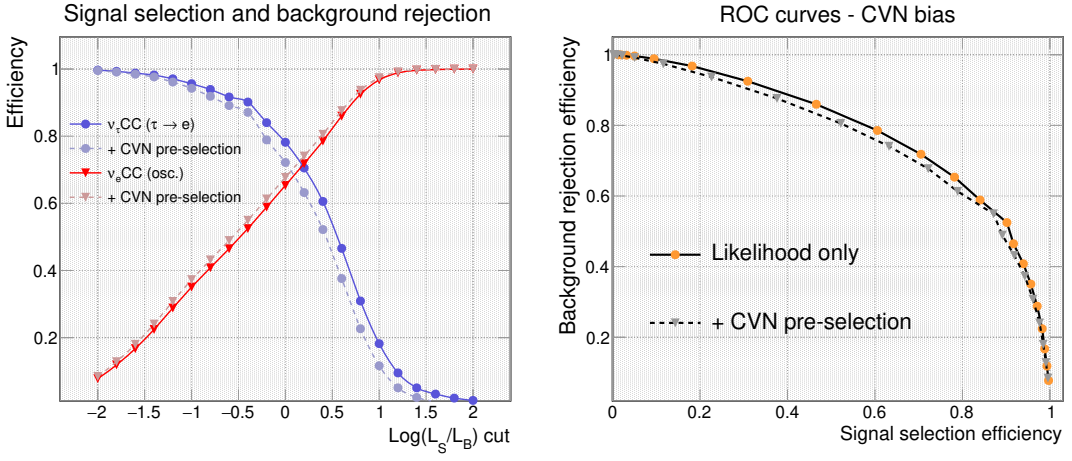


Figure 3.7: Log-likelihood ratio distributions of signal (blue) and background (red), using the 2-dimensional kinematical variables $[p^{(tr)}_{miss}\,;\, p^{(tr)}_{lep}]$ combined with $[\phi^{(tr)}_{hl}\,;\, \phi^{(tr)}_{hm}]$, as defined in equation 3.3. The CVN biased distributions are given with dashed histograms following the same colour convention. This bias is rather small, and mostly impacts the signal distribution which is shifted to the left.



(a) (b)

Figure 3.8: Left: signal (blue) and background (red) efficiencies as a function of the log-likelihood ratio cut with (filmy color) and without (bright color) the CVN pre-selection. Right: ROC curves of these efficiencies with (dashed) and without (full) the CVN pre-selection.

the level of 10% for a given cut value. One can note that this effect on the background events is only at the level of few percents, for which there is conversely a better log-likelihood rejection efficiency.

We emphasize that overall the bias introduced by the CVN pre-selection is small. This is easily understandable by the fact that the CVN relies on visual features of the neutrino events, while the likelihood relies on kinematic features in the transverse plane, and both are not, *a priori*, correlated. In figure 3.8b (right), we also display the ROC curves with and without the CVN bias, to further illustrate the small bias introduced by the CVN.

It is important to note at this stage that the convention most often used in this thesis is that a ROC curve displays the background rejection efficiency against the signal selection efficiency. This convention gives a symmetric role to the signal and the background and was chosen for this reason. However traditionally a ROC curve shows the true positive rate against the false positive rate, which in our case would correspond to the signal efficiency against the background contamination rate. We emphasize that the information carried by the two conventions is the same. The latter will be used in sections making use of machine learning techniques while our convention will be used for the likelihood based analysis.

We observe that the likelihood analysis allows for a 60% selection efficiency of the signal and a 20% background contamination, and that the use of the CVN does not affect much this likelihood performance. However it introduces an additional efficiency which disfavours the signal (46%) in comparison to the two backgrounds (respectively 72% and 58% for the oscillated and beam ν_e). The 95% background rejection is achieved for a corresponding 20% of the signal efficiency. As for the oscillated ν_e, we observe that the CVN pre-selection does not bias the likelihood efficiency in terms of beam ν_e rejection.

We now propose to assess the behavior of the second background component, the beam ν_e, in regards of the likelihood analysis. Before doing so, we will have a small digression to show the likelihood analysis consistency.

3.3.3 *A posteriori* check

It is always a good idea to check that an analysis performs the way we intend it to. We propose in this section to make an *a posteriori* check on the likelihood analysis. We presented in figure 3.4 that the $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}]$ distribution of the signal populated regions of $[0\,;\,180]^\circ$, while the background (oscillated ν_e) did not. Since this 2-dimensional distribution is used in the optimized likelihood search we performed earlier, we should observe that indeed ν_τ events populating this specific region have a high log-likelihood ratio. That is, the likelihood analysis tends to select these events because they constitute an easily recognizable signal.

We show in figure 3.9 the distribution of $[\phi_{hl}^{(tr)}\,;\,\phi_{hm}^{(tr)}]$ for ν_τ events for which the log-likelihood ratio exceeds 1.2. These events are indeed associated to the highly signal like region $[0\,;\,180]^\circ$ of the plane.

3.3.4 Oscillated and beam ν_e backgrounds

In this section we propose to add the beam ν_e background to the analysis. We have so far used the oscillated ν_e to optimize the likelihood search. We give in figure 3.10 an extension of the figure 3.7 to the beam ν_e (without the CVN bias). All distributions are normalized to unity.

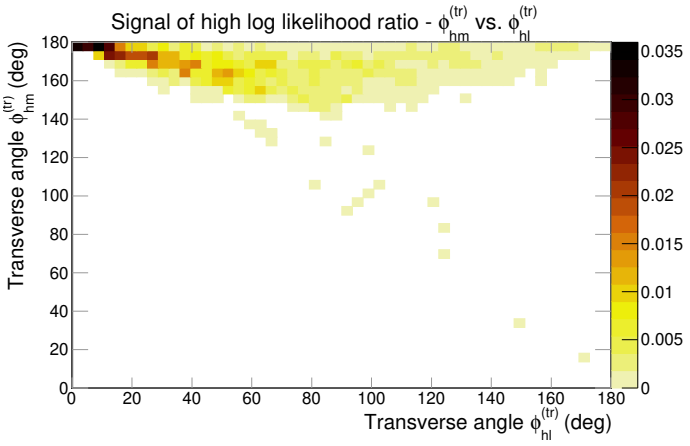
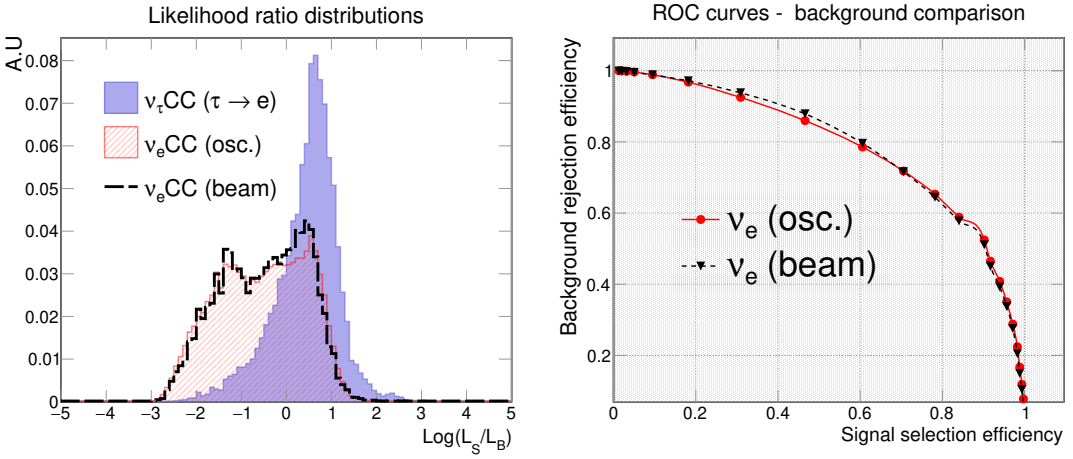
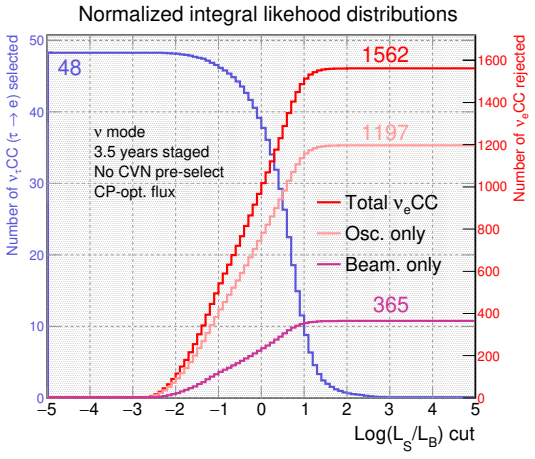


Figure 3.9: Signal $[\phi^{(tr)}_{hl}\,;\,\phi^{(tr)}_{hm}]$ distributions associated to a log-likelihood ratio greater than 1.2 (using the likelihood defined in 3.3.1.)



(a) (b)

Figure 3.10: Left: log-likelihood ratio distributions of ν_τ (blue filled), oscillated (red filled) and beam (black dashed line) ν_e events. We did not use the CVN pre-selection. We observe a close behavior between the two background components in regards of the likelihood analysis. Right: same but shown as ROC curves, *i.e* background rejection efficiency as a function of signal efficiency for likelihood cuts varying in $[-2\,;\,2]$.



(a)

Cut	ν_τ	ν_e (beam+osc)	S/B
0.6	22.5 ± 0.2	121.1 ± 0.7	0.19 ± 0.00
0.8	14.9 ± 0.2	112.4 ± 0.5	0.13 ± 0.00
1.0	8.8 ± 0.1	48.9 ± 0.3	0.18 ± 0.00
1.2	4.6 ± 0.1	17.9 ± 0.2	0.26 ± 0.01
1.4	2.4 ± 0.1	6.1 ± 0.1	0.39 ± 0.02
1.6	1.5 ± 0.1	2.7 ± 0.1	0.56 ± 0.03
1.8	1.0 ± 0.1	1.2 ± 0.0	0.83 ± 0.07

(b)

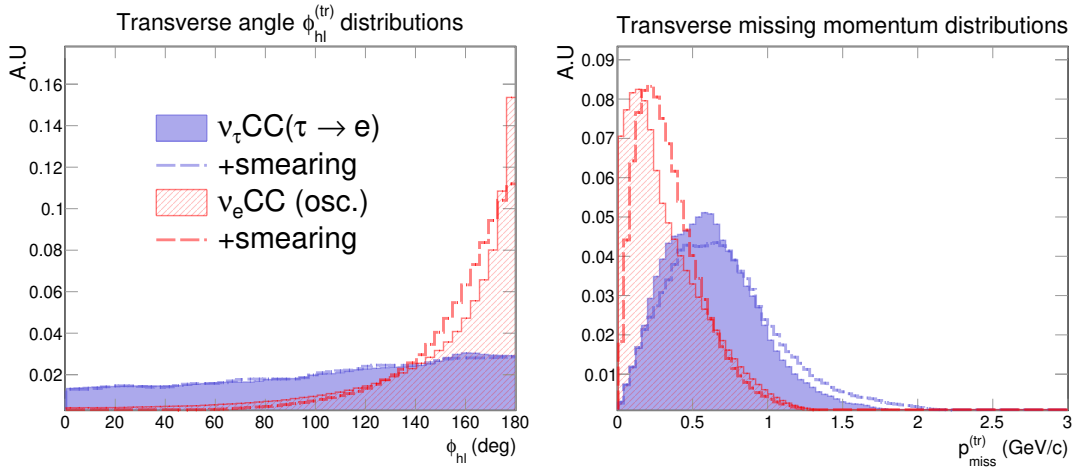
Figure 3.11: Left: integral likelihood distributions normalized to the 3.5 years staged deployment plan of DUNE, assuming normal ordering of mass hierarchy and oscillation parameters of table 2.1. Number of events are taken from table 2.4. Right: corresponding number of events selected for several cut values of the log-likelihood ratio. Errors are statistical only, they were calculated from the statistics available in the simulation files.

We do not observe a significant behavior difference between the two background components with respect to the likelihood analysis we earlier exposed. We showed in figure 3.6 that the two backgrounds behaved differently only in terms of the kinetic energy of the electron and the transverse momenta of the electron and the hadronic systems. We did not observe major differences for the transverse angles. Since we optimized the likelihood without the kinetic energy of the electron and that the likelihood is not sensitive to the two latter transverse momenta (*i.e* same distributions for ν_τ and oscillated ν_e), this result sounds reasonable. Let us now normalize the previous results in relation to the expected number of events at the DUNE far detector.

3.3.5 Normalization to 3.5 years staged

In this section we make use of the calculated event rates presented in 2.2. Assuming the constant density model of earth, we expect 270 ν_τ events, 1197 oscillated ν_e events and 365 beam ν_e events. We assume the normal ordering hypothesis and the neutrino (FHC) beam configuration. The event rates correspond to the so-called 3.5 years staged deployment plan, in which the four 10 kton modules of the far detector site are deployed step by step (see discussion around equation 2.14). We also need to take into account the $\tau \to e$ branching ratio of 17.83%. This implies that the effective number of signal events is about 48. The combined ν_e backgrounds amount to 1562 events (initial S/B ratio of about 3%).

We give in figure 3.11a the efficiencies, without the CVN pre-selection, of the signal and backgrounds, normalized to the expected number of events, as a function of the likelihood cut. The signal (blue) axis is on the left, while the background (red) axis is on the right of the plot. The three background histograms must be read as rejection efficiencies, that is to say the value read for a given likelihood cut value corresponds to the number of events rejected (or selected as ν_e). A cut value of 1.0 must be used to ensure at least 95% of background rejection. However, as shown in table 3.11b, even likelihood cut values greater than 1 do not allow for having a signal over background ratio greater than 1. Uncertainties on the results are statistical only, they are



(a) (b)

Figure 3.12: Smeared (dashed lines) and not smeared (full) distributions of the transverse angle $\phi_{hl}^{(tr)}$ (left) and the transverse missing momentum (right). Blue is the signal (ν_τ) and red the background (oscillated ν_e).

function of the statistics available in the simulation files. They are shown to illustrate that we have enough statistics even in the low background signal regions.

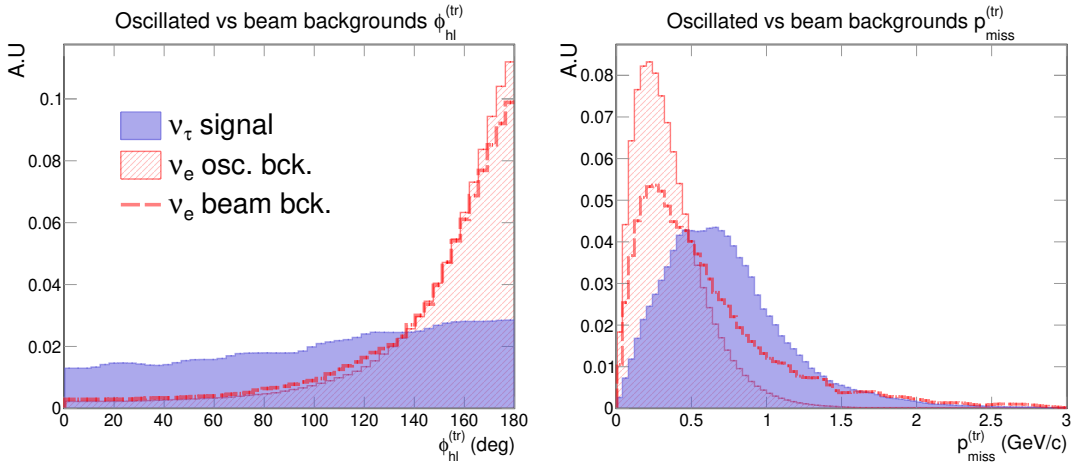
Given that the CVN efficiency selection was less good for the signal (45.7%) than for the background (72.3% for the oscillated ν_e and 58.3% for beam ν_e), the results exposed above are less good applying these additional efficiencies. The exposed results are now revisited in the light of the smearing effects.

3.4 Adding smearing effects

3.4.1 Kinematical distributions

The smearing designates a process to include experimental measurement uncertainties at the level of single particles in the generated events. Until now in this chapter, the true Monte Carlo information of particles was used to produce the results. This procedure is useful in a first instance to evaluate the best capability of the experiment, but it does not reflect its actual potential. The smearing process we use is described in 2.5, and the summary of reconstruction effects can be found in table 2.4. Figure 3.12 shows the effect of the smearing on two kinematical distributions: the transverse missing momentum and the transverse angle $\phi_{hl}^{(tr)}$ in figure 3.12. The previous distributions (not smeared) are given for comparison in filled histograms.

One can observe that the peak at 180° of the background transverse angle distribution was softened by including the smearing process while the corresponding signal distribution remains rather flat. This implies the variable to be slightly less selective. The signal and background transverse missing momentum distributions were shifted to higher values. The signal now has a longer tail above 1 GeV/c. Indeed, without (with) smearing there are 14% (21%) of the signal events above 1 GeV/c. This is an effect of the hadronic misreconstruction in the transverse plane, mostly. This tailing effect is not observed on the background distribution.



(a) (b)

Figure 3.13: Comparison of the two ν_e background components in terms of kinematical distributions. The ν_τ signal is still in blue, the oscillated ν_e background in red filled and the beam ν_e background in red dashed lines. Left: transverse angle between the hadronic and leptonic systems. Right: modulus of the transverse missing momentum.

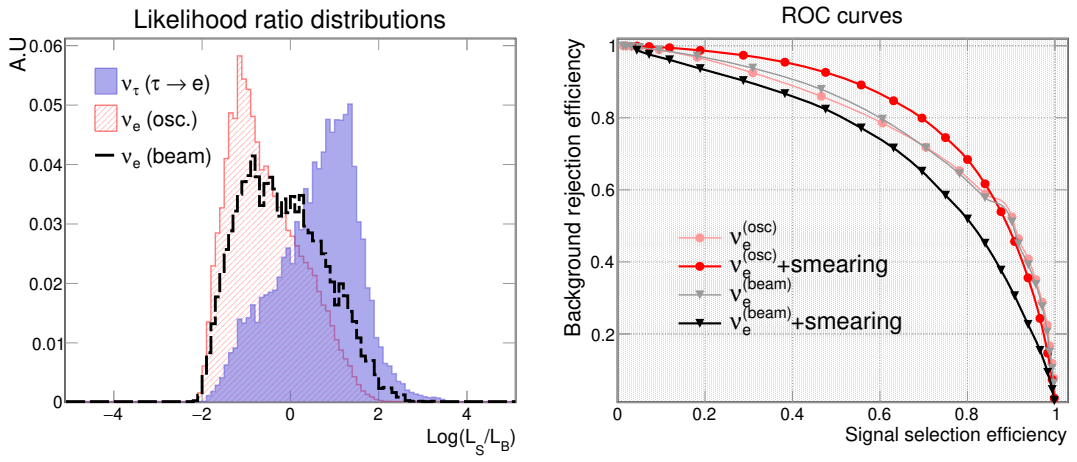
We now show the effects induced by the smearing by comparing the two components of the background, the oscillated and beam ν_e, with the transverse angle $\phi_{hl}^{(tr)}$ and the transverse missing momentum modulus distributions. Apart from the kinetic energy of the electron not used in the analysis, the differences observed between the two background types at the level of the kinematical distributions are the same as the ones without the smearing effects, that we already discussed in 3.2.3. There is no much difference between the $\phi_{hl}^{(tr)}$ distributions (see figure 3.13a). The only significant change is the deterioration of the transverse missing momentum distributions of the beam ν_e which can be seen in figure 3.13b. Remember we had observed similar distributions between the two types of background in figure 3.6b. This can be understood by the fact that since the beam ν_e have larger transverse hadronic and leptonic momenta than the oscillated ν_e, the smearing brought more missing momentum in the transverse plane of the beam ν_e.

3.4.2 Likelihood distributions and ROC curves

In this section the CVN pre-selection in not used in the analysis. We will make use of the same combination of 2-dimensional distributions than in 3.3.1 to build the likelihood analysis. We show the log-likelihood ratio distributions of the signal and the two backgrounds, as well as the corresponding ROC curves in figure 3.14. On the left the ν_τ signal is shown in blue, the oscillated ν_e in a red filled style and the beam ν_e in a black dashed line style. On the right the ROC curves obtained with (without) the smearing effects are shown in bright (filmy) colors for the two types of background separately.

There are two features to discuss at this point.

1. A better signal to oscillated ν_e background separation is obtained when taking into account the detector effects via the smearing. This can be seen on the ROC curves (filmy red and bright red). This is quite surprising because usually when we deploy the smearing, we take into account the detector effects and thus deteriorate the results obtained without



(a) (b)

Figure 3.14: Left: log-likelihood ratio distributions built using the combination of 2-dimensional variables described in 3.3.1. The ν_τ signal is shown in blue, the oscillated ν_e in red and the beam ν_e in a black dashed line histogram. Right: corresponding ROC (background rejection efficiency against signal selection efficiency), where the oscillated (beam) ν_e contribution is given in red (black). The filmy curves show the ROC curves obtained without the smearing effects as for comparison.

smearing which normally indicate the best sensitivity. We emphasize that for the "no smearing" framework we refer to, we actually fully discarded the neutrons from the final states (this is a kind of smearing already). This had a quite significant impact on the kinematical distributions, as we showed in figure 3.5. The smearing we use assumes a bad reconstruction of the neutrons energy, and they are given a 10% chance to fully escape detection. As a consequence, we met a specific case in which the "no smearing no neutrons" framework performed less good than with the smearing method. We checked that if we take into account the neutrons and do not apply the smearing, we find even better results than the ones obtained with the smearing.

2. Secondly, the separation between ν_τ and beam ν_e was deteriorated. This effect is mostly to be attributed to the transverse missing momentum of the beam ν_e (see figure 3.13b). There is large tail above $1\,\mathrm{GeV}/c$ that the oscillated ν_e do not have. As a consequence, since we build the likelihood distribution of the beam ν_e using the ν_τ and oscillated ν_e kinematical distributions, a significant fraction of the beam ν_e will have a signal-like score in terms of transverse missing momentum, deteriorating this background rejection.

3.4.3 Normalization to 3.5 years staged

We now normalize the previous efficiencies in relation to the expected number of events at the DUNE far detectors site. Again, we refer to the 3.5 years staged deployment plan explained in 2.4.2, which is equivalent to 10 years of exposition of one 10 kt far detector module. Because of the smearing, in the simulation files we can find neutrino events in which the electron goes undetected. In this case the event is discarded from the analysis, and we take it into account in the normalization. There can also be events without reconstructed hadronic system (only the electron of the leptonic system is reconstructed). Such events are discarded as well. The effect of both criteria is at the level of 1% for any of the samples considered here. The final

normalizing factors are 47 ν_τ, 1178 oscillated ν_e and 361 beam ν_e events which gives an initial S/B ratio of 47/1539 (3%).

The efficiencies presented in the previous section normalized to these number of events are given in table 3.1. The two types of background and their sum are shown. The errors are statistical only. This table reveals that it is difficult to find a low background region, for which the S/B would be at least greater than 1. As for illustration, a log-likelihood ratio cut at 2.4 removes more than 99.9% of the oscillated ν_e and 99.1% of the beam ν_e, leading to a total background of 3.2 ± 0.3. However, in this region, the signal efficiency is about 2% only, which is very low.

$\log(L_S/L_B)$ cut	0.6	0.8	1.0	1.2	1.4
ν_τ	26.2 ± 0.2	22.4 ± 0.2	18.0 ± 0.2	13.5 ± 0.2	8.9 ± 0.1
ν_e osc.	128.2 ± 0.5	87.0 ± 0.5	56.7 ± 0.4	30.9 ± 0.3	15.0 ± 0.2
ν_e beam	82.5 ± 1.7	63.6 ± 1.5	48.1 ± 1.3	34.9 ± 1.5	22.9 ± 0.9
ν_e total	210.6 ± 1.7	150.6 ± 1.5	104.7 ± 1.3	65.9 ± 1.5	38.0 ± 0.9
$\log(L_S/L_B)$ cut	1.8	2.0	2.2	2.4	2.6
ν_τ	5.6 ± 0.1	3.4 ± 0.1	2.1 ± 0.1	1.3 ± 0.1	0.8 ± 0.1
ν_e osc.	6.5 ± 0.2	2.7 ± 0.2	1.0 ± 0.0	0.4 ± 0.0	0.1 ± 0.0
ν_e beam	14.1 ± 0.7	8.8 ± 0.5	4.6 ± 0.4	2.8 ± 0.3	1.3 ± 0.2
ν_e total	20.6 ± 0.7	11.4 ± 0.6	5.6 ± 0.4	3.2 ± 0.3	1.5 ± 0.2

Table 3.1: Expected number of events selected depending on the log-likelihood ratio cut used. The two backgrounds types are shown, as well as their sum. Errors are statistical only, and are function of the available statistics in the simulation files.

3.4.4 QEL-like events

The results of the previous section highlight the difficulty to isolate a rather clean sample of τ neutrino interactions using the electron decay mode of the charged τ lepton, for which the dominant background are the true simulated ν_eCC events. Also remember that these results are obtained assuming to use the standard neutrino flux (optimized for CP violation study) running in ν mode. The high oscillation probability $\nu_\mu \to \nu_\tau$ (see figure 2.8) is largely compensated by the unfavourable ν_τCC cross section at energies ranging in $[0\,;\,6]$GeV, where most of the neutrino flux is contained.

Though the likelihood performance is satisfactory (it allowed for instance for a 90% of background rejection and a 40% signal efficiency), it is not enough to compensate the large amount of background. Indeed, for the 3.5 years staged hypothesis framework we expect a total signal of 47 signal *vs.* 1539 total background events, that is to say an initial signal to background ratio of 0.03. This means we could achieve a S/B ratio of about unity if we reached a signal selection efficiency of 50% and a background rejection of 99%. Clearly the likelihood did not reach such performance (see figure 3.14b).

We can possibly improve the performance of the analysis by reducing the analysis to exclusive final state events, such as the quasi-elastic (QEL). It has the advantage of having a low multiplicity hadronic system (only one proton) so the reconstruction effects impact should be less important. It should be noted that the QEL final state is required regardless of the truth at the generator level, and because of nuclear effects (final state interaction, pion absorption in the nucleus...), not all true QEL events will satisfy this criteria, and some non-QEL will. We should reserve the term QEL for the MC truth considerations, and we shall rather employ the term "QEL-like" to refer to exclusive final states "$1p1e^-$".

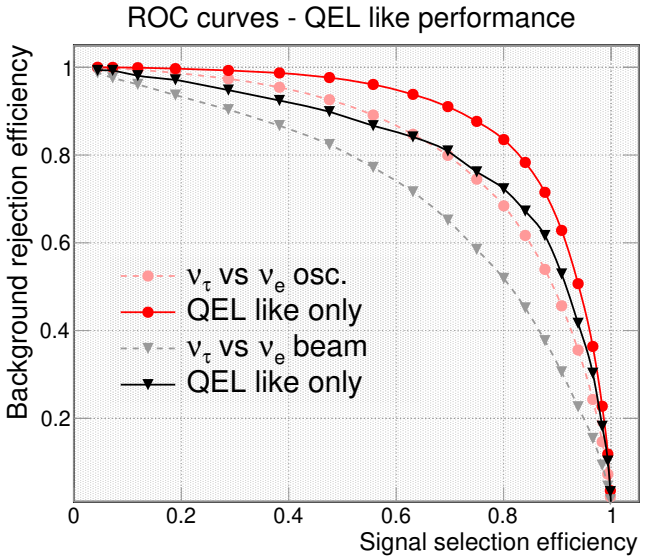


Figure 3.15: Comparison of the ROC curves obtained with (bright colors) and without (filmy colors) the QEL-like criteria. Curves with circles (triangles) correspond to the oscillated (beam) ν_e background.

We mentioned in 2.2.2 that for DUNE it is expected that about half (46%) of the ν_τCC interactions be quasi-elastic with the standard LBNF neutrino beam, optimized for the CP violation search. This fraction only reaches 22% for the oscillated ν_e and 13% for the beam ν_e. Thus there is a possibility of significant background reduction. Therefore we perform the same likelihood analysis as previously but we require in addition the final state of each neutrino event to be QEL-like, *i.e* to be composed only of one proton and one electron. This condition is imposed taking into account the detector effects (smearing).

We observe that:

- Among all the ν_τ events available for the analysis (removing events without hadronic system or for which the electron was not reconstructed), 27% of them satisfy the QEL-like criteria (while 46% are actual QEL). About 3% of this QEL-like population are not true QEL interactions.
- 13% of the oscillated ν_e are QEL-like (against 22% true QEL), and 6% of this QEL-like population are not true QEL interactions.
- 8% of the beam ν_e are QEL-like (against 13% true QEL), and 6% of this QEL-like population are not true QEL interactions. Beam ν_e events stand at higher energy than oscillated ν_e, so it is expected that the QEL-like fraction is smaller.

It is worth to start by evaluating the performance of the likelihood on these subsets of neutrino events, and compare them with results shown in figure 3.14b. The final states of these events are rather clean, so it is expected that the transverse plane kinematics are more discriminating. Indeed, we give the comparison of the background *vs* signal efficiencies with and without the QEL-like criteria in figure 3.15. Filmy (bright) curves are obtained without (with) the QEL-like criteria. Curves with triangles (circles) correspond to beam (oscillated) ν_e. There is a clear increase of the likelihood performance when we focus only on the QEL-like samples.

$\log(L_S/L_B)$ cut	0.6	0.8	1.0	1.2
ν_τ	6.6 ± 0.2	5.4 ± 0.2	4.0 ± 0.1	2.6 ± 0.1
ν_e osc.	6.2 ± 0.1	3.7 ± 0.1	2.0 ± 0.1	1.1 ± 0.1
ν_e beam	3.9 ± 0.5	2.9 ± 0.4	2.2 ± 0.4	1.5 ± 0.3
ν_e total	10.0 ± 0.5	6.6 ± 0.4	4.2 ± 0.4	2.6 ± 0.3
$\log(L_S/L_B)$ cut	1.4	1.6	1.8	
ν_τ	1.3 ± 0.1	0.6 ± 0.1	0.1 ± 0.0	
ν_e osc.	0.5 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	
ν_e beam	0.8 ± 0.2	0.6 ± 0.2	0.2 ± 0.1	
ν_e total	1.3 ± 0.2	0.7 ± 0.2	0.3 ± 0.1	

Table 3.2: Expected number of QEL-like events selected for various log-likelihood cut values, assuming the 3.5 years staged deployment plan of the DUNE experiment. Errors are statistical only.

We quote the corresponding expected number of selected events for various log-likelihood ratio cut values for the signal and the two types of backgrounds in table 3.2. In comparison to table 3.1, a much better signal to background ratio is achieved. As for illustration, a ratio of 1 is achieved with a log-likelihood cut value of 1.0 whereas without the QEL-like criteria the S/B ratio of 1 was not reached at all.

Another point of view we can adopt is the one of the significance. The significance measures the sensitivity of an experiment in relation to a given phenomenon. Typically in counting experiments, the expected total number of events we detect n is Poisson distributed. It is the sum of a signal s and a background b. Often we compute the ratio $s/\sqrt{b}$ and interpret the number as a significance. For instance, a significance of 5 means that given a predicted background b, the observation of n events can not be explained by a statistical fluctuation of the background with a confidence level of 5σ. Assuming that the observation can be attributed to the background only hypothesis would occur only in one among about 2 million repeated identical experiments. This is the standard accepted significance to claim for a discovery (*i.e* reject the background only hypothesis) in high energy physics.

Here the purpose is not to claim for discovery, but to assess *a priori* the sensitivity of DUNE to the ν_τ appearance phenomenon. We emphasize that the formula $s/\sqrt{b}$ is valid for large statistics (*i.e* large $n = s+b$) only, when the Poisson distribution of n can be approximated to a Gaussian distribution. Given the expected number of events we gave in tables 3.1 and 3.2, we should rather use the Asimov significance, for which the Poisson distributions are not approximated by Gaussian distributions:

$$Z_A = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)}. \tag{3.4}$$

It is possible to prove that for $s \ll b$, $Z_A \to s/\sqrt{b}$. For low $s+b$, however, the formula $s/\sqrt{b}$ tends to overestimate the real significance.

We observe in figure 3.16 that reducing our analysis to QEL-like events only does not improve the significance. Only for log-likelihood cut values ranging in $[-0.5\,;\,1.0]$, we barely improved the significance.

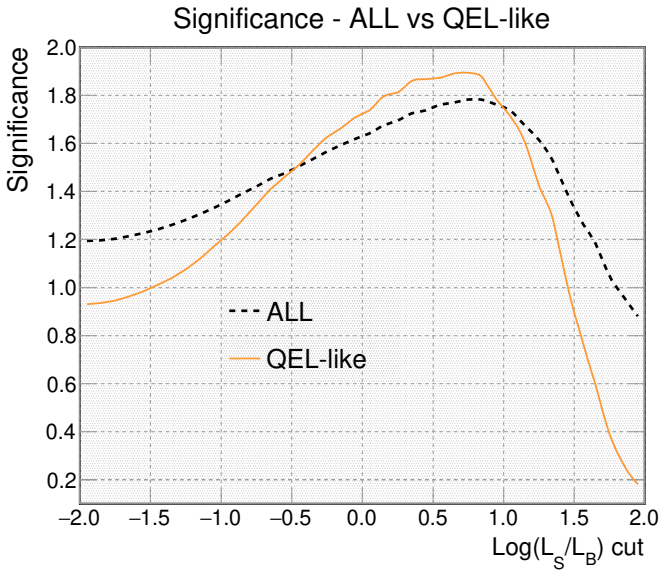


Figure 3.16: Asimov significance as a function of the log-likelihood ratio cut with (without) the QEL-like criteria, corresponding to the orange full (dashed black) curve.

We conclude that the QEL-like criteria helps at improving the signal to background ratio, though it implies to reduce the signal by a factor of about 5. However it does not increase the DUNE sensitivity to the $\nu_\mu \to \nu_\tau$ appearance in beam events.

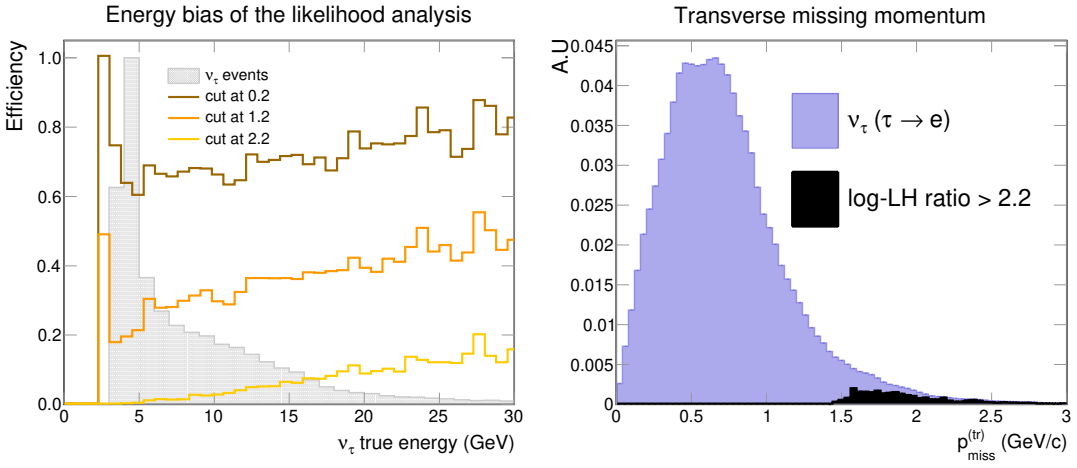
3.4.5 Energy bias of the analysis

It is worth to display the behavior of the likelihood analysis with respect to the impinging neutrino true energy. It gives a global picture of the way the likelihood selection of the signal biases the neutrino events.

In figure 3.17a we show the signal (ν_τ ($\tau \to e$)) efficiency as a function of the τ neutrino true energy. We also show the ν_τCC event rate (arbitrarily normalized) as a gray filled histogram. The efficiency peak at 3.5 GeV does not reflect any physical behavior, it is an energy bin with a very low statistics.

We observe that for a relatively low cut value (0.2, dark orange histogram), the selection efficiency is almost flat and is about 70%. Moving to higher cut values (1.2 and 2.2), we observe that the likelihood prefers higher energy events. On the yellow histogram (log-likelihood ratio cut value of 2.2), the signal efficiency is up to few percents only in the energy range smaller than 10 GeV, and is almost null in the energy region where the event rate is maximal ($\simeq$ 5 GeV). The signal efficiency increases up to 15% at energies greater than 20 GeV, where the event rate is very low.

We understand this high energy bias with the transverse missing momentum variable. We show in figure 3.17b that the signal events with the highest log-likelihood ratios (>2.2) correspond to events populating the tail (starting from 2.5 GeV/c) of the transverse missing momentum distribution. In this tail region of the distribution, there is however almost no oscillated ν_e events, which makes it very signal-like.



(a) (b)

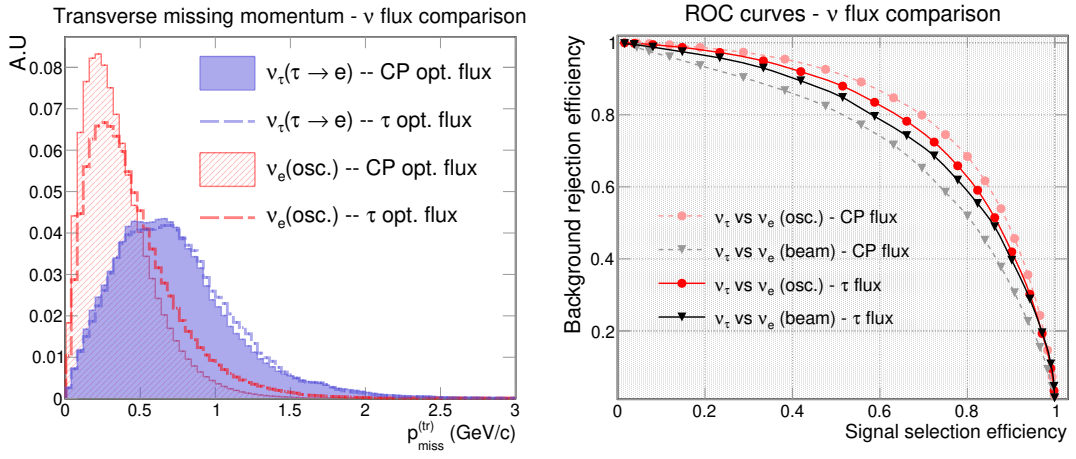
Figure 3.17: Left: Signal (ν_τ ($\tau \to e$)) selection efficiency as a function of the τ neutrino true energy. The gray filled histogram represents the ν_τ event rate arbitrarily normalized. Right: transverse missing momentum of all $\nu_\tau(\tau \to e)$ (blue) and the highest likelihood signal events (black). The black histogram is correctly normalized with respect to the blue one.

3.5 τ optimized beam flux

3.5.1 Impact on the kinematical distributions

We mentioned in section 2.2.2 that an alternative beam neutrino could be considered for the DUNE experiment. It would require a modification of the horns and their positioning, and would allow for selecting higher energy neutrinos in the beam. As far as the τ flavour is concerned, it would at some extent compensate for the kinematic suppression factor of the ν_τ cross section at energies of few GeV. We indeed showed in 2.2.2 that the use of this alternative neutrino beam would multiply the ν_τ statistics by a factor of 6. It is all the more interesting to note that the electron neutrino event rate is little modified (see table 2.3) by the change of beam configuration (1197 $\to$ 1199 for oscillated ν_e and 365 $\to$ 543 for beam ν_e). Indeed, we expect the same event rate for the oscillated ν_e, while the beam ν_e would be multiplied by a factor of 1.5 only. This alternative neutrino beam brings a considerable initial (*i.e* without likelihood analysis) signal to background ratio increase, and offers a promising scope to improve the beam ν_τ appearance using the $\tau \to e$ decay mode.

A priori, it is not clear to figure out how the likelihood performance itself is affected by the beam configuration change. In order to do so we must have a look at the signal and background kinematical distributions. We show in figure 3.18a the transverse missing momentum distributions for signal (blue) and oscillated ν_e background (red) for the two beam configurations. The filled (non filled) histogram correspond to the standard CP optimized (τ optimized) neutrino flux. The signal distribution is little affected, while the oscillated ν_e one is slightly shifted to higher values. Similar observations are found for the beam ν_e. We did not observe significant change for transverse angles distributions.



(a) (b)

Figure 3.18: Left: transverse missing momentum distributions for signal (blue) and oscillated ν_e (red) for the CP optimized (τ optimized) beam configuration in filled (dashed line) histograms. Right: background rejection efficiency versus signal selection efficiency (ROC curve) for the two types of backgrounds and with the two beam configurations. Oscillated (beam) ν_e correspond to red (black) curves, and CP optimized (τ optimized) to dashed (full) curves.

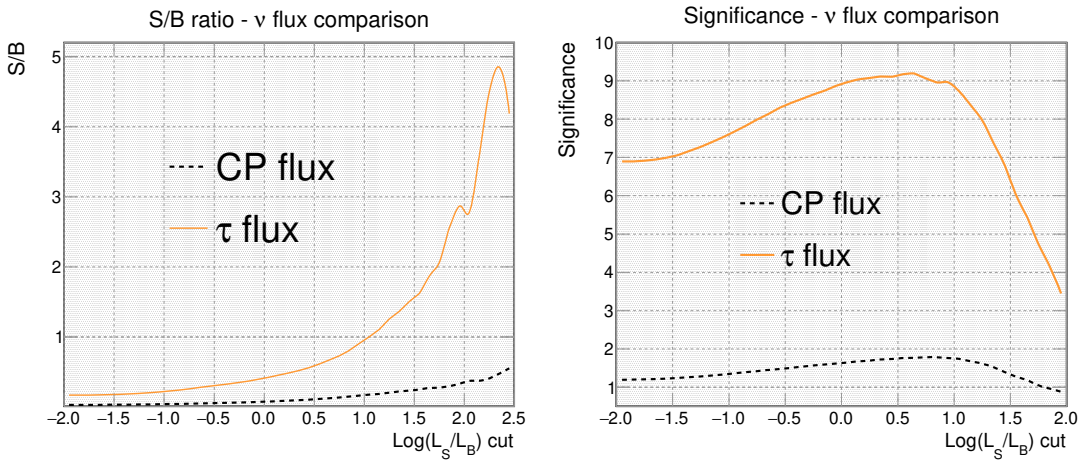
3.5.2 Impact on the likelihood analysis

The likelihood analysis is repeated with the alternative beam configuration. We calculate the log-likelihood ratios of the signal and the two background types under the test hypothesis "signal *vs* oscillated ν_e". The beam ν_e likelihood distributions are calculated with the oscillated ν_e distributions, so once more the likelihood is more of a pseudo-likelihood. We use the same combination of 2-dimensional variables that we defined in 3.3.1. Indeed, we did not find other combinations of variables which allowed for a better signal/background discrimination. However there are others combinations which lead to similar results (as already discussed in 3.3.1). Figure 3.18b shows the likelihood performance for the two types of background. We observe that it performs slightly less better for the oscillated ν_e and a bit better for the beam ν_e.

$\log(L_S/L_B)$ cut	0.6	0.8	1.0	1.2
ν_τ	151.6 ± 1.2	123.4 ± 1.1	98.2 ± 1.0	68.8 ± 0.9
ν_e osc.	143.6 ± 0.5	95.5 ± 0.4	60.0 ± 0.3	32.0 ± 0.2
ν_e beam	82.3 ± 2.0	57.2 ± 1.7	38.1 ± 1.4	22.7 ± 1.0
ν_e total	225.9 ± 2.1	152.8 ± 1.7	98.1 ± 1.4	54.8 ± 1.1
$\log(L_S/L_B)$ cut	1.4	1.6	1.8	
ν_τ	43.6 ± 0.8	23.8 ± 0.7	11.3 ± 0.5	
ν_e osc.	15.4 ± 0.2	6.1 ± 0.1	1.7 ± 0.1	
ν_e beam	13.6 ± 0.8	6.6 ± 0.6	2.7 ± 0.4	
ν_e total	29.0 ± 0.8	12.7 ± 0.6	4.4 ± 0.4	

Table 3.3: Expected number of selected events for various log-likelihood cut values, assuming the 3.5 years staged deployment plan of the DUNE far detectors and the use of the alternative τ optimized neutrino flux. Errors are statistical only.

Table 3.3 summarizes the evolution of the values presented in table 3.1 under the beam configuration change, assuming a 3.5 years staged deployment plan. The expected number of signal events is drastically increased, and the S/B ratio of unity is reached for a cut value of about



(a) (b)

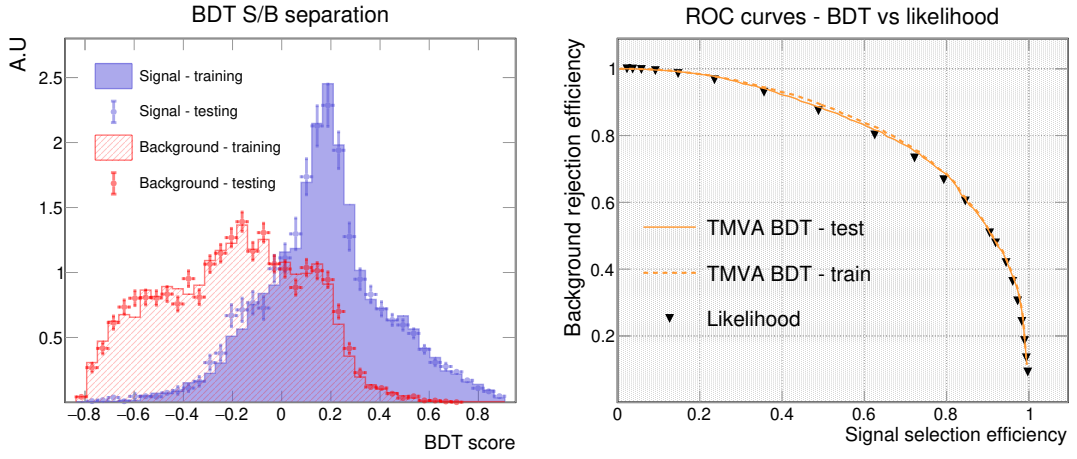
Figure 3.19: Left: S/B ratio as a function of the log-likelihood ratio cut for the CP optimized (τ optimized) beam configuration in dashed black (full orange). Right: Asimov significance as a function of the log-likelihood ratio cut for the standard (alternative) beam configuration in dashed black (full orange). Both plots were obtained under the assumption of the 3.5 years stage deployment plan hypothesis.

1.0. We display this tremendous change in figure 3.19a, where we represent the 3.5 years staged expected S/B ratio as a function of the log-likelihood cut for the two beam configurations. We recover the fact that we barely reached a S/B ratio of unity with the CP optimized beam for any cut value used. However, we achieve a much more purified sample with the alternative τ optimized beam, where we reach a S/B ratio of about 3 for a cut value at 2.0 (it corresponds to 5 signal events selected). We also display the Asimov significance (see equation 3.4) as a function of the log-likelihood ratio cut value used. We reach a maximum of 9 with the alternative beam configuration, where we barely reached 2 with the standard LBNF beam.

3.6 Machine learning techniques

It is worth to compare our analysis based on a likelihood approach to the performance of other tools. Machine learning techniques have proved to provide a huge contribution in high energy physics during the last years, both at reconstruction and signal identification levels. This section proposes to compare the likelihood performance exposed in 3.3 to the ones of machine learning tools.

For sake of simplicity, we will limit ourselves to signal-background separation using only one type of background, the oscillated ν_e. We also do not take into account the reconstruction effects (smearing), so the inputs to the machine learning tools are the distributions such as the ones shown in figure 3.2. We make use of the seven variables defined in 3.1.2, except the asymmetry ratio. We integrate the kinetic energy of the electron in the analysis, which we did not use in the previously exposed likelihood results. We updated the performance of the optimized likelihood search presented including this new variable to allow for a proper comparison with the machine learning tools. This section explores the S/B separation with Boosted Decision Trees (BDT) and Artificial Neural Networks (NN).



(a) (b)

Figure 3.20: BDT response to the signal/background classification. We show on (a) the BDT score distributions of the signal (blue) and background (red), for both the training (full) and testing (marks) sets. The corresponding ROC curves of background rejection efficiency versus signal selection efficiency (orange) and its comparison with the likelihood performance (black triangles) is shown on the right (b).

3.6.1 Boosted Decision Trees (BDT) TMVA from ROOT

ROOT provides a machine learning environment called TMVA (Toolkit for MultiVariate Data Analysis), composed of various tools. We evaluate the performance of a Boosted Decision Tree (BDT). A tree classifier uses cuts in the kinematical distributions provided and finds cut values to allow for a maximal signal/background separation. When boosted, many trees (400 in our case) act together one after an other, the tree $i+1$ being typically trained on the set of events where the tree i failed to have a correct classification. Trees get more and more specific and are trained on harder tasks. This method to use many trees, each individually "not so good", allows to have an overall robust decision algorithm.

Our simulation sample contains $\mathscr{O}(30\,000)$ ν_τCC (signal) events and $\mathscr{O}(700\,000)$ oscillated ν_e (background) events. The first step is to randomly select about 30 000 events of the background set as to equalize the two sets. 3/4 of both sets are used for the training of the BDT, and the last 1/4 is used for testing the tree with an independent set of events. The BDT is composed of 400 trees.

We give in figure 3.20a the BDT response on the training and testing samples. Blue (red) histograms designate signal (background). Full histograms are the training sets, and dots with statistical error bars are the testing sets. The BDT assigns a score of 1 (-1) to events that are, without doubt, signals (backgrounds). We illustrate the corresponding efficiencies on the signal and background samples in figure 3.20b. The ROC curve displays the background rejection efficiency versus the signal efficiency. Orange curves are the BDT response, and the triangle markers show the best likelihood performance we obtained in parallel. We find that the BDT performs similarly than the likelihood method.

3.6.2 Neural Network (NN) analysis using the TensorFlow platform

The artificial Neural Networks (NN) are another class of quite popular machine learning tools. Unlike BDT which merely apply cuts on kinematical distributions, the NN performs many non-linear operations on data (activation functions). The network is structured in layers of individual neurons, and each neuron of layer i is connected to neurons of layers $i-1$ and $i+1$ via activation functions, each with its own parameters. The first layer of the network is composed of the data (here there are 7 layers in input), and the output layer depends on the task assigned to the NN. We are interested in signal/background separation, so the network's final layer is composed of one neuron, a signal score which varies between 0 (targeted background events) and 1 (targeted signal events).

When an event is given in input to the NN (in the form of 7 variables), all the calculations from one neuron to another using the activation functions are performed until the last neuron which contains the signal score. The training is supervised, indicating each time whether the event was signal or background. A loss function allows to penalize the NN in case it takes a wrong decision (for instance, attributing a low score to a true signal event). For a binary decision, we typically use the cross-entropy loss function:

$$F_{loss}(data) = -\frac{1}{N}\sum_{i=1}^{N}\left(t_i \log(p_i) + (1-t_i)\log(1-p_i)\right), \tag{3.5}$$

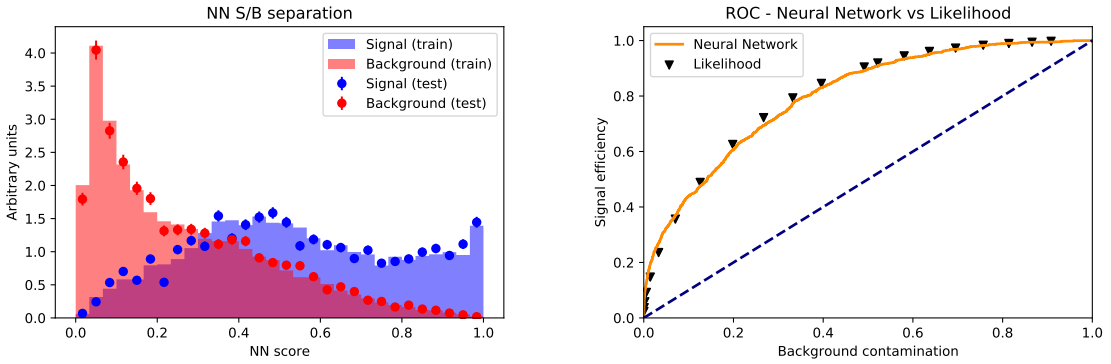
noting N the total sample size (signal + background), p_i the NN score attributed to the i^{th} event, and $t_i = 1$ for signal, 0 otherwise. For signal events, only the first term is non-null, and the loss function increases as p_i tends to 0, which corresponds to signal events misclassified as background.

TensorFlow is an open source platform dedicated to machine learning, based on the Python programming language. The library **Keras** provides a framework to create and train deep neural networks. We used a NN with one hidden layer composed of 1024 neurons. Modifications of the structure of the network (adding hidden layers with variable sizes) was not found to be of substantial help. An epoch designates one run of the neural network over the total dataset. Between two epochs, the NN modifies its internal parameters in order to decrease the loss function. Thus, the neural network acts like a huge minimization problem, trying to find a set of internal parameters which allow for a minimal loss function.

We show in figure 3.21a the output NN score (varying in $[0\,;\,1]$) for the signal (blue) and background (red) samples. The filled (dot) histograms show the training (testing) samples which accounts for 3/4 (1/4) of the signal and background sets. No clue of over training was found. As for comparison with the likelihood, the ROC curve of the signal efficiency against the background contamination is shown in figure 3.21b. Both methods perform similarly.

3.7 The $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ decay mode

The branching ratio of the $\tau \to \mu$ decay mode is equivalent to the one of the $\tau \to e$ decay mode, which means that the expected ν_τ event rate associated to these two decays modes is the same. However, the analysis associated to the $\tau \to e$ decay mode takes advantage of the small presence of prompt electron neutrinos in the beam and of the fact that the $\nu_\mu \to \nu_e$ oscillation channel is sub-leading in comparison to the two other oscillation channels. As a consequence,



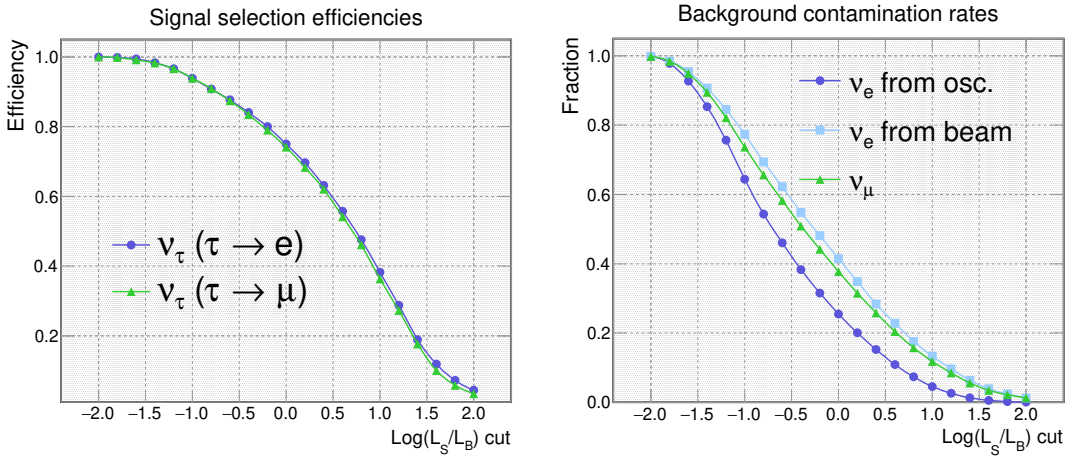
(a) (b)

Figure 3.21: Artificial Neural Network built with the Keras library performance on the signal/background separation task. The distributions of signal and background NN scores are shown on (a), while the ROC curve on (b) allows to visualize (orange) the signal efficiency versus background contamination, for cuts values varying in $[0\,;\,1]$. For comparison, the best likelihood performance obtained is shown in black triangles.

the ν_e background is much smaller than the ν_μ background associated to the $\tau \to \mu$ decay mode analysis. However a large fraction of the ν_μ originally present in the beam oscillate in ν_τ and this contributes in decreasing the ν_μ background at the far detector. The table 2.2 indeed illustrates that the 3.5 years staged initial signal to background ratio is $17.41\% \times 270/9660 = 47/9660$, that is to say about 0.5% (17.41% is the branching ratio of the $\tau \to \mu$ decay mode). As for comparison the initial signal to background ratio of the $\tau \to e$ analysis was at the level of 3% which is 6 times more. A demanding signal/background separation power is thus needed for DUNE to be reasonably sensitive to the ν_τ appearance using the muon decay mode. For this reason, this decay mode has been discarded so far.

We propose in this section, still, to evaluate the signal/background separation power of the $\tau \to \mu$ decay mode. We consider at first order that the transverse plane kinematics of ν_μCC interactions are equivalent to the one of the ν_eCC interactions which come from the $\nu_\mu \to \nu_e$ oscillations. We also consider that the transverse plane kinematics of $\nu_\tau(\tau \to \mu)$ is equivalent to the one of $\nu_\tau(\tau \to e)$. This way, the $\nu_\tau(\tau \to \mu)/\nu_\mu$ separation is based on the kinematical distributions already obtained in 3.4.1. In this section the detector effects are taken into account via the smearing method.

Then the simulated $\nu_\tau(\tau \to \mu)$ and ν_μCC events are tested under the signal/background hypothesis and the corresponding log-likelihood ratio distributions are built. No further discriminating optimization was found than the one presented in 3.3.1. The corresponding signal efficiency is shown in figure 3.22 (left). It reveals that the likelihood based selection efficiency of $\nu_\tau(\tau \to \mu)$ relying on transverse plane kinematics is very close to the one obtained for the $\nu_\tau(\tau \to e)$. The background contamination as a function of the log-likelihood ratio cut is shown in figure 3.22 (right). At first order it shows that the likelihood equivalently rejects the ν_μ background and the beam ν_e of the $\tau \to e$ analysis. However we find that the likelihood rejects better the oscillated ν_e than the ν_μ of the $\tau \to \mu$ decay mode analysis. A possible explanation is that the likelihood better rejects low energy backgrounds. Since the low energy ν_μ are disfavoured by the $\nu_\mu \to \nu_\mu$ oscillations, the ν_μ rejection efficiency is smaller than the one for the ν_e (coming from oscillations) rejection. Further investigations would however be necessary to fully understand this feature.



(a) (b)

Figure 3.22: Signal efficiencies (left) and background contamination rates (right) for the two $\tau \rightarrow e$ and $\tau \rightarrow \mu$ decay mode analysis. The results for the $\tau \rightarrow \mu$ analysis are shown in green with triangle markers.

We finally report a significance estimation. A log-likelihood ratio cut value of 0.4 is found to optimize the Asimov significance at the level of 0.6σ which corresponds to a signal/background ratio of $29/2496 \simeq 1\%$. In comparison a statistical significance of 1.8σ was found for the $\tau \rightarrow e$ analysis using the same 3.5 years staged normalization.

3.8 Summary

This chapter proposed a kinematic search of ν_τCC interactions (coming from the $\nu_\mu \rightarrow \nu_\tau$ oscillation channel) using the $\tau \rightarrow e$ leptonic decay mode. The main backgrounds associated are the ν_eCC interactions coming from the muon neutrino oscillations (oscillated ν_e) and the ν_e beam flavour contamination (beam ν_e).

$\bar{\nu}_\tau$ and $\bar{\nu}_e$, though present in minority in the neutrino beam, were discarded from the analysis. A more complete analysis would require to take them into account since it is not clear that DUNE will be able to separate neutrino from antineutrino events. Based on kinematics in the transverse plane of the interactions, the oscillated ν_e, which constitute the dominant background component, are better rejected than the beam ν_e. A likelihood based analysis allows for a 40% signal efficiency and a 95% (85%) rejection of oscillated (beam) ν_e with the standard LBNF neutrino flux optimized for the CP violation study and taking into account the LArTPC detector performance.

These efficiencies are normalized following the 3.5 years staged deployment plan of DUNE, and it is found that a significance of 1.8σ is reached in regards of the $\nu_\mu \rightarrow \nu_\tau$ appearance phenomenon. It corresponds to a signal/background ratio of about 22/151. Only a slight improvement of the significance is achieved when the analysis is restricted to events with a QEL-like topology for which the final state is composed of only one proton and one electron.

A major improvement is however found when the high energy τ optimized neutrino beam is used instead of the CP optimized one. The boost in ν_τCC statistics by a factor of 6 allows to reach a ν_τ appearance significance of 9 assuming the same 3.5 years staged deployment plan which is

equivalent to 10 years of exposure with one single 10 kt far detector module. A parallel study with machine learning techniques (Boosted Decision Trees and Artificial Neural Networks) was found to have a similar signal/background separation power as the likelihood-based analysis.

Finally a quick overview of the $\tau \to \mu$ decay mode analysis was performed, in which the kinematics of the signal and the background are assumed to be that of the $\tau \to e$ decay mode analysis. The initial large amount of background ($S/B \simeq 0.5\%$) makes this decay mode unfavourable for the search of ν_τCC interactions in DUNE. A maximal significance of 0.6σ for the 3.5 years staged normalization and the standard LBNF neutrino flux was reached, in comparison to 1.8σ for the ν_τ search via the $\tau \to e$ decay mode.

⋆ ⋆ ⋆

Chapter 4

The $\tau^- \to \rho^- \nu_\tau \to \pi_0 \pi^- \nu_\tau$ decay mode analysis

Contents

The leptonic decay channels of the τ amount to about 35% of all possible decay modes, which means that with an analysis on these decay modes only would allow to catch at best one third of the tau neutrino charged current events. Moreover it should be noted that the muon decay mode has a large level of ν_μCC background. A complementary analysis including semi-leptonic decay modes is thus required to extend DUNE's capability of identifying ν_τ interactions. The τ lepton has a large branching ratio (25.52%) involving the creation of a ρ meson in a resonant state which quickly decays to a pair of pions (100%). Consequently, reconstructed invariant mass from the final state particles corresponding to that of the ρ meson would constitute a characteristic signature of this τ decay mode.

This chapter first gives a brief description of the ρ meson and then tackles its reconstruction in ν_τ charged current interactions in which the τ decays into a ρ (further abbreviated $\nu_\tau(\tau \to \rho)$). The electrical charge of the lepton τ^- and the meson ρ^- created in $\nu_\tau(\tau^- \to \rho^-)$ interactions might later be omitted for simplicity. Then a likelihood based analysis to enhance signal/background ratio for this channel is presented. The possibility to identify the charge of the pions via specific final state topologies is explored in a dedicated section, as well as the possibility for DUNE to run with the so-called τ optimized neutrino beam tuned to foster the detection of ν_τCC interactions.

4.1 The ρ meson

The ρ meson is an isospin triplet with spin 1 whose three states are $\rho^+ = (u\bar{d})$, $\rho^- = (\bar{u}d)$ and $\rho_0 = (u\bar{u} - d\bar{d})/\sqrt{2}$. Since the triplet is made with the same quarks as the pion triplet (which has spin 0), the ρ meson can be thought as an excited state in which the spins of the quarks are parallel, resulting in a larger mass. Indeed the mass of the charged $\rho^\pm$ is $0.776\,\mathrm{GeV}/c^2$ while the mass of the charged pion is $0.140\,\mathrm{GeV}/c^2$

The ρ decay is mediated by gluons. Its very short lifetime, at the level of 10^{-24}s, implies that it can not be observed in any detector. The interesting point is that the charged $\rho^\pm$ decays into a pair of pions ($\pi^\pm\pi_0$) which are particles observable in LArTPC detectors. The reconstruction of neutral pions is not trivial since it relies on the reconstruction of the two electromagnetic showers produced by the decay photons. Each photon can be identified by the electromagnetic shower produced in LAr after conversion and by the specific ionization energy losses at the start of the shower. Indeed the first created e^-/e^+ pair leaves a track for which the ionizing signal is twice the one a single electron/positron would leave. This point is for instance discussed in [60].

4.2 Two-body decay process

The rate of a two-body decay $A \to B + C$ is given by the Fermi Golden rule:

$$d\Gamma = |\mathscr{M}_{fi}|^2 (2\pi)^4 \delta^4(\mathbf{p}_b + \mathbf{p}_c - \mathbf{p}_a) \frac{1}{2m_a} \prod_{f=b,c} \frac{d^3\vec{p}_k}{2E_f}. \tag{4.1}$$

$\mathscr{M}_{fi}$ is the matrix containing the physics describing the decay and is calculated following the Feynman rules. An additional correcting factor $1/n!$ must be taken into account in case of n

identical final state particles to avoid multiple counting. The decay rate Γ defined in equation 4.1 is related to the probability that the unstable particle survives after a given time t.

Assuming indeed a system at time $t = 0$ containing $N(t = 0)$ unstable particles decaying with the rate Γ, the number of remaining non-decayed particles $N(t)$ at time t is given by:

$$N(t) = N(t = 0) \exp(-\Gamma t). \tag{4.2}$$

$T = 1/\Gamma$ corresponds to the time after which the particles system has $1/e$ of its original population. In particle physics this time is defined as the lifetime of the particle.

The lifetimes for the neutral pion and the ρ meson are taken from the Particle Data Group:

$$\left\{ \begin{array}{lcl} T_{\pi_0 \to \gamma\gamma} & = & (8.52 \pm 0.18) \times 10^{-17}\text{s} \\ \Gamma_{\pi_0 \to \gamma\gamma} & = & 7.7\,\text{eV} \end{array} \right. \quad \left\{ \begin{array}{lcl} T_{\rho^\pm \to \pi^\pm \pi^0} & = & 4.5 \times 10^{-24}\text{s} \\ \Gamma_{\rho^\pm \to \pi^\pm \pi^0} & = & 0.146\,\text{GeV} \end{array} \right. \tag{4.3}$$

An important consequence of the decay rate is that if an experiment aims to measure the mass of an unstable particle by measuring the invariant mass of its daughter particles, the distribution of the invariant mass will come up with an irreducible width related to the decay rate Γ. This feature can be understood thanks to the time-energy uncertainty principle $\Delta E \Delta t \simeq \hbar/2$. For a short lived particle there is a large uncertainty on its energy which implies a width in the distribution of the invariant mass. For this reason the decay rate may sometimes be expressed as the decay width (though the units of the two are not the same). The mass of an unstable particle is commonly defined as the peak of the corresponding distribution. The width of the invariant mass distribution reflects the lifetime of the particle.

The energy resolution of the DUNE detectors will not allow resolving the small decay width of the neutral pions, however that of the ρ mesons will be accessible. The ρ invariant mass is defined as:

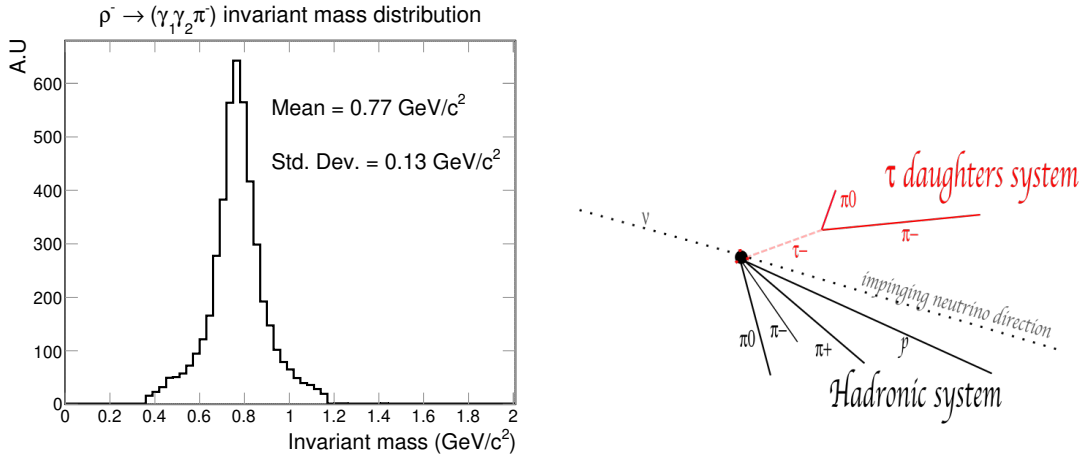
$$M_\rho^{(inv)\,2} = (\mathbf{p}_{\pi_0} + \mathbf{p}_{\pi^-})^2 = (\mathbf{p}_{\gamma_1} + \mathbf{p}_{\gamma_2} + \mathbf{p}_{\pi^-})^2, \tag{4.4}$$

where the square are Lorentz squares and the bold notations $\mathbf{p}$ designate the 4-momenta of the ρ daughter particles.

We selected, in the DUNE simulation files produced for the Technical Design Report, a sample of ν_τCC events for which the τ^- decays into a ρ^- meson. About 30 000 such events were found and the corresponding ρ^- invariant mass distribution calculated from the 4-momenta of the daughter system is shown in figure 4.1a. The mean and the standard deviation of the distribution are in agreement with the mass and the decay width of the ρ^- meson.

4.3 ρ reconstruction

Within a ν_τCC event where the charged lepton decays as $\tau^- \to \rho^- + \nu_\tau$, the definition of the set of ρ^- daughter particles (*i.e* the pair of pions $\pi_0\pi^-$) can be blurred by combinatorics due to additional pions belonging to the hadronic system of the neutrino interaction. Thus there is a necessity to first develop an analysis to correctly select the ρ decay daughters.



(a) Invariant mass plot

(b) Scheme of a ν_τCC($\tau \to \rho$) event

Figure 4.1: Left: ρ invariant mass distribution of $\mathcal{O}(30\,000)$ ν_τCC events for which the τ^- decays into a ρ^-. Right: Scheme of a ν_τCC($\tau \to \rho$) where the ρ^- decays into a pair of pions ($\pi_0\pi^-$) (red). The hadronic system is shown in black and is composed of one proton, one neutral pion and two charged pions.

Neutral pions can only be reconstructed via their decay daughters. In this chapter they are assumed to decay 100% into a pair of photons, isotropically in their rest frame. A 1% branching ratio corresponding to the Dalitz decays into a photon plus an electron/positron pair is neglected. If a neutrino interaction produces two or more π_0, there exists an ambiguity on the π_0 reconstruction. We assume that any possible pair of photons could define a π_0 candidate. However if the reconstructed invariant mass square of the two photons is negative or if the reconstructed kinetic energy of the corresponding neutral pion (taken as the sum of the energies of the two photons, from which the mass of the π_0 is subtracted) is negative, the neutral pion candidate is discarded from the analysis flow and will not be used to define a ρ candidate.

In the following, the term "**true** ρ" will be employed to denote the triplet of final state particles ($\gamma_1\gamma_2\pi^-$) which really corresponds to the full decay chain of the charged lepton $\tau^- \to \rho^-\nu_\tau \to \pi_0\pi^-\nu_\tau \to \gamma_1\gamma_2\pi^-\nu_\tau$ as identifiable by using the Monte Carlo truth. Any other ($\gamma_1\gamma_2\pi^\pm$) triplet that is not entirely composed of τ daughter particles will be referred as a "**fake** ρ". By extension the labels "true" and "fake" will also be employed to denote the photons and pions themselves. A fake decay daughter candidate is a final state photon or charged pion which is not a daughter of the decayed τ.

A subtlety must be noted: the hadronic system of the ν_τ interactions might also contain real ρ resonances which also decay into a pair of pions. We emphasize that the terminology "true" and "fake" refers only to the selection of the τ decay daughters.

A schematic view of a ν_τCC($\tau \to \rho$) interaction is depicted in figure 4.1b. The leptonic (hadronic) system is represented in red (black). Neutral pions are not decayed in this example for sake of visual simplicity. We propose to illustrate in the next section the number of ρ candidates available in the final state for this specific interaction example.

4.3.1 Illustration of the level of ambiguity in the ρ reconstruction

The correct reconstruction of the genuine ρ from the τ decay in ν_τCC event requires identifying the triplet of particles in the final state which are the daughter system of the ρ. The ν_τCC interaction illustrated in figure 4.1b corresponds to a deep inelastic scattering event with a final state including the charged lepton τ^-, a proton and three pions. The leptonic system (red) provides the true particles γ_1, γ_2 and π^- and the hadronic system provides one fake neutral pion (two fake photons) and two fake charged pions. It should be noted that in general the hadronic system may contain additional photons not coming from a neutral pion decay.

Deep inelastic interactions on neutrons are about twice more probable than on protons since the neutron contains two d valence quarks, so this configuration is not unlikely to happen but it is the most probable for a deep inelastic event. Assuming that we are not able to distinguish in the event reconstruction the sign of the pion, both the negative and positive pions from the hadronic system may be selected when looking for the charged pion of the ρ decay.

One way of enumerating the ρ combinatorics is to distinguish the cases where we correctly picked up 0, 1, 2 or 3 correct daughter particles of the true ρ system:

- There is only **1 candidate** ($\gamma_1\gamma_2\pi^-$) corresponding to the true ρ.
- If the triplet contains two true particles, then one particle is fake. Two cases are possible:
  - The fake particle is a photon. There are two fake photons available which can replace one of the two true photons. Thus there are **4 such candidates**.
  - The fake particle is the charged pion, there are **2 such candidates** since there are two fake charged pions available.
- Only one particle of the triplet is true and the two others are fake. Three sub-cases can be distinguished:
  - If the charged pion is the true particle, then the two photons are fake. Since the hadronic system only provides two fake photons, there is **1 such candidate**.
  - One of the two photons, say γ_1, is the true particle. Then γ_2 and the charged pion are fake. Two fake photons and two fake charged pions are available, so there are **4 such candidates**.
  - This last case is obtained by swapping the role of γ_1 and γ_2. This adds **4 candidates**.
- The three particles of the triplet are fake, there are **2 such candidates**.

This ν_τCC event with an hadronic system providing among others two charged pions and one neutral pion (two photons) thus contains 18 ρ candidate combinations among which the true one must be chosen.

4.3.2 Fake ρ multiplicity in the simulated $\nu_\tau(\tau \to \rho)$ events in DUNE

The fake ρ multiplicity of the simulated $\nu_\tau(\tau \to \rho)$ is shown in table 4.1a. It appears that about half of the events have no fake ρ candidate. The cases where the hadronic system provides one neutral pion and no charged pion result in a fake ρ multiplicity of 5. This is why in the table there is a bump at 7% for this multiplicity. Table 4.1b distinguishes the fake ρ multiplicity with

Fake ρ multiplicity per event	0	1	2	3	4	5	>6
fraction (%)	54	18	4	2	1	7	15

(a)

Fake ρ multiplicity per event	0	1	[2;5]	[6;50]	>50
QEL (46%)	93	4	2	<1	<0.1
RES(23%)	21	47	25	8	< 0.1
DIS(26%)	6	18	26	41	9

(b)

Table 4.1: Fake ρ multiplicity in the sample of simulated ν_τCC ($\tau \to \rho$) events(a). The second table (b) splits the fake ρ multiplicity with respect to the dominant scattering processes QEL, RES and DIS. The relative abundance of each is given between brackets.

respect to the QEL, RES and DIS scatterings, and for each of them the corresponding fraction observed in the ν_τCC interactions is shown between brackets.

For the QEL which accounts for 46% of the interactions, the fake ρ multiplicity is low and is zero in 93% of the cases. Nuclear effects which can result in the creation of pions allow explaining that the fake ρ multiplicity is non-zero in few percents of the cases. A small non-zero (<0.1%) fraction of QEL scatterings have a very high ρ multiplicity (>50). It is found that these cases correspond to charmed baryon production which can decay into numerous pions. The charmed baryon production in QEL scatterings was discussed in 2.1.3. RES and above all DIS scatterings have a higher fake ρ multiplicity than QEL scatterings.

The first requirement for the ν_τ search analysis in the $\tau \to \rho$ decay mode is to develop a method that makes it possible to correctly reconstruct the true ρ of the ν_τCC($\tau \to \rho$) events. In order to get some help in doing so, we will exploit the scatter plots of the invariant masses. For a given ρ candidate, one can indeed calculate the invariant mass of the two photons $M_{\pi_0}^{(inv)}$ (or equivalently the invariant mass of the π_0 candidate contained in the ρ candidate) and the invariant mass of the whole triplet $M_\rho^{(inv)}$. It is expected that the 2-dimensional distribution $[M_{\pi_0}^{(inv)}\,;\, M_\rho^{(inv)}]$ of the true ρ be distributed close to the point with coordinates $[0.135\,;\,0.776]\,\mathrm{GeV}/c^2$ which are the masses of the neutral pion and the charged ρ meson.

4.3.3 Study of the ρ invariant mass distribution

It has been pointed out that the lifetime of the π_0 is large enough so that its decay width can be neglected in the analysis for DUNE. For the ρ meson the decay width is not negligible and we propose to model it with a Breit-Wigner distribution:

$$\sigma(M_\rho^{(inv)}, \Gamma, M_0) = \sigma_{max} \frac{\Gamma^2/4}{(M_\rho^{(inv)} - M_0)^2 + \Gamma^2/4}, \tag{4.5}$$

where the parameters Γ and M_0 are respectively the width and the central mass value of the resonance. Indeed, when $M_\rho^{(inv)} - M_0 = \pm\Gamma/2$, the height of the distribution is half of its peak value. σ_{max} is merely a global factor that in our case does not carry any physical information since it is related to the raw number of events considered. However in the general case this factor is spin dependant (see for instance discussion in [104, section 7.3]). For relativistic particles the Breit-Wigner distribution is modified as:

$$\begin{aligned} \sigma^{(Rel)}(M_\rho^{(inv)}, \Gamma, M_0) &= \sigma_{max}^{(Rel)} \frac{k}{(M_\rho^{(inv)} - M_0)^2 + M_0\Gamma^2} \\ k &= 2\sqrt{2M_0\Gamma\gamma}/(\pi\sqrt{M_0^2+\gamma}) \\ \gamma &= \sqrt{M_0^2(M_0^2+\Gamma^2)}. \end{aligned} \tag{4.6}$$

In the simulated neutrino events generated for the DUNE Physics TDR, the intermediate ρ states of the τ decays were not explicitly recorded, however the pions from the ρ decay were identified as the charged τ lepton daughters. The $(\pi_0\pi^-)$ invariant mass distribution shown in figure 4.1a confirms the presence of the ρ resonance which is well simulated underneath. An histogram with a bin size of $0.03\,\mathrm{GeV}/c^2$ is filled with the calculated ρ invariant mass for each τ neutrino event and is then fitted with the Breit-Wigner and relativistic Breit-Wigner functions defined respectively in equations 4.5 and 4.6.

Results are shown in figures 4.2a (Breit-Wigner fit) and figure 4.2b (relativistic Breit-Wigner fit). Surprisingly the non-relativistic Breit-Wigner function fits the simulated data better than the relativistic one, though one would expect the ρ to be relativistic at DUNE energies. The fitting parameters numerical values confirm the presence of the ρ meson with a mass of $0.77\,\mathrm{GeV}/c^2$ and a decay width of $0.15\,\mathrm{GeV}/c^2$.

A complementary cross-check is performed using a sample of ρ resonances explicitly selected from the hadronic systems of the simulated τ neutrino events. The corresponding invariant mass of the ρ daughter particles is calculated and fitted with Breit-Wigner and relativistic Breit-Wigner functions. The results are shown respectively in figures 4.2c and 4.2d.

The results obtained for the fitted parameters slightly differ from that obtained with the $\tau \to \rho$ sample. The decay widths of the two are rather similar but the peak of the resonance has a smaller value for the ρ simulated in the hadronic systems ($0.75\,\mathrm{GeV}/c^2$ with the Breit-Wigner fit and $0.76\,\mathrm{GeV}/c^2$ with the relativistic Breit-Wigner fit) than the ρ simulated in the τ decays ($0.77\,\mathrm{GeV}/c^2$). Moreover the Breit-Wigner function fits better the ρ from τ decays while the relativistic Breit-Wigner function fits better the ρ simulated in the hadronic systems. The difference is better appreciated in figure 4.3 where the two distributions are put on the same scale. The two ρ samples are simulated with different libraries (TAUOLA for $\tau \to \rho$ *vs* PYTHIA for ρ in the hadronic system), so that can explain the small differences. Though we emphasized these differences, a further discussion on which library better reflects data goes beyond the scope of this thesis.

It should be noted that the differences observed between the two ρ samples are rather small since they are at the level of 10 MeV for the mass of the ρ meson. In the following, we will use for the ρ analysis selection criteria the values: $m_{\pi_0} = 0.135\,\mathrm{GeV}/c^2$ and $m_\rho = 0.776\,\mathrm{GeV}/c^2$.

4.3.4 The invariant mass selection criterion

We pointed out that within a ν_τCC event where the τ lepton decays into a ρ meson, the decay system of the τ could be blurred by pions coming from the hadronic system. We propose in this section a method to identify the true ρ meson daughter particles. The sample used in the present study consists of about 30 000 $\nu_\tau(\tau \to \rho)$ charged current events. It was simulated by the DUNE collaboration using GENIE v2.12.00 for the DUNE Technical Design Report [41]. We recall that the neutral pions are decayed isotropically in their rest frame into a pair of

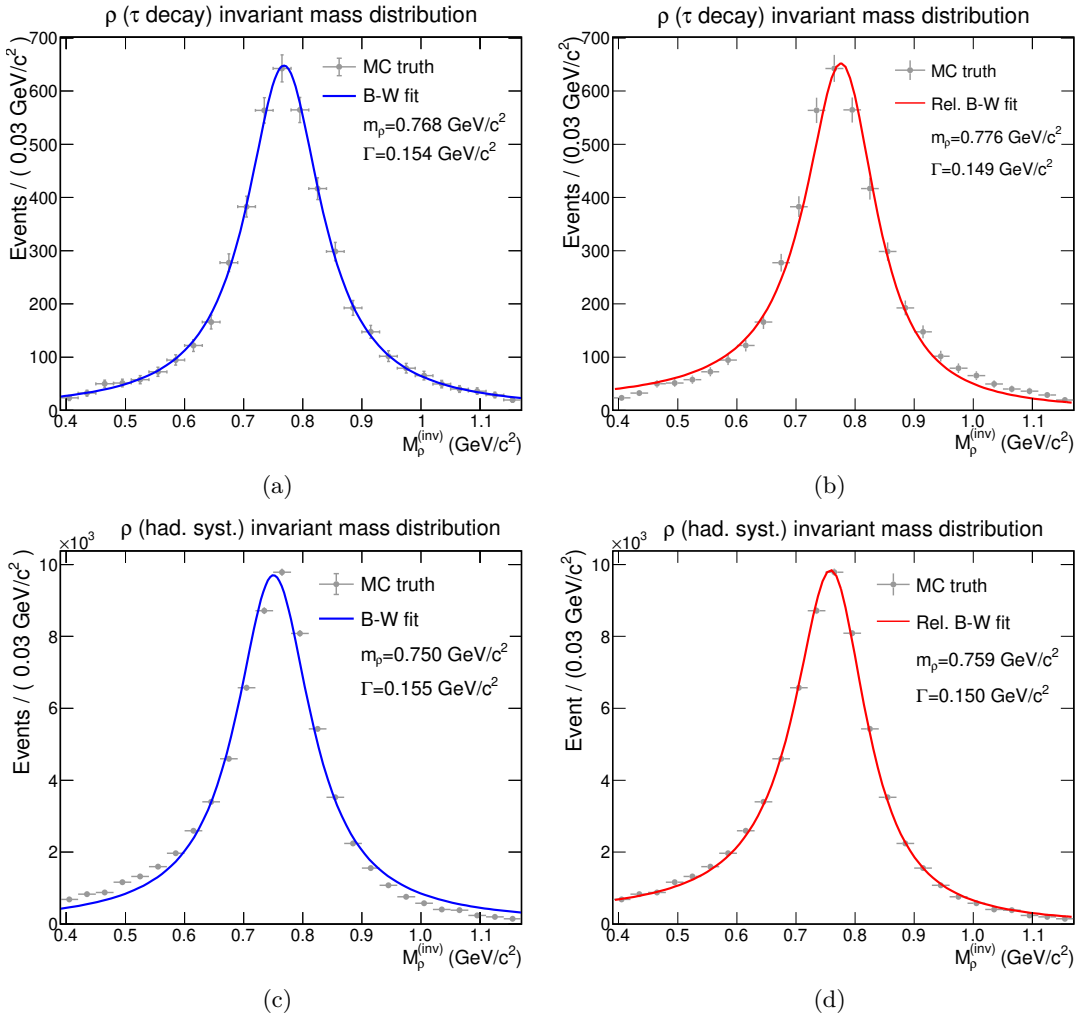


Figure 4.2: Top: Breit-Wigner (left) and relativistic Breit-Wigner (right) fits of ρ invariant mass distribution coming from τ decays. Bottom: Breit-Wigner (left) and relativistic Breit-Wigner (right) fits of ρ invariant mass distribution belonging to hadronic systems.

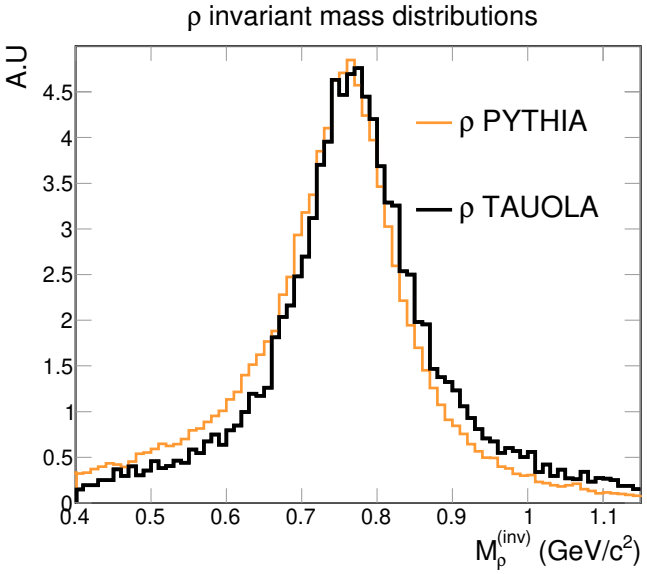


Figure 4.3: ρ invariant mass distributions of hadronic ρ simulated by the library PYTHIA (orange), and ρ of τ decays simulated by TAUOLA (black).

photons and that the decay mode into a photon and a pair e^+/e^- with a branching fraction of 1% is neglected. If more than two π_0 are simulated in the final state, there exists an ambiguity in the reconstruction of the π_0 in the sense that for some pairs of reconstructed photons ($\gamma_1\gamma_2$) have γ_1 and γ_2 stemming from different neutral pions.

In this section the detector resolution and efficiency effects are taken into account via the smearing method which was described in 2.5. The detector is assumed to have an ideal particle identification performance. Final state photons and charged pions with a null energy after smearing or an energy below their expected detection threshold are considered undetected and thus do not participate in the ρ multiplicity counting.

All available photons and charged pions are grouped in triplets following the method exposed in 4.3.1. For each ρ candidate thus defined the $M_{\pi_0}^{(inv)}$ and $M_\rho^{(inv)}$ invariant masses are calculated, from which a corresponding score $d_{M_{inv}}$ is deduced:

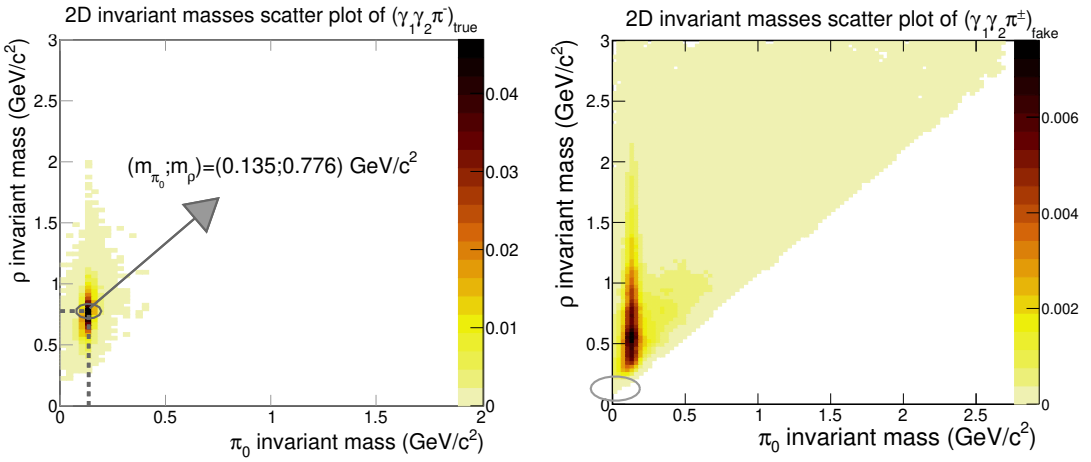
$$M_{\pi_0}^{(inv)} = \sqrt{(\mathbf{p}_{\gamma_1} + \mathbf{p}_{\gamma_2})^2} \tag{4.7}$$

$$M_\rho^{(inv)} = \sqrt{(\mathbf{p}_{\gamma_1} + \mathbf{p}_{\gamma_2} + \mathbf{p}_{\pi^\pm})^2} \tag{4.8}$$

$$d_{M_{inv}} = \sqrt{(M_{\pi_0}^{(inv)} - m_{\pi_0})^2 + (M_\rho^{(inv)} - m_\rho)^2}. \tag{4.9}$$

$$m_{\pi_0} = 0.135\,\mathrm{GeV}/c^2 \;\;;\;\; m_\rho = 0.776\,\mathrm{GeV}/c^2. \tag{4.10}$$

It should be noted that the squares of the two first lines are Lorentz squares of 4-momenta whereas the squares of the third line are the usual squares of real numbers. $d_{M_{inv}}$ is merely the 2-dimensional distance between the point with coordinates $[M_{\pi_0}^{(inv)}\,;\,M_\rho^{(inv)}]$ and the point whose coordinates are the known masses of the π_0 and the ρ. For each ν_τCC event all the ρ candidates are classified in ascending order with respect to this invariant mass metric so that the true ρ daughter system is given a rank ≥ 1.



(a) (b)

Figure 4.4: 2-dimensional distributions of the ρ invariant mass $M_\rho^{(inv)}$ against the $\pi_0 = (\gamma_1\gamma_2)$ invariant mass $M_{\pi_0}^{(inv)}$. The left (right) distribution contains only true (fake) ρ candidates. On the left plot there is a one to one correspondence between the number of simulated ν_τCC events studied (given that the true ρ is reconstructed) while on the right a given event may contribute several times to the distribution. See later in the text for the gray circle discussion.

It was pointed out in table 4.1a that in about half of the ν_τCC events there are no fake ρ candidates. In this case the true ρ, providing that it gets reconstructed, is ranked 0 (no ambiguities) as distinct from the case where the true ρ is ranked 1 among other candidates. If the true ρ candidate is not reconstructed, the triplet is ranked -1 even if there is no fake ρ candidate available. The -1 rank enquires on the fraction of signal that will be lost even before the start of the analysis. The rank 0 enquires on the fraction of signal where there is no need to deploy an analysis to correctly reconstruct the ρ system from the τ decay.

The 2-dimensional distributions of the invariant masses $[\ M_{\pi_0}^{(inv)}\ ;\ M_\rho^{(inv)}\ ]$ are shown in figure 4.4a for the true ρ and in figure 4.4b for all the reconstructed fake ρ. The distribution associated to the true ρ clusters in the region distributed around the point of coordinates $[m_{\pi_0}\,;\,m_\rho]$ with some dispersion due to the smearing effects and the decay width of the ρ which is about $0.140\,\mathrm{GeV}/c^2$. The distribution associated to the reconstructed fake ρ candidates in all ν_τCC events is more scattered on the plane. One can observe a vertical band at $0.13\,\mathrm{GeV}/c^2$, corresponding to fake ρ candidates for which a real π_0 was used in the triplet. This vertical band is pretty populated in the $M_\rho^{(inv)}$ region between $[0.3\,;\,0.9]\,\mathrm{GeV}/c^2$ which is quite close to the points of coordinates $[m_{\pi_0}\,;\,m_\rho]$. It indicates the existence of fake ρ candidates for which the reconstructed invariant masses are compatible with that of a ρ decay, and thus a contamination is to be expected in the true ρ reconstruction.

Figure 4.4b gives an idea of the fake ρ invariant mass distribution, however it is built in a way that one ν_τ event contributes to the histogram as much as the fake ρ multiplicity of the event. The distribution also displays an interesting forbidden region in the invariant mass plane that we propose, as for curiosity, to investigate.

The forbidden region of the invariant masses scatter plot

The fake ρ candidates of figure 4.4b never populate the invariant masses region below the line of coordinate $M_\rho^{(inv)} \simeq M_{\pi_0}^{(inv)}$. Let us investigate more closely this relation:

$$
\begin{aligned}
M_{\rho}^{(inv)2} &= (\mathbf{p}_{\pi^{\pm}} + \mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2})^2 && (4.11)\\
&= \mathbf{p}_{\pi^{\pm}}^2 + (\mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2})^2 + 2\mathbf{p}_{\pi^{\pm}} \cdot (\mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2}) && (4.12)\\
&= m_{\pi}^2 + M_{\pi_0}^{(inv)2} + 2\mathbf{p}_{\pi^{\pm}} \cdot (\mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2}), && (4.13)
\end{aligned}
$$

where m_{π} is the mass of the charged pion. This relation illustrates that in the limit where the 3 particles of the triplet have a null energy, the ρ invariant mass can not decrease below the mass of the charged pion ($0.140\,\mathrm{GeV}/c^2$). This feature can indeed be observed in figure 4.4b (see gray circle at the bottom left).

To further extend the Lorentz scalar product of equation 4.13, let us write $\theta_{\pi\gamma_1}$ the angle between the 3-momentum of the charged pion and the first photon. Using the notation $\mathbf{p}_{\pi^{\pm}} = (E_{\pi}\ \vec{p}_{\pi})$ and $\mathbf{p}_{\gamma_1} = (E_{\gamma_1}\ \vec{p}_{\gamma_1})$:

$$
\begin{aligned}
\mathbf{p}_{\pi^{\pm}} \cdot \mathbf{p}_{\gamma_1} &= E_{\pi}E\gamma_1 - ||\vec{p}_{\pi}||E_{\gamma_1}\cos(\theta_{\pi\gamma_1})\\
&\geq E_{\gamma_1}(E_{\pi} - \underbrace{||\vec{p}_{\pi}||}_{=\beta_{\pi}\gamma_{\pi}m_{\pi}})\\
&\geq E_{\gamma_1}E_{\pi}(1-\beta_{\pi}),
\end{aligned}
$$

where β_{π} and γ_{π} denote the standard Lorentz factors of the charged pion. Then the ρ invariant mass is greater than:

$$
M_{\rho}^{(inv)2} \geq m_{\pi^{\pm}}^2 + M_{\pi_0}^{(inv)2} + 2E_{\pi}(1-\beta_{\pi})\underbrace{(E_{\gamma_1} + E_{\gamma_2})}_{=E_{\pi_0}}. \tag{4.14}
$$

E_{π_0} is the total energy of the decayed neutral pion candidate and it can be related to the π_0 invariant mass through the relation:

$$
(E_{\gamma_1} + E_{\gamma_2})^2 = M_{\pi_0}^{(inv)\,2} + (\overrightarrow{p}_{\gamma_1} + \overrightarrow{p}_{\gamma_2})^2. \tag{4.15}
$$

It is rather obvious that $(E_{\gamma_1} + E_{\gamma_2}) \geq M_{\pi_0}^{(inv)}$. It implies that the ρ invariant mass is greater than the function f_{min} defined as:

$$
f_{min}(M_{\pi_0}^{(inv)}) = \sqrt{m_{\pi^{\pm}}^2 + M_{\pi_0}^{(inv)2} + 2E_{\pi}(1-\beta_{\pi})M_{\pi_0}^{(inv)}}. \tag{4.16}
$$

At the DUNE considered energies it is likely that the $\pi^{\pm}$ energy rarely exceeds 5 GeV which corresponds to $\beta_{\pi} = 0.999$.

The function f_{min} is superimposed to the 2-dimensional invariant mass distribution of the fake ρ candidates in figure 4.5 for a fixed value $E_{\pi} = 5\,\mathrm{GeV}$. The function describes quite well the forbidden region of the scatter plot. It should be noted that the arbitrary choice $E_{\pi} = 5\,\mathrm{GeV}$ only matters in the low ($< 1\,\mathrm{GeV}/c^2$) $M_{\pi_0}^{(inv)}$ region. For high $M_{\pi_0}^{(inv)}$ we indeed recover the relation $M_{\rho}^{(inv)} \simeq M_{\pi_0}^{(inv)}$.

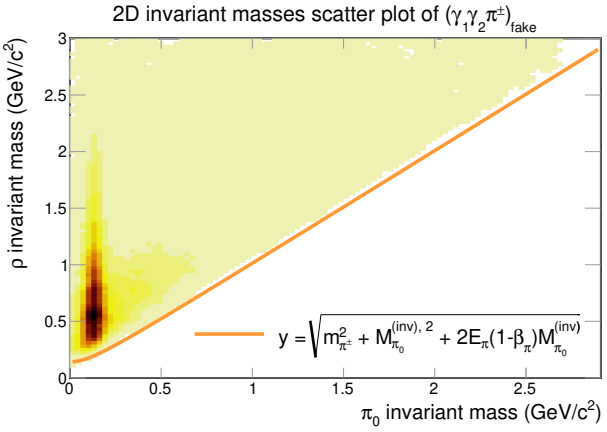


Figure 4.5: 2-dimensional invariant mass distribution of the fake ρ candidates where the function f_{min} defined in equation 4.16 was superimposed for a fixed value $E_\pi = 5\,\text{GeV}$.

Rank	-1	0	1	2	3	>3
fraction (%)	2.9	52.6	21.9	10.0	3.1	9.6

Table 4.2: Results of the ranking method based on the invariant mass metric $d_{M_{inv}}$ to identify the true ρ in the sample of simulated $\nu_\tau(\tau \to \rho)$ events. A null rank is attributed to true ρ for which there exists no ambiguity in the reconstruction. A negative rank of -1 is attributed to true ρ for which the reconstruction is not possible. It occurs for instance when one of the particles of the true ρ decay is undetected.

Performance of the invariant mass criterion for the true ρ triplet selection

The purpose of the metric $d_{M_{inv}}$ is to help at reconstructing the true ρ out of all the ρ candidates in each ν_τCC event. The true ρ is expected to have a low score and thus be identifiable that way. The performance of the ranking method based on the invariant mass score is given in table 4.2.

To begin with, for about half of the ν_τ events the true ρ is ranked 0 which means that the corresponding fake ρ multiplicity is null. The observed fraction (52.6%) is a bit smaller than the 54% observed in table 4.1a. The difference is due to cases where the true ρ is not reconstructed even though there is no fake ρ candidate available. In these cases the true ρ is ranked -1. The table shows that in about 3% of the cases, the true ρ is not reconstructed.

The fraction of events with a rank greater than or equal to 1 is 44.5% and the fraction of events with a rank equal to 1 is 21.9%. It means that the invariant mass score allows to identify the correct ρ for about half of the ν_τ events where the true ρ is reconstructed but there exists an ambiguity in the ρ reconstruction.

In total the correct ρ is identified in $52.6 + 21.9 = 74.5\%$ of the ν_τ interactions for which the τ decays into a ρ. It can be noted that in about 13% of the cases the true ρ got ranked either 2^{nd} or 3^{rd}. In other words, there is a rather significant amount of events where the reconstruction of the true ρ was almost correct. The invariant mass criteria can be associated to other kinematical variables in order to improve the ρ reconstruction performance.

Rank	-1	0	1	2	3	>3
Invariant mass ranking method (%)	2.9	52.6	21.9	10.0	3.1	9.6
Medal Game (%)	2.9	52.6	29.2	6.7	1.9	6.7

Table 4.3: Results of the ρ reconstruction based on the Medal Game and comparison with the method making use of the invariant mass metric $d_{M_{inv}}$ alone. The Medal Game allows for an increase from 21.9% to 29.2% of true ρ ranked 1.

4.3.5 Improved ρ ranking method: the Medal Game

Several kinematical variables were tested to assess their ability to improve the performance of the true ρ reconstruction. Two of them were selected:

- The ρ energy defined as $\rho_{E_K} = E_K^{\pi^\pm} + E_K^{\pi_0}$, where E_K denotes a kinetic energy. The variable is merely the sum of the kinetic energies of the particles defining the ρ candidate. The kinetic energy of the π_0 is the sum of the reconstructed energy of the two photons from which the mass of the π_0 is subtracted. Pions from the hadronic system in general have less energy than the pions coming from the true ρ decay so we shall favour ρ candidates with the highest ρ_{E_K}.
- The momenta of the particles of a triplet ($\overrightarrow{p}_{\gamma_1}$, $\overrightarrow{p}_{\gamma_2}$, $\overrightarrow{p}_{\pi^\pm}$) allow us to define a ρ direction as $\overrightarrow{p}_\rho = \overrightarrow{p}_{\gamma_1} + \overrightarrow{p}_{\gamma_2} + \overrightarrow{p}_{\pi^\pm}$. The motivation comes from the observation that in the majority of the cases where a fake ρ was identified as the true ρ via the ranking method based on the invariant mass metric $d_{M_{inv}}$, the winning fake candidate was actually composed of one or two true particle(s). Such hybrid candidates (in the sense that they contain at least one true and one fake particles) tend to have a higher angular dispersion. The third variable is noted $\bar{\theta}$ and is defined as:

$$\cos(\theta_{\gamma_1}) = \frac{\overrightarrow{p}_{\gamma_1} \cdot \overrightarrow{p}_\rho}{||\overrightarrow{p}_{\gamma_1}|| \times ||\overrightarrow{p}_\rho||} \quad (4.17)$$

$$\bar{\theta} = \frac{1}{3} \sum_{i=\gamma_1,\gamma_2,\pi^\pm} \theta_i, \quad (4.18)$$

  where θ_i is the angle between a particle of the triplet and the ρ direction carried by $\overrightarrow{p}_\rho$.

The two latter variables and the invariant mass metric $d_{M_{inv}}$ defined in equation 4.9 are used to define a Medal Game (MG).

For a given ν_τ event where there exists an ambiguity in the ρ reconstruction, the three variables $d_{M_{inv}}$, ρ_{E_K} and $\bar{\theta}$ are calculated for each ρ candidate. They are classified according to each of these variables separately following an ascending order for $d_{M_{inv}}$, a descending order for ρ_{E_K} and an ascending order for $\bar{\theta}$. At this stage one obtains three separate rankings. For each of them, the best three candidates are rewarded respectively with 3, 2 and 1 medals. Then all candidates are ranked following their total number of medals. In case of tied candidates, an internal global score is calculated to separate the candidates.

The results of the improved ranking method based on the Medal Game are given in table 4.3 where the performance of the invariant mass method alone is also recalled. The Medal Game allows for a significant improvement in the correct ρ reconstruction since the rank 1 fraction increased from 21.9% using only the invariant mass method to 29.2% with the Medal Game. In

total the true ρ is reconstructed in $52.6 + 29.2 = 81.8\%$ of the cases. Additional tries with more variables in the Medal Game were not found to help significantly reducing the $\simeq 18\%$ fake ρ contamination.

4.4 Signal selection and background evaluation

The previous section studied the first step of the ν_τ search analysis via the $\tau \to \rho$ decay mode which is to correctly reconstruct the ρ resonance associated to the τ decay. A 82% ρ reconstruction efficiency was achieved. The analysis in now completed with a signal/background study. Remind that the searched signal corresponds to the ν_τCC interactions where the full decay chain of the τ^- is $\tau^- \to \rho^- \nu_\tau \to \pi_0 \pi^- \nu_\tau \to \gamma_1 \gamma_2 \pi^- \nu_\tau$. The absence of electron and muon in the leptonic system makes the searched events quite similar to neutral current interactions containing pions in the final state. They constitute the dominant source of background which will be the only one studied in this chapter. The minimal final state topology of neutral current interactions that can mimic a ρ decay signature is $1\pi_0 1\pi^\pm$. These events will later be noted NC($\geq 1\pi_0 1\pi^\pm$).

Assumptions: here is a list of the assumptions made in the analysis of this chapter.

- The DUNE neutrino beam is assumed to operate in the neutrino (forward horn current) mode.
- The standard LBNF neutrino beam, optimized for the CP violation search, is used by default. The alternative high energy τ optimized neutrino flux is used when explicitly mentioned.
- Neutrino oscillations are calculated with GLoBES using the oscillation parameters of table 2.1. In this chapter, the Normal Ordering hypothesis is always assumed.
- Detector effects are taken into account via the smearing method (see 2.5 for the corresponding discussion) when explicitly mentioned.

The ν_τ charged current events for which the τ decays into a ρ represent 25.52% of the expected total number of ν_τ interactions. Given the event rates of table 2.2, one expects $25.52\% \times 270 \approx 69$ signal events assuming a time of exposure corresponding to the DUNE 3.5 years staged deployment plan discussed in 2.4.2. One also expects a smaller contribution from $\bar{\nu}_\tau$ charged current interactions at the level of 10%. This contribution is neglected but a more complete analysis would need including these events in the signal definition since DUNE far detectors will only partially be able to resolve the matter/antimmater reconstruction of interacting neutrinos and antineutrinos.

As far as the neutral current background are concerned, the contribution from the unoscillated ν_μ, $\bar{\nu}_\mu$, ν_e and $\bar{\nu}_e$ fluxes are all taken into account with the correct relative abundance. Since the neutral current interactions are flavour independent, it is useless to apply neutrino oscillations to the four corresponding neutrino fluxes and then combine the neutral interactions due to all flavours. 8832 neutral current interactions are expected at the DUNE far detectors assuming the 3.5 years staged deployment plan. With the standard LBNF neutrino flux it is found that 19.4% of these NC have the minimal final state topology NC($\geq 1\pi_0 1\pi^\pm$) which constitute the background. Thus the initial signal/background ratio is $69/1713 \simeq 4\%$. It is more favourable than the initial signal/background ratio of the $\tau \to e$ analysis which was at the level of 3%.

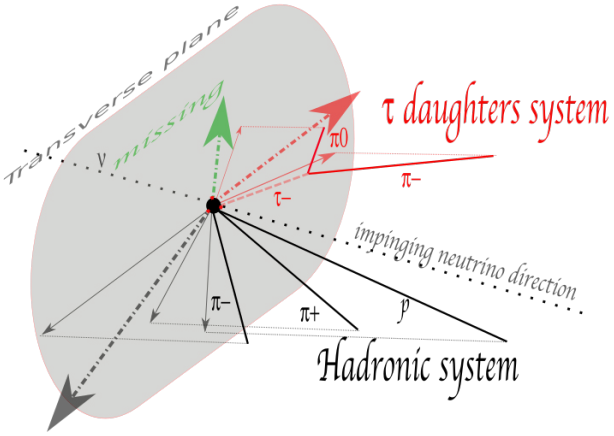


Figure 4.6: Schematic view of a ν_τ deep inelastic charged current interaction. The hadronic system (black) is composed of two charged pions and one proton. The leptonic system (red) is composed of the daughter particles associated to the $\tau^- \to \rho^- \to \pi_0\pi^-$ decay. The neutral pion is not decayed for sake of visual simplicity. The beam neutrino direction is shown and allows to define the transverse plane (gray) of the interaction where the momenta of the particles are projected. A green arrow indicates the transverse missing momentum.

The present analysis makes use of a likelihood approach where first a set of kinematical variables with presumed discrimination power are studied. The corresponding probability density functions are built and discussed. For the NC($\geq 1\pi_0 1\pi^\pm$) the Medal Game is used to select only one ρ candidate per background event and for the signal the MC truth is used to build the distributions with the correct ρ. This allows to avoid biasing the signal distributions with the Medal Game. The obtained distributions are then used to build log-likelihood ratio distributions. For this second step the Medal Game is this time used both for the signal and the background events to reconstruct one ρ per event.

4.4.1 The set of kinematical variables

To help understand the definitions of the kinematical variables defined in this section, a schematic view of a DIS ν_τCC interaction is shown in figure 4.6. The final state is split in two components: the red one (leptonic system) and the black one (hadronic system). The ρ^- of the τ^- decay is not shown. In total the final state contains one reconstructed neutral pion, three reconstructed charged pions and one reconstructed proton. The neutral pion is not decayed for sake of visual simplicity. The impinging neutrino direction is shown as well as the transverse plane of the interaction on which we projected the particles momentum.

The following 17 kinematical variables are defined:

- The kinetic energy of the pions $E^K_{\pi_0}$ and $E^K_{\pi^\pm}$ of the ρ candidate. Their sum $\rho_{E_K} = E^K_{\pi_0} + E^K_{\pi^\pm}$ defines what will be referred to as the ρ energy. We also define $r^K_\pi = E^K_{\pi^\pm}/\rho_{E_K}$ which is representative of the energy sharing between the two pions, using the charged pion as a reference. The kinetic energy of the charged pion is taken as the reconstructed energy, while for the neutral pion it is computed from the sum of the energies of the two photons minus the π^0 rest mass. If the latter operation leads to a negative energy, the neutral pion candidate is discarded.

- The two invariant masses $M_{\pi_0}^{(inv)}$ and $M_\rho^{(inv)}$. They were already defined respectively in equations 4.7 and 4.8.
- The four laboratory angles $\theta_{\rho h}$, $\theta_{\rho tot}$, $\theta_{h\nu}$, $\theta_{\rho\nu}$ which are respectively the angles between the ρ and the hadronic momenta, the ρ and the total momenta of the final state, the hadronic system momentum and the impinging neutrino direction, the ρ momentum and the impinging neutrino direction. The hadronic momentum is defined as the sum of the particles momentum belonging to the hadronic system. The ρ momentum is the sum of the particles momentum belonging to the leptonic system.
- The modulus of the projections of the ρ momentum and the hadronic momentum in the transverse plane $p_\rho^{(tr)}$ and $p_{had}^{(tr)}$. In this plane the two momenta generally do not sum to zero. As in the $\tau \to e$ decay mode analysis the transverse missing momentum is defined to account for this non-zero component:

$$\overrightarrow{p}_{miss}^{(tr)} = -\left(\overrightarrow{p}_\rho^{(tr)} + \overrightarrow{p}_{had}^{(tr)}\right), \tag{4.19}$$

  and the modulus of the transverse missing momentum $p_{miss}^{(tr)}$ is used in the analysis.
- The relative orientation of the three transverse momenta is described with the angles $\phi_{h\rho}^{(tr)}$, $\phi_{hm}^{(tr)}$, $\phi_{\rho m}^{(tr)}$. For instance $\phi_{h\rho}^{(tr)}$ is the angle between the transverse ρ momentum and the transverse hadronic momentum.
- The transverse mass defined as

$$M^{(tr)} = 2\sqrt{p_\rho^{(tr)} p_{miss}^{(tr)}} \left|\sin\left(\frac{\phi_{\rho m}^{(tr)}}{2}\right)\right|. \tag{4.20}$$

It should be noted that most of the previous variables are ρ dependent in the sense that their numerical value depends on the ρ reconstruction. Only the transverse missing momentum is independent of the choice of the ρ candidate.

Six of the kinematical variables distributions which show signal/background discriminating power are plotted in figure 4.7 in blue for the signal and red for the background. It must be noted that though the detector effects are not taken into account in this section, the neutrons are considered undetected. For a signal or background event, if there is no reconstructed hadronic system in addition to the ρ candidate, the event is discarded. This consideration mostly concerns NC($\geq 1\pi_0 1\pi^\pm$) for which the final state topology can be exactly that of the minimal topology required. With the present considerations, 17.9% of the NC events are NC($\geq 1\pi_0 1\pi^\pm$) with a non-null hadronic system to go with the ρ candidate(s) while the fraction of NC($\geq 1\pi_0 1\pi^\pm$) is 19.4%.

The features of the ρ energy (top left), the invariant mass $M_\rho^{(inv)}$ (top right), the transverse missing momentum (middle left) and the transverse angle $\phi_{hm}^{(tr)}$ (middle right) are briefly discussed.

The ρ energy of the signal events is higher than the background events. This is understandable by the fact that the pions of NC($\geq 1\pi_0 1\pi^\pm$) are pions only coming from the hadronic system, which have less energy than the pions of the ρ decay of ν_τ interactions.

The ρ invariant mass distribution of the signal peaks at $0.7\,\mathrm{GeV}/c^2$ while the corresponding background distribution has a large fraction below $0.6\,\mathrm{GeV}/c^2$. $M_\rho^{(inv)}$ will have a key role in the signal/background separation in addition to its role in the Medal Game.

The transverse missing momentum modulus of the signal and the background are comparable which was not the case in the $\tau \to e$ analysis. Indeed, both $\nu_\tau(\tau \to \rho\ \nu_\tau)$ and NC($\geq 1\pi_0 1\pi^\pm$) events have one undetected final state neutrino in the leptonic system which carries away missing energy.

The background distribution of $\phi_{hm}^{(tr)}$ favours the region close to 180° which corresponds to the transverse hadronic momentum and the transverse missing momentum in an antiparallel configuration. In reality the final state neutrino momentum compensates the hadronic momentum but since it is undetected the missing momentum is observed to go in opposite direction with respect to the hadronic system.

In addition the 2-dimensional distributions built with the 17 kinematical variables (offering 136 possibilities) are studied. It must be noted that most of them are irrelevant. We show in figure 4.8 two correlations for which the signal and background favour distinct regions of the plane: the correlation between the transverse missing momentum and the transverse ρ momentum (left) and the correlation between the ρ energy and the laboratory angle $\theta_{\rho h}$ (right, corresponding to the angle between the ρ momentum and the hadronic momentum). The 1-dimensional and 2-dimensional kinematical distributions can now be used to proceed to the likelihood analysis of the signal/background study. They are all smoothed and normalized to unity to be used as probability density functions.

4.4.2 Likelihood search: test hypothesis

The neutrino events studied are the $\nu_\tau(\tau \to \rho)$ (signal) and the NC($\geq 1\pi_0 1\pi^\pm$) (background). For each of them the Medal Game defined in 4.3.5 is used to reconstruct one ρ candidate. For the signal events this reconstruction is now done blindly with respect to the MC truth. The 17 kinematical variables previously defined are calculated as well as the corresponding log-likelihood ratios:

$$\mathscr{L}^{(1d)} = \log\left(\frac{L_S^{(1d)}}{L_B^{(1d)}}\right), \tag{4.21}$$

where L_S is the likelihood for the event of being signal and L_B the likelihood of being background. In addition the 136 log-likelihood ratios associated to the 2-dimensional variables are calculated:

$$\mathscr{L}^{(2d)} = \log\left(\frac{L_S^{(2d)}}{L_B^{(2d)}}\right). \tag{4.22}$$

This brings the total set of likelihood ratios to 153. In the case where L_S or L_B vanishes, the minimal value of the corresponding probability density function is used.

The most discriminating 1-dimensional and 2-dimensional variables are searched manually by looking at the log-likelihood ratio distributions obtained for the signal and background events. Figure 4.9a describes the log-likelihood distributions associated to the correlation of the laboratory angle $\theta_{\rho h}$ and the ρ energy ρ_{E_K} for which we have shown the 2-dimensional probability density functions in the previous section (see figure 4.8 on the right). The blue (red) histogram corresponds to the signal (background), and both histograms are normalized to unity to allow

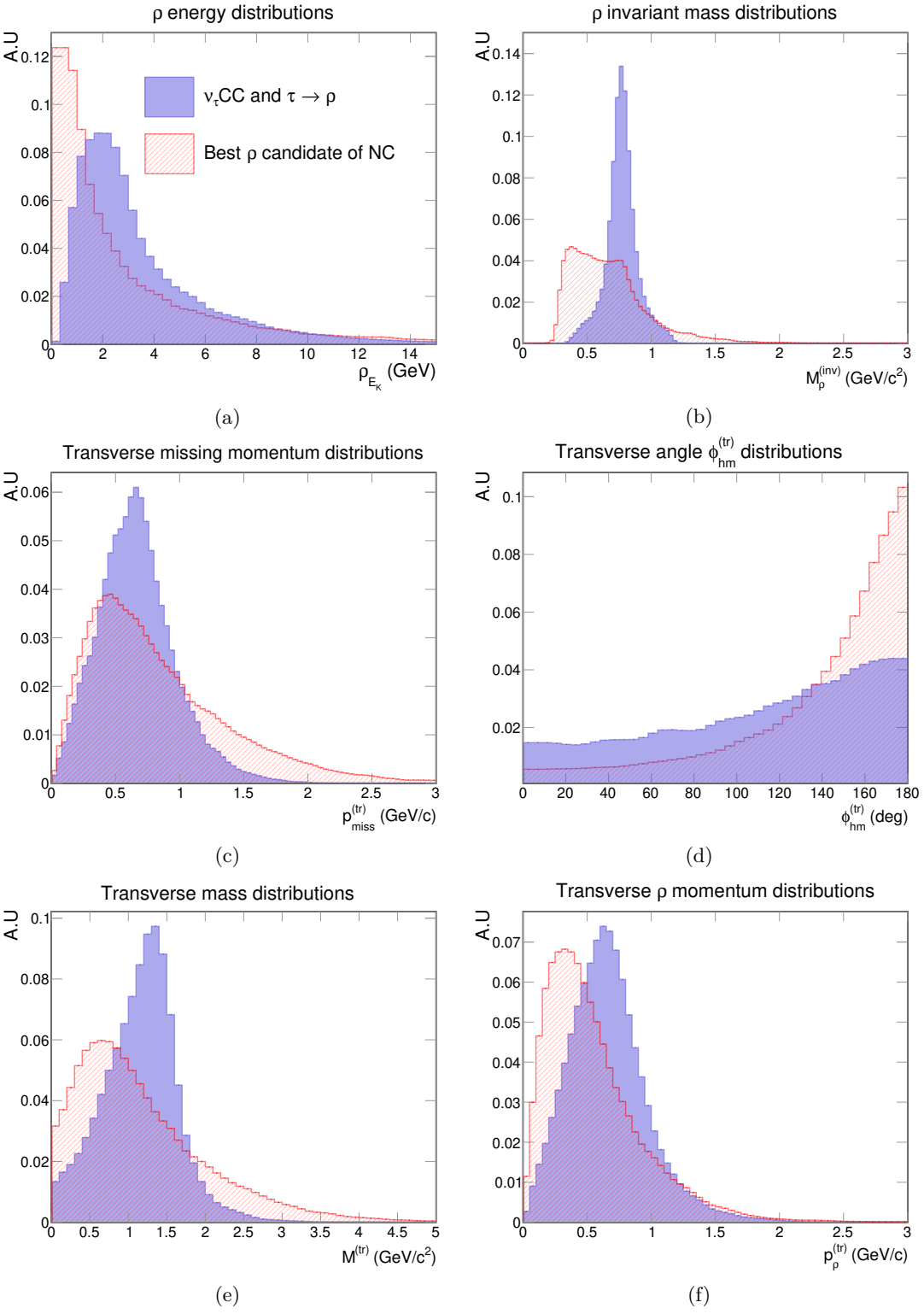


Figure 4.7: Kinematical distributions normalized to unity for 6 of the 17 defined variables for the analysis. The signal (background) is shown in blue (red). For the signal the MC truth is used to select the true ρ whereas for the NC($\geq 1\pi_0 1\pi^\pm$) background the Medal Game is used to select the best ρ candidate (if the ρ multiplicity is greater than 2) in order to have only one contribution per background event.

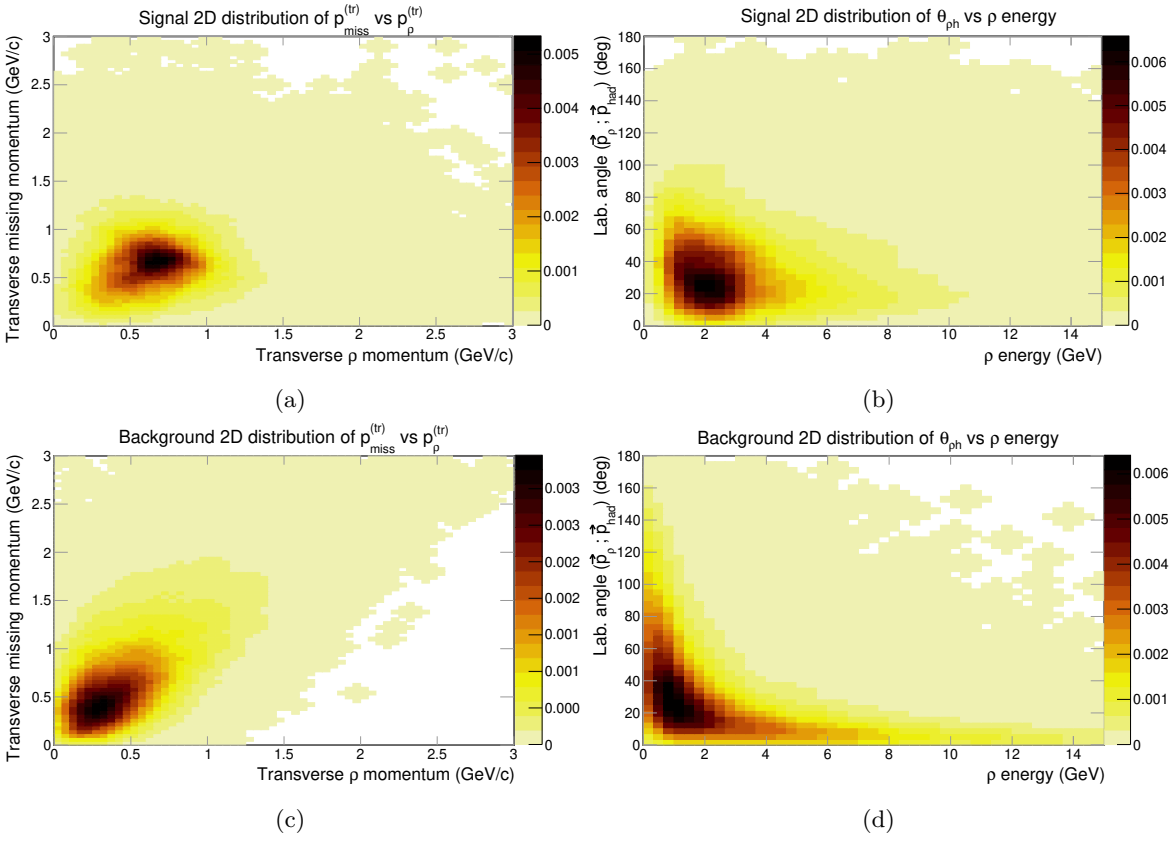


Figure 4.8: 2-dimensional distributions for two sets of kinematical correlations. Left: transverse missing momentum against transverse ρ momentum $[p_{miss}^{(tr)}$ vs $p_{\rho}^{(tr)}]$. Right: laboratory angle between the ρ momentum and the hadronic momentum against the ρ energy $[\theta_{\rho h}$ vs $\rho_{E_K}]$. The top (bottom) distributions correspond to the signal (background).

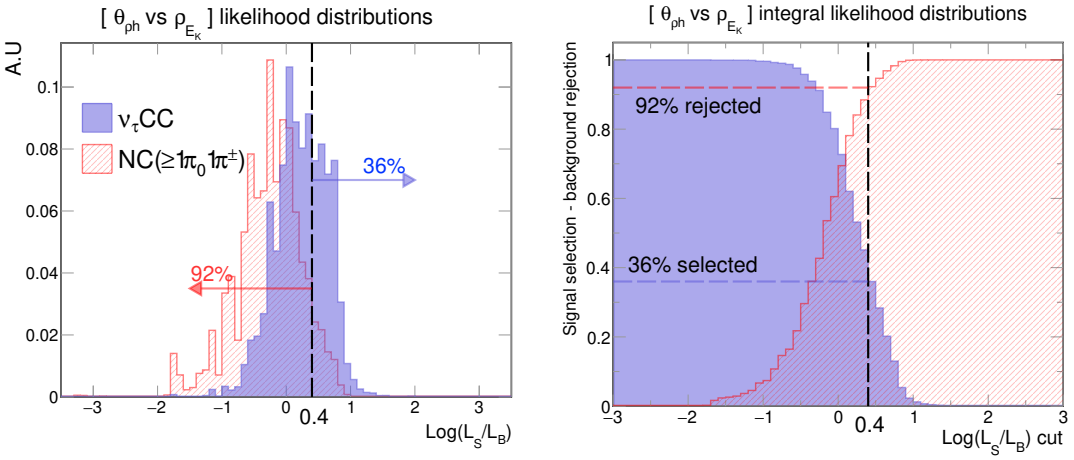


Figure 4.9: Left: log-likelihood ratio distributions of the signal (background) in blue (red) associated to the correlation of the variables $\theta_{\rho h}$ (laboratory angle between the hadronic and the ρ momenta) and ρ_{E_K} (the ρ energy). Right: the integrated distributions of the two. The signal (background) convention is to represent the selection efficiency (rejection efficiency) as a function of the log-likelihood ratio cut value. An illustrative cut at 0.4 is shown. It allows for a 36% selection efficiency and a 92% background rejection efficiency.

for a direct visual comparison. Figure 4.9b shows the corresponding integral likelihood distributions. It conventionally represents the signal selection efficiency and the background rejection efficiency as a function of the log-likelihood ratio cut value.

An illustrative log-likelihood ratio cut placed at 0.4 is shown on the two figures and it allows for a signal selection efficiency of 36% and a background rejection efficiency of 92%. The log-likelihood distributions associated to the correlation of the variables $p_{miss}^{(tr)}$ and $p_\rho^{(tr)}$ for which we also showed the distributions in the previous section leads to similar quantitative results.

This first approach allows to isolate discriminating variables that can be further combined to improve the signal/background separation.

4.4.3 Optimized likelihood search

A possible way of improving the analysis is to combine several likelihoods. Given a and b two kinematical variables (whether 1-dimensional or 2-dimensional), they can be combined within a unique log-likelihood ratio of the form:

$$\mathscr{L}_{Opt} = \log\left(\frac{L_S^{(a)} \times L_S^{(b)}}{L_B^{(a)} \times L_B^{(b)}},\right). \tag{4.23}$$

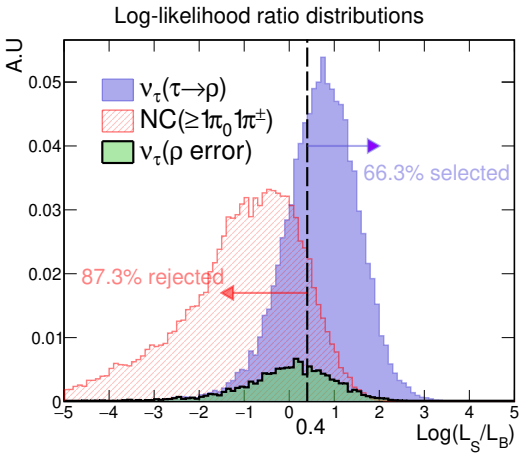
where the subscript *Opt* stands for Optimized. Two remarks must be made about this definition.

- First equation 4.23 is mathematically correct if a and b are independent variables. Since they describe the same physical process it is probably not the case here. The term "pseudo-likelihood" should then rather be employed. This terminology distinction will further be omitted, and one can think of the proposed method as a sophisticated way of applying cuts to kinematical distributions.
- Secondly redundant information should be avoided. For instance if the transverse missing momentum information is in variable a then it should not be in variable b. Such repetitions would not improve the analysis but only widen the gap between very signal-like and background-like events.

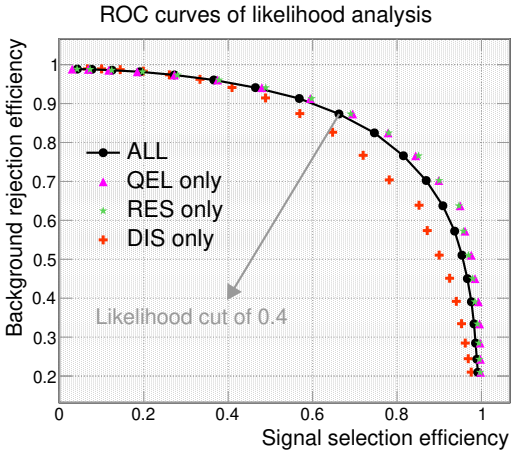
The combination of the 2-dimensional variable $[p_{miss}^{(tr)}\,;\,p_\rho^{(tr)}]$ with the 2-dimensional variable $[\theta_{\rho h}\,;\,\rho_{E_K}]$ and the ρ invariant mass $M_\rho^{(inv)}$ is found to optimize the signal/background separation. The results associated are reported in figure 4.10.

First, the log-likelihood ratio distributions of the signal (background) is shown in figure 4.10a with the blue (red) histogram. Both histograms are normalized to unity. Mind the smaller green histogram at the bottom which corresponds to the fraction of ν_τCC signal for which the ρ was misreconstructed by the Medal Game ($\simeq$ 18% contamination). It populates the left tail of the blue histogram but not the right tail, indicating that the likelihood tends to reject this contaminated signal population. An illustrative cut at 0.4 allows for a 66.3% signal selection efficiency and 87.3% background rejection efficiency.

Figure 4.10b displays the background rejection efficiency against the signal selection efficiency (ROC curve) for cuts varying between -2 and 2. The point corresponding to the cut value of 0.4



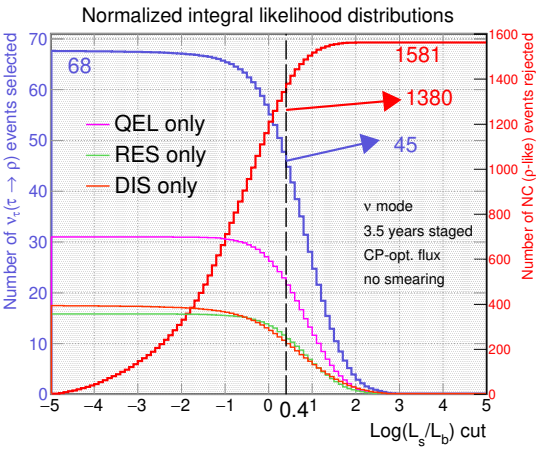
(a)



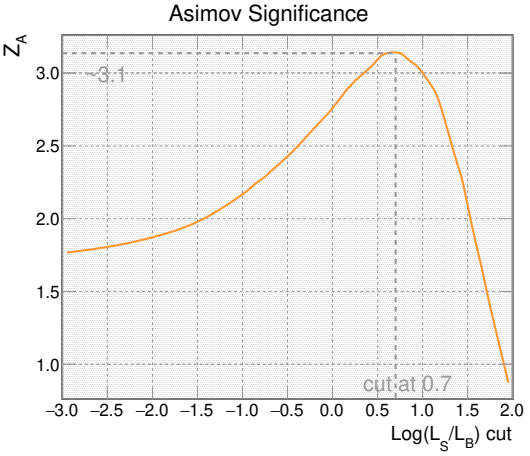
(b)

Likelihood cut	0	0.2	0.4	0.6	0.8	1.0	1.2
fraction of signal kept (%)	81.5	74.7	66.3	56.7	46.5	36.6	27.3
Cases of fake ρ propagated	6.8	5.6	4.6	3.6	2.6	1.9	1.3
fraction of bck contamination	23.4	17.5	12.7	8.7	5.9	3.9	2.6

(c)



(d)



(e)

Figure 4.10: Results of the signal-background study obtained with the optimal likelihood search without the detector effects. See text for a detailed discussion.

is indicated. A distinction is made for the three main scattering processes (quasi-elastic QEL, resonant RES and deep inelastic DIS). The message is that the likelihood performs significantly worse on the DIS sample than on the QEL and RES for which it performs similarly.

Table 4.10c gives some quantitative signal efficiencies and background contaminations associated to the log-likelihood distributions. A raw is dedicated to the case where the ρ of the ν_τ was misreconstructed by the Medal Game. As for illustration, a cut at 0.4 selects 66.3% of the signal among which 4.6% correspond to misreconstructed ρ. Thus, the fraction of signal with a correct ρ reconstruction is 61.7%. A 12.7% background contamination is conceded.

The previous efficiencies are normalized following the 3.5 years staged deployment plan of DUNE which was discussed in 2.4.2. Table 2.2 indicates the expected number of events which are 270 ν_τCC interactions and 8832 NC interactions. We must take into account the branching ratio of the $\tau \to \rho$ decay mode which is 25.52% and the fraction of NC($\geq 1\pi_0 1\pi^\pm$) which was found to be 19.4%. Additional events are discarded if the final state has either no ρ candidate available or if there is no reconstructed hadronic system in addition to the ρ candidate(s). This mostly concerns the background population for which 17.9% of the NC satisfy both the minimal topology NC($\geq 1\pi_0 1\pi^\pm$) and the latter condition. These criteria lead to an initial S/B (signal to background) ratio of 68/1581. The cut at 0.4 selects 45 signal events and rejects 1380 background events (contamination of 201). The signal QEL/RES/DIS contributions are also shown.

Finally the figure of merit of this analysis is shown on 4.10e. It represents the Asimov significance as a function of the log-likelihood cut value. Remind that given an observed signal s and a background b it is defined as:

$$Z_A = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)}. \tag{4.24}$$

The significance shows a maximum at 3.1σ obtained for a cut value of 0.7. It corresponds to a S/B ratio of 35/114.

A further step is required to better estimate the DUNE sensitivity to the $\nu_\mu \to \nu_\tau$ appearance via the study of the $\tau \to \rho$ decay mode. Indeed the results exposed in this section do not yet include the detector effects which are simulated in this thesis with the smearing process described in 2.5.

4.5 Including smearing effects

The smearing designates a process to simulate in a fast manner the detector response at the single particle level. Our reference smearing method takes into account both the energy and direction reconstruction resolutions of the DUNE far detectors. This section proposes to review the results obtained in the previous one in the light of the smearing effects.

4.5.1 Impact of detector effects on the kinematical distributions

The kinematical distributions shown in figure 4.7 are now reported in figure 4.11 to illustrate the smearing effects. The signal (background) is shown with the blue (red) color. The dashed (filled) histograms correspond to the smeared (unsmeared) distributions.

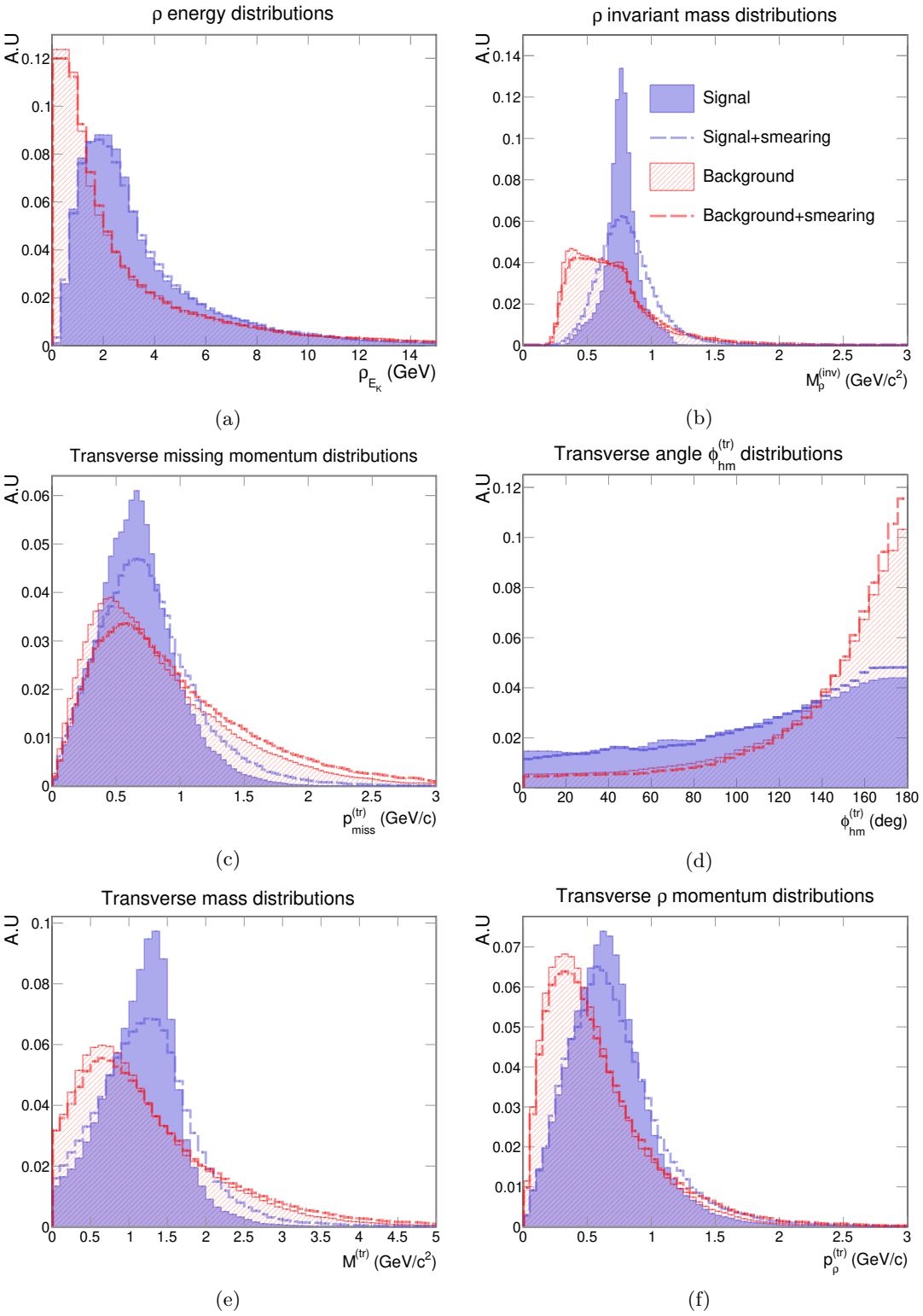


Figure 4.11: Kinematical distributions normalized to unity of 6 of the 17 defined variables for the analysis including the smearing effects described in 2.5. The signal (background) is shown in blue (red). The distributions obtained with (without) the smearing correspond to the dashed (filled) histograms.

The main effect is observed on the signal $M_\rho^{(inv)}$ distribution for which the peak at $0.77\,\mathrm{GeV}/c^2$ was significantly widened. The transverse missing momentum distributions of both the signal and background were slightly shifted to higher values indicating a transverse missing momentum component attributed to the misreconstruction of particles momenta in the transverse plane. As a consequence the transverse mass distributions are also affected since $M^{(tr)}$ depends on the transverse missing momentum. Other kinematical distributions were not much affected.

4.5.2 Impact of detector effects on the likelihood analysis performance

The likelihood results evaluated including smearing effects are reported in figure 4.10. The same optimized combination of likelihoods as in 4.4.3 is used.

Figure 4.10a gives the log-likelihood ratio distributions of signal (background) in blue (red). In addition the green histogram illustrates the fraction of signal events for which the Medal Game misreconstructed the ρ. In comparison to the no smearing case, the signal and background distributions are found to some extent to have pulled together. For instance the fraction of background with a likelihood ratio below -3 was drastically suppressed. This is a direct consequence of the signal variable $M_\rho^{(inv)}$ for which the peak at $0.77\,\mathrm{GeV}/c^2$ (see figure 4.11b) was found to widen because of the smearing effects. The illustrative cut at 0.4 is repeated and allows for a 56.8% (66.3% without smearing) signal efficiency and 88.4‰ (87.3% without smearing) background rejection. The notable effect of the smearing is the decrease of about 10% in the signal selection efficiency. Table 4.10c extends the illustrative cut at 0.4 to several log-likelihood cut values, indicating for each the signal selection and background rejection efficiencies as well as the fraction of signal selected for which the ρ was misreconstructed by the Medal Game.

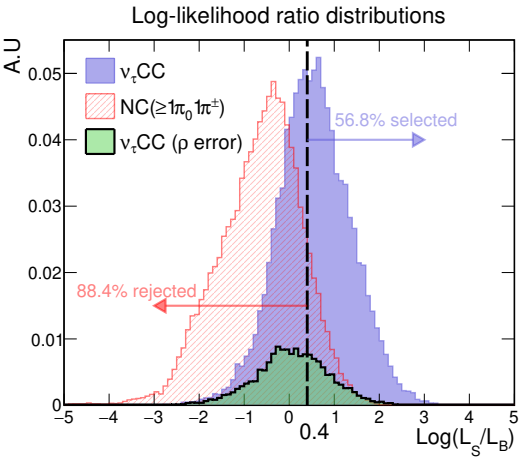
Figure 4.10b shows the background rejection efficiency against the signal selection efficiency for cuts varying between -2 and 2. An arrow indicates the point corresponding to the 0.4 illustrative cut of figure 4.10a. The ROC curve splits the contributions of the signal QEL, RES and DIS interactions and reveals that the likelihood performs significantly less well on the DIS.

Figure 4.10d shows the integral likelihood distributions corresponding to the 3.5 years staged deployment plan. It is found that 18.0% of the NC events have the minimal topology NC($\geq 1\pi_0 1\pi^\pm$) and a non-null hadronic system in addition to the ρ candidate(s). The initial signal/background ratio is 67/1590 and is slightly different than the initial ratio observed without the smearing effects. The QEL, RES and DIS contributions of the signal are also shown. The illustrative cut at 0.4 would select 38 signal events and reject 1405 background events (contamination of 185).

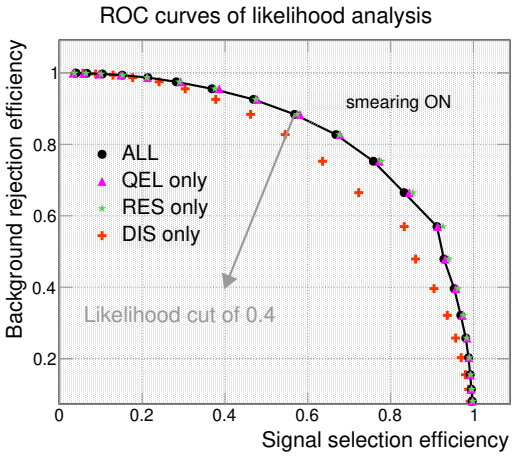
Figure 4.10e reports the corresponding Asimov significance as a function of the log-likelihood ratio cut. In comparison with the study without the smearing effects, the maximum observed for a cut value at 0.7 was decreased from 3.1σ to 2.8σ. It now corresponds to a S/B ratio of 28/91.

4.5.3 Convolutional Visual Network (CVN) events selection bias

In the $\tau \to e$ decay mode analysis the CVN deployed by the DUNE collaboration has been suggested as a pre-selection tool. The CVN was presented in 2.7. The idea was to explore the combination of the CVN with the likelihood analysis in order to search for ν_τ events. The role of the CVN is to classify the neutrino interactions as ν_e, ν_μ, ν_τ and NC. Since the DUNE



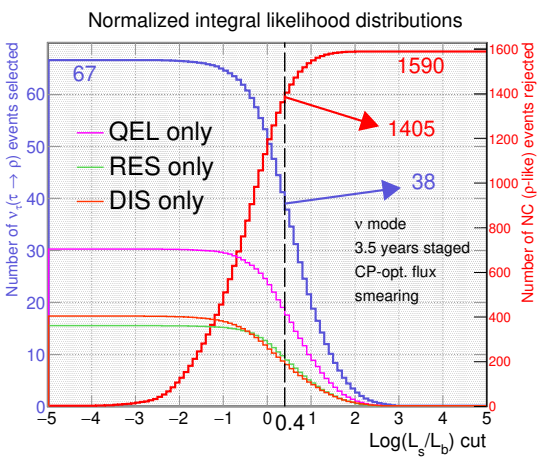
(a)



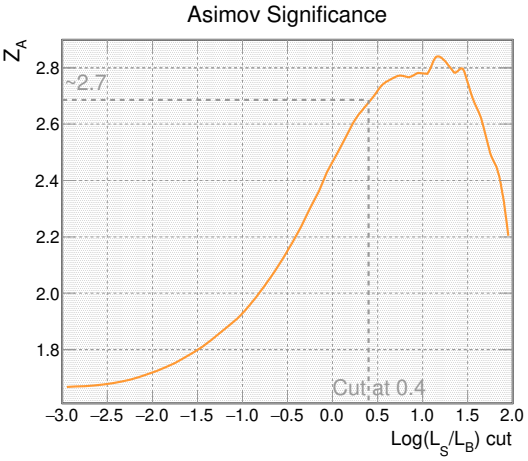
(b)

Likelihood cut	0	0.2	0.4	0.6	0.8	1.0	1.2
fraction of signal kept (%)	75.8	66.8	56.8	46.9	36.9	28.2	21.4
Cases of fake ρ propagated	8.8	7.1	5.6	4.2	2.9	2.0	1.3
fraction of bck contamination	24.8	17.3	11.6	7.4	4.5	2.5	1.3

(c)

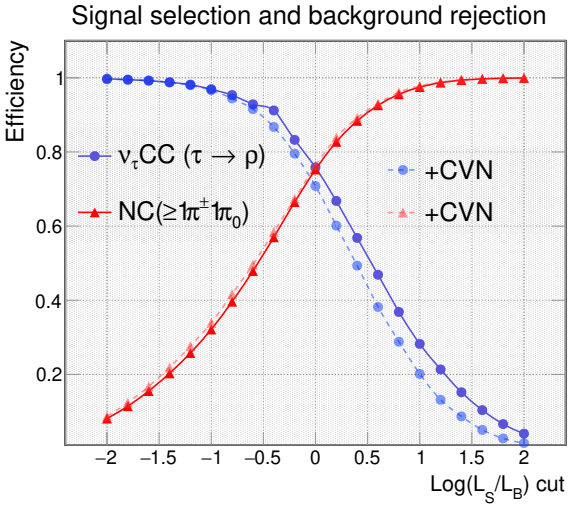


(d)



(e)

Figure 4.12: Results of the signal/background study obtained with the optimal likelihood search including the detector effects. See text for a detailed discussion.



Cut	S % (CVN)	B % (CVN)
0.0	76 (71)	75 (76)
0.2	67 (60)	83 (84)
0.4	57 (49)	88 (89)
0.6	47 (38)	93 (93)
0.8	37 (29)	96 (96)
1.0	28 (20)	98 (98)
1.2	21 (13)	99 (99)

Figure 4.13: Effect of the CVN pre-selection (NC score ≥ 0.5) on the likelihood efficiencies. The signal selection efficiency (background rejection efficiency) is shown in blue with circle markers (red with triangle markers). Full bright curves (dashed filmy curves) correspond to the case without (with) the CVN pre-selection. The right table reports the corresponding signal selection (S) and background rejection (B) efficiencies in percentage. The efficiencies obtained with the CVN bias are indicated between brackets.

collaboration mostly tuned it for the $\nu_\mu \to \nu_e$ and $\nu_\mu \to \nu_\mu$ oscillation studies, the performance of the CVN on the ν_τ classification was rather low (at the level of 15%). The prompt decay of the charged lepton τ does not allow the CVN to rely on a visual hint to recognize the ν_τ interactions.

Table 2.7 reveals that 53.5% of the $\nu_\tau(\tau \to \rho)$ events are classified as NC while only 21.2% are correctly classified as ν_τ. The idea is to explore the fraction of ν_τ interactions misclassfied as NC that can be retrieved thanks to the likelihood analysis. The previous analysis is run again imposing this time that the signal and background events are classified by the CVN as neutral currents, using a score of 0.5. We find that 70.2% of the neutral current backgrounds (*i.e* the NC($\geq 1\pi_0 1\pi^\pm$) for which there is a non-null hadronic system in addition to the ρ candidate(s)) are classified as NC by the CVN. Other NC events misclassified by the CVN are discarded from the analysis.

Figure 4.13 describes the effect of this CVN bias on the likelihood signal selection (blue) and background rejection (red) efficiencies. A filmy (bright) color is used to indicate the case where the CVN pre-selection is (is not) applied. We observe that the CVN bias has little impact on the background rejection, which means that the NC($\geq 1\pi_0 1\pi^\pm$) events correctly identified by the CVN are uncorrelated with the NC events correctly identified by the likelihood. However the signal efficiency is decreased at the level of 10% in most of the log-likelihood ratio axis. It indicates that the $\nu_\tau(\tau \to \rho)$ events misclassified as NC interactions by the CVN are a bit harder to retrieve with the likelihood analysis.

The efficiencies are normalized to the DUNE 3.5 years deployment plan. Without the CVN bias the initial S/B ratio is 67/1590. The CVN pre-selection leads to a less favourable initial S/B ratio of 36/1116. A maximal sensitivity is reached for a cut value of 0.2 corresponding to a S/B ratio of 21/180 and an Asimov significance of 1.6σ which is to compare to the 2.8σ obtained without the CVN pre-selection. This sharp decrease can be explained by the fact that the misclassification of $\nu_\tau(\tau \to \rho)$ events as NC is lower than the correct classification of

Fake ρ multiplicity	0	1	2	3	4	5	>6
No $\pi^{\pm}$ identification (%)	54	18	4	2	1	7	15
$\pi^{\pm}$ identification (%)	71	5	<1	<1	<1	12	10

(a)

Medal Game ranking	-1	0	1	2	3	>3
No $\pi^{\pm}$ identification (%)	2.9	52.6	29.2	6.7	1.9	6.7
$\pi^{\pm}$ identification (%)	2.9	69.6	17.1	3.9	1.6	6.6

(b)

Table 4.4: (a) shows the effect of the charged pion identification assumption on the fake ρ multiplicity per ν_τCC($\tau \to \rho$) event and (b) the evolution of the correct ρ reconstruction efficiency as performed by the Medal Game with and without the charged pion identification.

NC($\geq 1\pi_0 1\pi^{\pm}$). In addition the signal efficiency of the likelihood has a 10% decrease when the CVN is used as a pre-classifier tool.

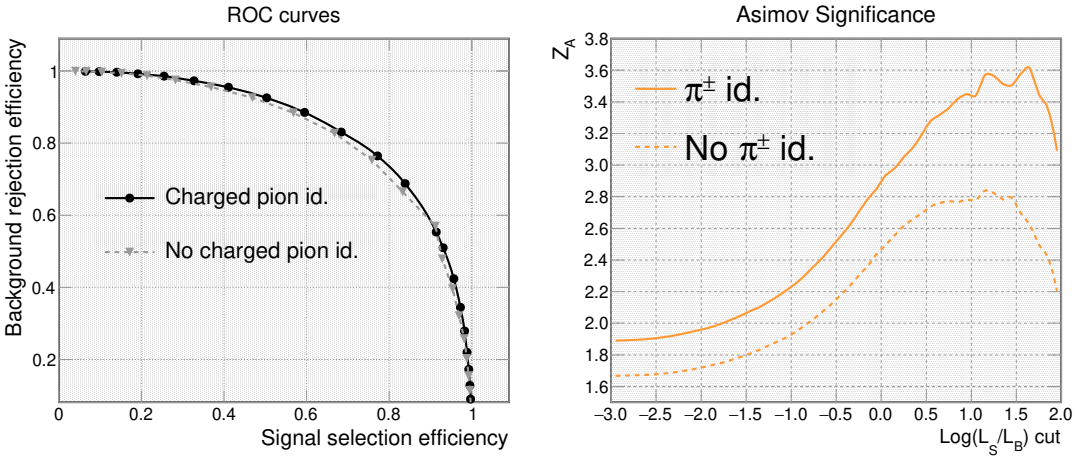
4.6 Charged pion identification

Recent publications by the DUNE collaboration about the protoDUNE single phase operating at CERN suggest the capability of the LArTPC detectors to perform particle identification via measurement of the mean energy deposited per unit of length $\langle dE/dx \rangle$ combined with the range [68]. For a given kinetic energy, protons typically leave a smaller track with a higher ionizing signal per unit of length than charged pions. However this method does not allow for the π^+/π^- differentiation. Deeper understandings of the pions interactions in liquid argon is required, and has been for instance studied in [105]. A difference observed between π^+ and π^- is that π^- can undergo a nuclear capture at rest while π^+ would rather decay at rest into a muon and an muon antineutrino. The resulting different topologies can to some extent be used to differentiate between the positively and negatively charged pions.

This section suggests to evaluate the effects of an idealistic full differentiation between the charged pions π^+/π^-. This assumption will obviously affect the ρ reconstruction of ν_τ events since the fake ρ multiplicity per event is expected to decrease as a result of the positively charged pions no longer used to define fake ρ candidates. In this section a fake ρ candidate can only be composed of two photons and a π^- since the signal considered in the study are the $\nu_\tau(\tau^- \to \rho^- \to \pi_0\pi^-)$.

Table 4.4a indeed gives the fake ρ multiplicity per $\nu_\tau(\tau \to \rho)$ event with and without the charged pion identification assumption. In general the fake ρ multiplicity is decreased. More specifically the 0 multiplicity increased from 54% to 71%. It should be noted however that the ρ multiplicity of 1 decreased from 18% to 5%. It is found that this mainly corresponds to events for which the hadronic system contains one π^+ and no photon. With the charged pion identification, such a topology populates the multiplicity 0 whereas without the charged pion identification it populates the multiplicity 1.

The ρ reconstruction efficiency using the Medal Game (defined in 4.3.5) is given in table 4.4b. The message is that the charge reconstruction of the pions allows for a moderate increase of the correct ρ reconstruction which goes from 81.8% to $69.6 + 17.1 = 86.7\%$.



(a) (b)

Figure 4.14: (a) shows the ROC curve (background rejection efficiency against signal selection efficiency) corresponding to the optimal likelihood analysis assuming (not assuming) the charged pion identification in black with round markers (dashed gray with triangle markers). (b) gives the achievable increase of sensitivity to the ν_τ appearance corresponding to a 3.5 years staged normalization. The full (dashed) curve is obtained with (without) the charged pion identification.

The second aspect is to re-evaluate the signal/background study in the light of the charged pion identification. The minimal final state topology of neutral current interactions to be considered as a background is now NC($\geq 1\pi_0 1\pi^-$). In addition if there is only one valid reconstructed ρ candidate, the final state must provide a non-null hadronic system to go with the ρ. It is observed that the corresponding NC fraction is 13.7% to compare with the 18.0% without the charged pion identification. It must be noted that this background reduction is not an effect of the present analysis but the best achievable with an idealistic π^+/π^- differentiation.

Figure 4.14a shows the ROC curves (background rejection efficiency against signal selection efficiency) obtained with the previous optimal likelihood search with and without the charged pion identification. Only a limited improvement of signal/background separation is achieved. However the overall NC background reduction allows the 3.5 years staged Asimov significance to reach a maximum of about 3.5σ against 2.8σ without the charged pion identification (see figure 4.14b).

4.7 τ optimized flux

An alternative neutrino beam tuned to enhance the ν_τ detection could be considered after several years of operation for DUNE, once the main scientific program (CP violation study) is achieved. This high energy flux was discussed in 2.2.2 and would indeed allow to multiply the ν_τCC event rate by a factor of 6 in comparison to the standard LBNF neutrino flux (see table 2.3). This section suggests to revisit the results of the optimal likelihood analysis (reported in 4.5) in regards of this alternative beam configuration. It proceeds in two steps: the assessment of the ρ reconstruction efficiency in the ν_τ events and the signal/background analysis. In this section we no longer assume the charged pion identification. The detector effects are modeled by the reference smearing method.

Fake ρ multiplicity	0	1	2	3	4	5	>6
CP optimized flux (%)	54	18	4	2	1	7	15
τ optimized flux (%)	43	25	5	2	1	10	15

(a)

Fake ρ multiplicity	0	1	[2;5]	[6;50]	>50
CP optimized flux					
QEL (46%)	93	4	2	<1	<0.1
RES(23%)	21	47	25	8	< 0.1
DIS(26%)	6	18	26	41	9
τ optimized flux					
QEL (32%)	94	4	2	<1	< 1
RES(36%)	18	48	25	9	< 1
DIS(27%)	7	23	28	37	5

(b)

Table 4.5: (a) gives the fake ρ multiplicity for the CP optimized and the alternative τ optimized neutrino beams. (b) is obtained by splitting the chart (a) with respect to the QEL/RES/DIS scatterings for each neutrino flux. The relative abundance of each scattering is indicated in brackets.

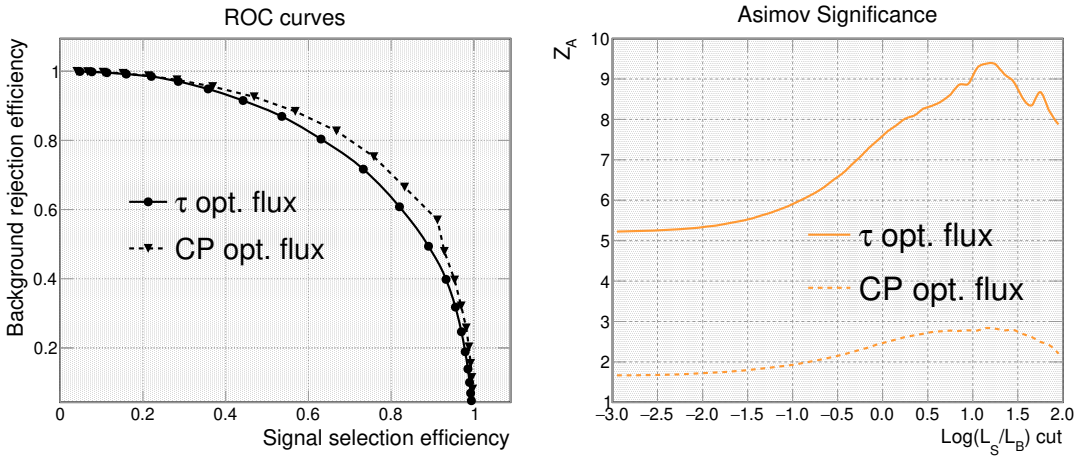
Medal Game ranking	-1	0	1	2	3	>3
CP optimized flux - fraction (%)	2.9	52.6	29.2	6.7	1.9	6.7
τ optimized flux - fraction (%)	2.6	42.8	37.4	8.5	2.3	6.5

Table 4.6: Comparison of the ρ reconstruction efficiency for both the CP optimized neutrino flux and the alternative higher energy τ optimized neutrino flux.

To begin with, table 4.5a shows the fake ρ multiplicity in the ν_τ events with the two beam configurations. The level of ambiguity in the ρ reconstruction is higher for the high energy τ optimized flux. The 0 multiplicity decreased from 54% to 43%. Two other changes are noted: the increase from 18% to 25% of the multiplicity 1 which corresponds to an hadronic system topology of the form $0\gamma 1\pi^{\pm}$, and the multiplicity 5 which increases from 7% to 10%. The latter corresponds to an hadronic system with a pion topology of the form $1\pi_0 0\pi^{\pm}$.

Table 4.5b allows to understand the underlying mechanism associated to this rise in the ρ reconstruction level of ambiguity. The fake ρ multiplicities of each QEL/RES/DIS scatterings in the two beam configurations are comparable. However the relative abundance of QEL and RES is modified with the τ optimized flux: the QEL proportion decreases from 46% to 32% and the RES increases from 23% to 36%. It should be noted that the DIS relative abundance is however unchanged. Since the RES scatterings produce more final state pions than the QEL, the ρ level of ambiguity for the high energy τ optimized neutrino flux is higher than for the standard LBNF neutrino flux.

The performance of the Medal Game for the reconstruction of the correct ρ is now assessed and results are shown in table 4.6. The previous results obtained with the standard CP optimized neutrino beam are also reported for comparison. By chance, the sum of the ranks 0 and 1 for the two fluxes are comparable. One obtains a correct ρ reconstruction efficiency of 81.8% with the standard beam and 80.2% with the alternative beam.



(a) (b)

Figure 4.15: Evolution of the figures of merit of the optimized likelihood search (4.5) with respect to the neutrino beam used. (a) illustrates the ROC curves showing the background rejection efficiency *vs.* the signal selection efficiency and (b) shows the significance assuming a 3.5 years staged running plan.

In a second time the signal/background study is re-evaluated with the alternative beam neutrino. Because the neutrino events will on average occur at higher energy than with the CP optimized neutrino flux, it is not clear *a priori* if the optimal likelihood combination used in 4.4.3 will hold as the optimal choice in this section. As reminder, the likelihood made use of the 2-dimensional variable $[p_{miss}^{(tr)}\,;\,p_\rho^{(tr)}]$ combined with $[\theta_{\rho h}\,;\,\rho_{E_K}]$ and $M_\rho^{(inv)}$. The likelihood analysis is run again and alternative combinations are tested. However no other likelihood combination is found to optimize the signal/background separation though other combinations led to comparable results. Thus the evolution of the ROC curve and the Asimov significance is solely reported in figure 4.15. The ROC curves illustrate that the likelihood performance is less good with the τ optimized flux than it is with the CP optimized flux.

The normalization of these efficiencies is applied assuming the 3.5 years staged deployment plan of DUNE. The corresponding event rates for the two neutrino beams are taken from table 2.3. One expects 1658 ν_τ events with the τ optimized neutrino beam. One must take into account the 25.52% branching ratio of the $\tau \to \rho$ decay mode. 17 564 NC events are expected with the alternative beam and 32.8% fill the condition to have the minimal final state topology NC($\geq 1\pi_0 1\pi^\pm$) and have a non-null hadronic system to go with the ρ candidates. The latter condition also removes few percents of signal events. The initial signal to background ratio found is 407/5761. Figure 4.15b reveals that though the likelihood discriminates less well the signal from the background with the τ optimized flux, it is largely compensated by the boost of signal statistics. Indeed the Asimov significance reaches a maximal value of 9σ while for the CP optimized beam the maximum significance is 3σ. A cut at 1.0 allows to reach both maxima. The corresponding S/B ratio is 116/123 (28/91) for the τ optimized (CP optimized) neutrino beam.

4.8 Performance of an Artificial Neural Network (NN)

The current wide literature on artificial neural networks and their use in particle physics makes it worth to investigate their use in the present ν_τ search analysis. So far the analysis has been

divided in two parts:

1. Developing a method to correctly reconstruct the ρ associated to the τ decay in the ν_τ events.
2. Developing a likelihood based analysis and assess its signal/background discriminating power, where the backgrounds are the neutral current interactions which can mimic a ρ decay (NC($\geq 1\pi_0 1\pi^\pm$)).

This section shall also follow these two steps. **TensorFlow** is an open source platform which was chosen to build and train simple neural networks on the two previous tasks mentioned. More specifically the NN are built with the library **Keras** which are the high level deep learning Application Program Interface (API) of the TensorFlow platform. The standard DUNE neutrino flux is used and the detector effects are modeled by the smearing method.

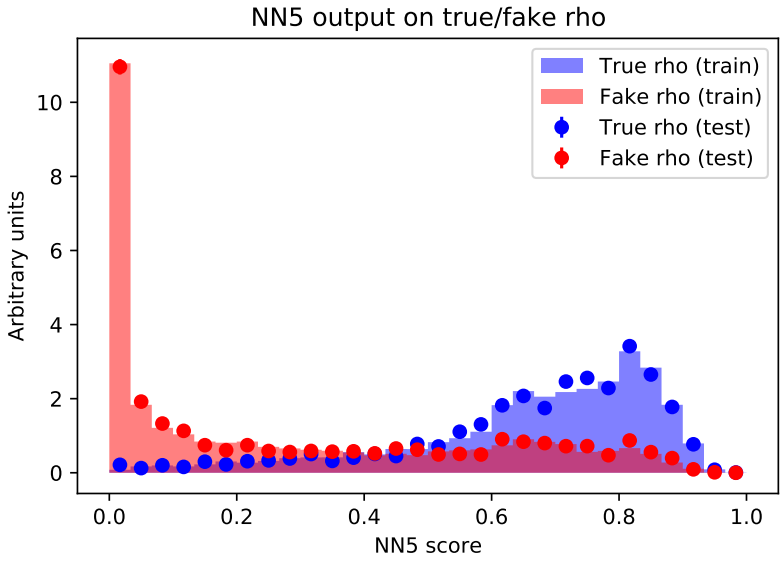
4.8.1 ρ reconstruction with NN

The Neural Network (NN) used in this part will compete with the previously defined Medal Game (MG) in 4.3.5 for which three variables were used: the invariant mass metric $d_{M_{inv}}$ defined in equation 4.9 rewarding ρ candidates with invariant masses close to the masses of the π_0 ($0.135\,\mathrm{GeV}/c^2$) and the ρ ($0.776\,\mathrm{GeV}/c^2$), the ρ energy defined as the sum of the kinetic energies of the pions, rewarding higher energy candidates, and finally an angle score to reward less scattered candidates around the mean ρ direction. A neural network is trained at identifying the correct ρ based on these three variables and is referred to as "NN_3".

Since neural networks are known for finding correlations in complex situations, two other NN are trained with additional variables. NN_5 is trained on the same variables than NN_3 adding the two invariant masses $M_{\pi_0}^{(inv)}$ and $M_\rho^{(inv)}$, totaling 5 kinematical variables. Finally NN_7 is trained adding to the two transverse angles $\phi_{\rho h}^{(tr)}$ and $\phi_{hm}^{(tr)}$, respectively the angle between the transverse ρ and hadronic momenta, and the angle between the transverse hadronic and missing momenta.

The NN are trained using a sample of $\simeq 40\,000$ simulated ν_τCC($\tau \to \rho$) interactions. For each of them the true ρ kinematical variables are calculated. This set constitute the signal target. Some 700 000 fake ρ candidates are available in these simulated interactions. 40 000 of them are randomly selected to equalize the sample sizes. The latter sample constitutes the background target. The two sets are randomly split into one training set of size 3/4 and one testing set of size 1/4. The three neural networks learn with the training sets to differentiate the true ρ from the fake ρ, and the testing set is used to check for overtraining. The NN tries to assign a rank of 1 (0) to signal (background) ρ candidates. 50 epochs and three hidden layers are used for each NN. The results of the training corresponding to NN_5 is shown in figure 4.16a. Full histograms show the training samples while the dotted ones show the testing samples. No indication of overtraining is found. All histograms are arbitrary scaled the same way.

Once the neural networks are trained the corresponding efficiencies are evaluated. For each ν_τCC event where there is an ambiguity in the ρ reconstruction, all the ρ candidates are ranked in descending order based on their NN score. The MC truth is used to find the rank of the true ρ candidate. Remind that rank 0 designates ν_τ events for which the fake ρ multiplicity is null and rank -1 the cases where the true ρ can not be reconstructed. Results are reported in



(a)

True ρ rank	-1	0	1	2	3	>3
Medal Game (%)	2.9	52.6	29.2	6.7	1.9	6.7
NN_3 (%)	2.9	52.6	26.7	8.7	2.7	6.5
NN_5 (%)	2.9	52.6	27.3	8.5	2.7	6.2
NN_7 (%)	2.9	52.6	28.2	8.6	2.6	5.3

(b)

Figure 4.16: Top: raw output of the neural network NN_5. Training (testing) samples correspond to full (dotted histograms). The signal (background) are associated to the blue (red) histograms. Bottom: true ρ reconstruction efficiency comparison between the Medal Game (MG) and the three trained neural networks.

table 4.16b. The method displaying the greatest rank 1 population is the Medal Game. NN_3 performs less well than the Medal Game though it uses the same variables.

As a conclusion, the use of a artificial neural networks to reconstruct the ρ of ν_τCC interactions did not bring substantial improvements. The Medal Game is even found to perform better than the three trained neural networks. The use of additional kinematical variables to train the neural networks did not prove to have an impact. No significant change was either found by changing the structure of the neural networks and the number of epochs used for training.

4.8.2 Signal/background discrimination with NN

The second step is to perform a signal/background separation analysis between ν_τCC ($\tau \to \rho$) events and NC($\geq 1\pi_0 1\pi^\pm$) events. The $\simeq$ 40 000 signal events used in the previous section and the sample of simulated NC($\geq 1\pi_0 1\pi^\pm$) which by chance has approximately the same size define the pool of events for the neural network. For each signal and background event, the 17 kinematical variables defined in 4.4.1 are calculated. For the signal events the MC truth is used to pick the correct ρ candidate while for the background the NN_5 defined in the previous section is used to select one ρ candidate per NC event. This methods allows to avoid training the neural network on a biased signal sample.

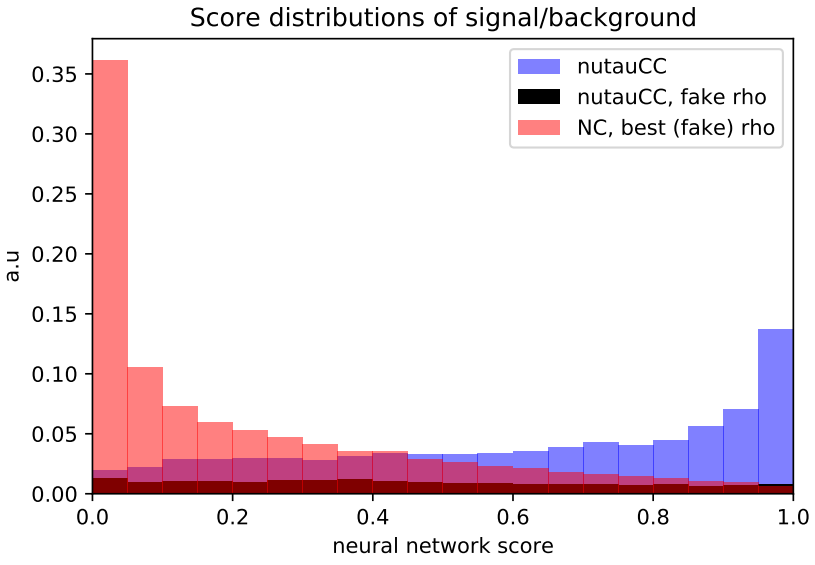
A neural network similar to the ones of the previous section is built (using three hidden layers). It is trained on 3/4 of the signal and background sets. No overtraining hint was found by comparing the training and testing sets. Once trained, for each signal and background event a ρ candidate is selected based on the NN_5 ranking and the event is assigned a score by the just trained neural network. Signal (background) events are targeted with a score of 1 (0).

The score distributions normalized to unity of the signal and background samples are shown in figure 4.17a. A third black histogram is shown at the bottom and it corresponds to signal events for which the NN_5 misreconstructed the ρ in ν_τ events. This signal component is slightly more populating the low neural network score, indicating the neural networks tends to reject them as background. Figure 4.17b displays the corresponding ROC curve. Mind the conventional change with previous ROC curves. Here it plots the signal efficiency against the background contamination rate. Black triangles show the efficiencies of the previous search based on the optimal likelihood. It is found that the Medal Game - Likelihood combination performs similarly as the Neural Network based analysis in terms of signal/background separation.

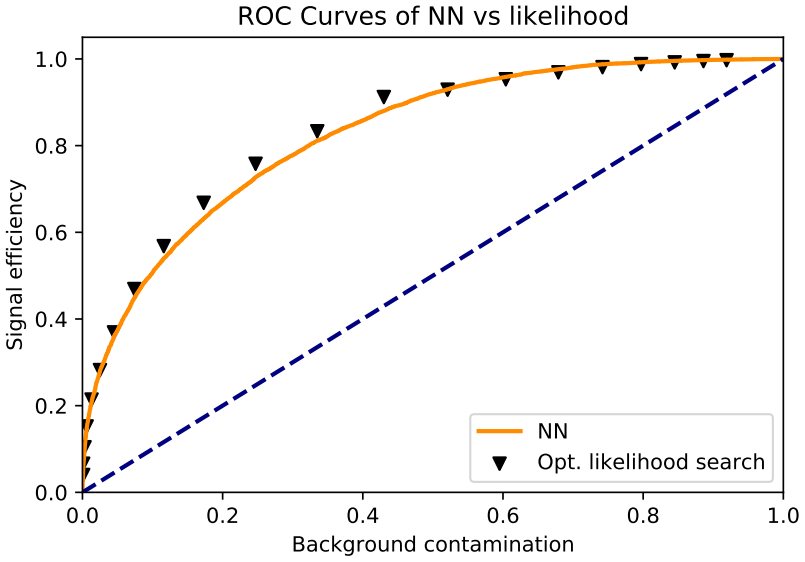
As a conclusion the neural networks built using the Keras libraries of the TensorFlow platform did not improve the likelihood analysis neither improved the correct ρ reconstruction in simulated ν_τ interactions. Modifications of the neural networks structures (number of neurons and hidden layers) did not prove to be of substantial help. The possibility to use larger sample sizes for the training of the neural networks has not been studied.

4.9 Summary

This chapter presented a likelihood based kinematic analysis to search for ν_τCC interactions in DUNE exploiting the large branching ratio (25%) of the $\tau \to \rho$ decay mode. The use of the ρ and π_0 invariant masses, among others, allowed to correctly reconstruct the ρ in ν_τ events with a 82% efficiency both for the standard CP optimized neutrino flux and the τ optimized neutrino flux. A total of 17 kinematical variables were studied to isolate the ν_τ events from the



(a) NN raw output



(b) ROC curves

Figure 4.17: Top: neural network score distributions of signal (blue) and background (red) events for which NN_5 is used to reconstruct one ρ candidate per event. Both histograms are normalized to unity. The smaller black histogram corresponds to signal events for which NN_5 misreconstructed the ρ in ν_τCC. Bottom: corresponding ROC curve (signal efficiency against background contamination) for the neural network based analysis (orange curve). For comparison the likelihood based performance is shown with black triangles.

main backgrounds which in this case are the neutral currents NC($\geq 1\pi_0 1\pi^{\pm}$) which mimic a ρ decay topology. Taking into account the detector effects with the smearing, a signal efficiency selection of 57% and a background rejection of 88% can be for instance achieved. We find that the likelihood performs significantly better on the QEL and RES scatterings than it does on the DIS. The use of the CVN as a pres-classification tool was found to bias the ν_τ sample in the sense that the signal efficiency on the biased sample was decreased at the level of 10%. However the CVN and the likelihood selections on the NC($\geq 1\pi_0 1\pi^{\pm}$) events are rather uncorrelated.

A normalization assuming the 3.5 years staged deployment plan of DUNE indicates a $\nu_\mu \rightarrow \nu_\tau$ sensitivity at the level of 2.8σ which corresponds to a signal/background ratio of 28/91. Assuming a 100% charged pion identification via specific final state topologies allows this significance to reach 3.4σ. The use of the optimized high energy τ neutrino beam is found to have, as in the $\tau \rightarrow e$ analysis, the greatest impact on DUNE sensitivity in regards of the $\nu_\mu \rightarrow \nu_\tau$ appearance. With the same level of normalization (3.5 years staged), the likelihood analysis allows to reach a statistical significance of 9σ.

In parallel the use of an artificial neural networks (NN) built with the Keras library of TensorFlow was not found to improve the analysis. A similar performance between the NN and the Medal Game is observed as far as the correct ρ reconstruction of ν_τ interactions is concerned. In both cases 82% efficiency is at best reached. The same conclusion applies for the signal/background separation analysis where a NN is found to perform similarly as the likelihood analysis.

$\star$ $\star$ $\star$

Chapter 5

The $\tau^- \to \pi^- \nu_\tau$ decay mode analysis

Contents

5.1 Motivations

This chapter proposes to extend the analysis deployed for the $\tau \to \rho$ decay mode to a more exclusive τ decay final state. The idea of exploiting the $\tau \to \rho$ decay mode relied on its combined large branching ratio (about 25%) and the opportunity to take advantage of the kinematical signature of the ρ resonance with the invariant masses (of the ρ itself and of the two photons coming from the decay $\pi_0 \to \gamma\gamma$). With the standard LBNF neutrino flux optimized for CP violation studies, we estimated DUNE to be statistically sensitive to the $\nu_\mu \to \nu_\tau$ appearance at the level of 3σ (which increased to 9σ by using the τ optimized neutrino flux) assuming a 3.5 years staged deployment plan (defined in 2.4.2).

One limitation encountered in this analysis was the sometimes large level of ambiguity of the ρ reconstruction within a ν_τCC event, that is to say the triplet $(\gamma_1\gamma_2\pi^-)$ which is the daughter system of the ρ decaying as $\rho^- \to \pi_0\pi^- \to \gamma_1\gamma_2\pi^-$. With both the CP optimized and τ optimized neutrino fluxes, the correct ρ was reconstructed in about 82% of the cases. This efficiency was hardly improvable. As for comparison the NOMAD experiment estimated in a similar analysis a correct ρ identification efficiency of 75% in [26] and 66% in [25]. However it is delicate to make a direct comparison between DUNE and NOMAD since NOMAD used a higher energy neutrino beam and as a consequence the hadronic system pions multiplicity was undoubtedly greater. The pions (neutral and charged) multiplicity was also the main reason of the method limitation.

It was illustrated in section 4.3.1 that a ν_τCC($\tau \to \rho$) event in which the hadronic final state is composed of one neutral pion and two charged pions leads to a level of ambiguity of 18 in the sense that the ρ reconstruction process must select among 18 ρ candidates. This multiplicity is mostly sensitive to the number of neutral pions.

One natural improvement to diminish the 18% bias (complementary of the ρ reconstruction efficiency of 82%) associated to the misreconstruction of the τ daughter particles can be reduced by studying an hadronic τ decay mode for which the final state is more exclusive, thus circumscribing *de facto* the multiplicity of τ daughter candidates. This is the key motivation to study the $\tau^- \to \pi^- \nu_\tau$ decay mode (later abbreviated $\tau \to 1\pi$) which accounts for 10.83% of the total branching ratio. It is all the more convenient that the analysis of this specific decay mode can be viewed as a simplified case of the one deployed for the $\tau \to \rho$ decay mode. In this chapter the detector effects are directly included with the smearing method and the charged pion (π^+/π^-) identification hypothesis is included only when explicitly mentioned. In the following it is implicitly assumed that in the studied ν_τCC interactions the charged lepton τ^- decays into one negatively charged pion and one tau neutrino, and that the electrical charge of the τ and its daughter π are negative ($\bar{\nu}_\tau$ interactions are not studied).

5.2 Identification of the correct decay daughter π

5.2.1 Medal Game

The first requirement is to correctly reconstruct the charged pion which is the daughter of the τ lepton, because the hadronic system may provide additional charged pions. The method is similar to the one deployed for the $\tau \to \rho$ decay mode analysis. The same terminology is also used:

Hadronic system pions	0	1	2	3	4	>4
Fraction (%)	63.7	23.6	7.1	3.6	1.3	0.7

(a)

True π ranking	-1	0	1	2	3	>3
Fraction (%)	0.4	63.5	30.5	4.7	0.7	0.2

(b)

Table 5.1: (a): Fake charged pion multiplicity in ν_τCC events. (b): Ranking of the τ daughter charged pion among all pion candidates, using the Medal Game defined in 5.2.1. A rank of -1 means the true pion was not reconstructed. A rank of 0 means the τ daughter pion is reconstructed and there is no pion belonging to the hadronic system (fake charged pion multiplicity=0).

- The **true** charged pion designates the pion which is the daughter of the charged τ lepton in a ν_τCC interaction.
- Any charged pion of the final state which is not the daughter of the τ lepton is defined as a **fake** pion.

The idea is to find suitable kinematical variables for which the fake pions and the true one behave differently. The three following variables are found to be rather discriminating:

- The kinetic energy of the pion $E^K_{\pi^\pm}$. It is expected that most of the neutrino energy goes to the leptonic system (τ lepton), so that the daughter pion inherits a substantial fraction of this energy while pions belonging to the hadronic system have less energy.
- It is also worth considering this energy normalized to the total visible energy. E^K_{had} is the total reconstructed energy visible in the hadronic system, excluding the pion candidate. We use the notation $\alpha^{(had)}_{\pi^\pm} = E^K_{\pi^\pm}/(E^K_{\pi^\pm} + E^K_{had})$.
- The third variable is the transverse pion momentum $p^{(tr)}_\pi$ normalized to the total transverse momentum: $\rho_L = p^{(tr)}_\pi/(p^{(tr)}_\pi + p^{(tr)}_{had} + p^{(tr)}_{miss})$. This variable has a value less than 1/2 since the transverse missing momentum is defined as $\overrightarrow{p}^{(tr)}_{miss} = -(\overrightarrow{p}^{(tr)}_\pi + \overrightarrow{p}^{(tr)}_{had})$.

Given a ν_τCC event, each of the charged pion candidates is ranked in descending order for each three variable separately. In each category the 1^{st} pion is rewarded with 3 medals, the 2^{nd} with 2 medals and the 3^{rd} with 1 medal. Afterward pion candidates are ranked according to their total number of medals. In case of tied candidates, the one with the highest reconstructed kinetic energy wins. This process elects without ambiguity a best pion candidate for each ν_τCC($\tau \to 1\pi$) event. Events containing 2 charged pions follow the same Medal Game rules except that there is no attribution of one single medal.

5.2.2 π reconstruction efficiency

The previously defined Medal Game performance is evaluated on the sample of ν_τCC events available in the simulation files used for the DUNE Technical Design Report. As for the $\tau \to \rho$ analysis, a rank of 0 is reserved for events in which the true pion is the only candidate because in this case there is no ambiguity in the pion reconstruction. A rank of -1 is attributed to

events in which the true pion is not reconstructed, even if there is no fake pion. For all the others events the Medal Game is used to sort all pion candidates and the rank of the true pion can be read at the MC truth level.

An overview of the fake pion multiplicity is shown in table 5.1. Approximately 94% of the ν_τCC events have a fake pion multiplicity smaller or equal to 2. It should be reminded that in the $\tau \to \rho$ analysis the corresponding fake ρ multiplicity fraction was 75% (see table 4.1a), and that 15% of events had a fake ρ multiplicity strictly greater than 6. The level of ambiguity of the τ reconstruction is indeed much lower in the $\tau \to 1\pi$ decay mode.

The Medal Game results are gathered in table 5.1b. A careful reader will observe that the fraction of true pions ranked 0 (63.5%) is smaller than the fraction of ν_τCC interactions with a fake pion multiplicity of 0 (63.7%): there exist few events for which the true pion, though the only candidate, is not reconstructed and the event is discarded (rank -1). Table 5.1b indicates that the true π is correctly identified in 94.0% of the cases. A normalization of the efficiency can be done removing the ranks 0 and -1 for which the Medal Game is not used. This leads to a corrected efficiency of 84%.

For completeness the distributions of the three variables used for the Medal Game are given in figure 5.1. The Monte Carlo truth is used to build the blue histograms with the true pions even if the Medal Game unsuccessfully reconstructed them. However, the red histograms (fake pions) were filled using the Medal Game to select only one fake pion candidate per ν_τCC event if such a candidate existed. Both histograms are afterward normalized to unity to allow for a direct comparison. This construction method thus bias the red histograms in the sense that the fake pion used is each time the one best ranked by the Medal Game. In the perspective of the separation between the true and the fake pions, the red histograms can be viewed as the "worst case scenario" distributions.

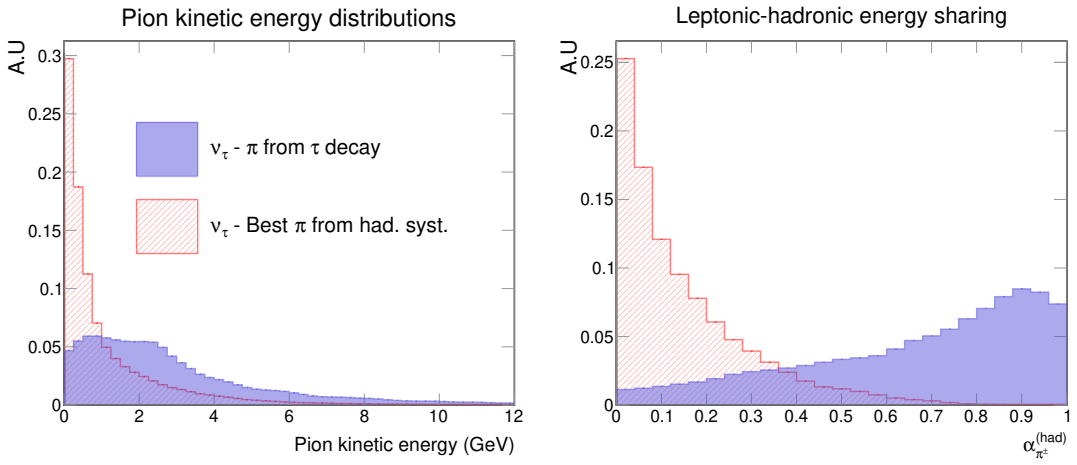
Figure 5.1a shows that the true pions have, as expected, more energy than the fake pions for which a large fraction is observed below 1 GeV. This is visible in figure 5.1b which shows the distribution of this energy normalized to the total reconstructed energy ($\alpha_{\pi^\pm}^{(had)}$). Finally it is seen from figure 5.1c that the true pion also carries the majority of the transverse momentum. These three variables are probably correlated and improving the performance of this Medal Game would require to add new uncorrelated information. But given that the total efficiency reached is 94%, the achievable gain is low and the present method is considered satisfactory.

5.3 Signal selection and background evaluation

The main background associated to the ν_τCC($\tau \to 1\pi$) signal is composed of the neutral currents for which the minimal final state topology has one charged pion. In the following, such events are noted NC($\geq 1\pi^\pm$). It should be noted that the strict topology "$1\pi^\pm$" is however not enough to constitute a background since the presence of an hadronic system in addition to the pion is required. This section is dedicated to the study of kinematical variables that help in discriminating between this neutral current background and the ν_τ signal.

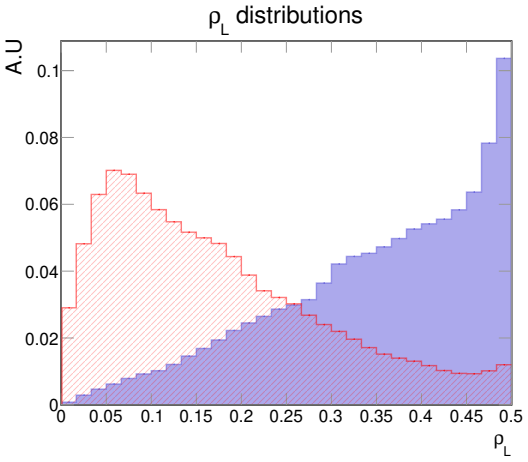
5.3.1 The set of exploitable kinematical variables

It was pointed out in the introduction of this chapter that the signal/background analysis of the $\tau \to 1\pi$ decay would be a simplified case of the analysis deployed for the $\tau \to \rho$ decay mode.



(a)

(b)



(c)

Figure 5.1: Kinematical distributions of the variables used in the Medal Game defined in 5.2.1. MC truth was used in order to build blue histograms with the τ daughter charged pion. For the red histograms the Medal Game was used to select in each ν_τ event the best pion candidate of the hadronic system (if such pion was available). A discussion of the three plots is provided in the text (see 5.2.2).

The kinematical variables used for the current analysis are indeed greatly inspired from 4.4.1. The following 14 kinematical variables are studied:

- The 3 variables already used in the Medal Game (see definitions in 5.2.1).
- 4 laboratory angles. $\theta_{\pi h}$ is the angle between the pion candidate and the hadronic reconstructed momenta, $\theta_{\pi tot}$ is the angle between the pion candidate and the total momenta of the event, $\theta_{\pi\nu}$ is the angle between the pion candidate momentum and the neutrino beam direction and $\theta_{h\nu}$ is the angle between the hadronic momentum and the neutrino beam direction.
- The modulus of the pion candidate momentum projected in the transverse plane of the interaction $p_\pi^{(tr)}$. $p_{had}^{(tr)}$ is defined in the same way for the hadronic system. These two variables allow as well defining the transverse missing momentum $p_{miss}^{(tr)}$.
- The relative orientations of the three transverse momenta are described with the angles $\phi_{h\pi}^{(tr)}$ (between the pion and the hadronic system), $\phi_{hm}^{(tr)}$ (between the hadronic system and the transverse missing momentum) and $\phi_{m\pi}^{(tr)}$ (between the transverse missing momentum and the pion).
- The transverse mass defined as

$$M^{(tr)} = 2\sqrt{p_\pi^{(tr)} p_{miss}^{(tr)}} \left| \sin\left(\frac{\phi_{m\pi}^{(tr)}}{2}\right) \right|. \tag{5.1}$$

The reader should remind that the numerical values of these variables depend on the choice of the τ decay reconstruction. For the signal the MC truth is used in order not to bias the distributions but for the background NC($\geq 1\pi^\pm$) the Medal Game is used to reconstruct a τ daughter candidate. The histograms are then smoothed (with a standard algorithm of the ROOT libraries) and normalized to unity to be used as probability density functions.

5.3.2 Most sensitive kinematical variables for signal selection

The distributions of 6 of the 14 variables defined in the previous section are presented in figure 5.2 where the signal (background) is shown in blue (red).

Figure 5.2a shows that for NC events the pions reconstructed as the τ daughter have less energy than the true pions of the ν_τ events. It is all the more visible in figure 5.2b where this energy is normalized to the total reconstructed energy of the interaction. In addition the true pion of ν_τ carries substantially more transverse momentum than the pions of NC($\geq 1\pi^\pm$) (see figure 5.2c).

The angles $\theta_{\pi tot}$ and $\theta_{\pi h}$ are also shown. The large tail of the background distributions up to 180° might be surprising because the fake pion candidates of NC($\geq 1\pi^\pm$) *in fine* belong to the real hadronic systems. These two tails are in fact correlated with low energy pions (typically $\leq 500\,\text{MeV}$) which tend to spread more than high energy pions with respect to the hadronic system direction and the total reconstructed momentum.

The last plot (figure 5.2f) shows the transverse angle $\phi_{hm}^{(tr)}$ for which the background distribution peaks at 180° while the signal distribution is rather flat. In NC events, the real leptonic system is actually composed of a neutrino only. As the (real) hadronic system and the neutrino go in

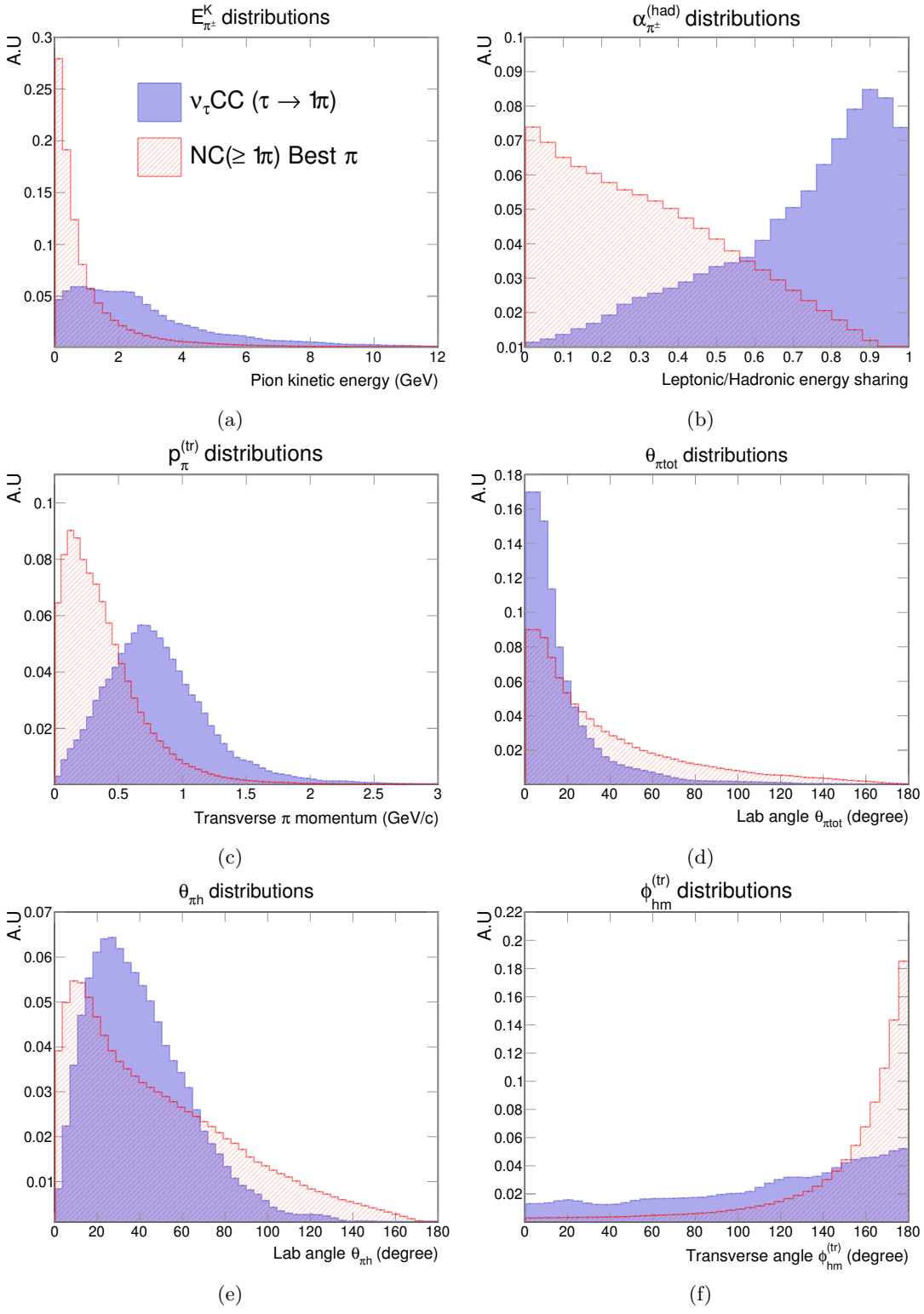


Figure 5.2: Signal/background distributions of six kinematical variables showing discriminating power to discriminate between ν_τCC($\tau \to 1\pi$) (blue) and NC($\geq 1\pi^{\pm}$) events (red). From left to right and top to bottom: the kinetic energy of the pion which is then normalized with respect to the total reconstructed energy ($\alpha^{(had)}_{\pi^{\pm}}$), the transverse pion momentum, the laboratory angle between the pion momentum and the total reconstructed momentum of the interaction, the laboratory angle between the pion and the hadronic system and finally the angle between the transverse hadronic momentum and the transverse missing momentum.

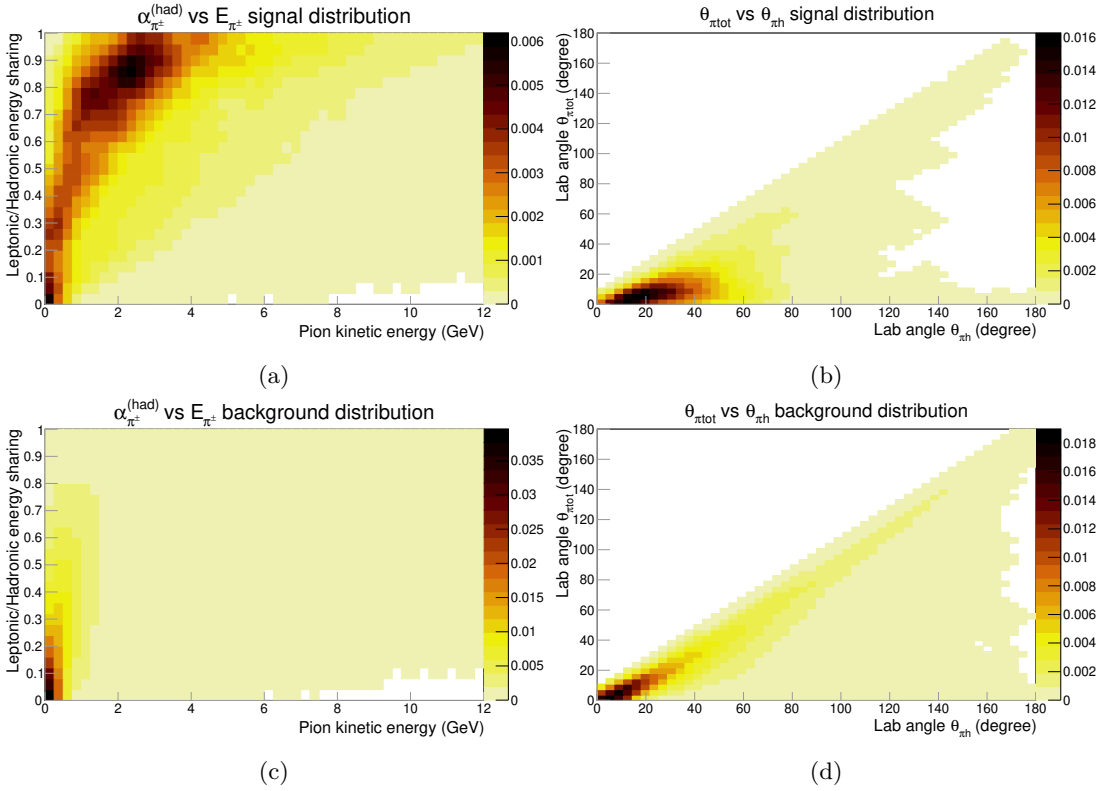


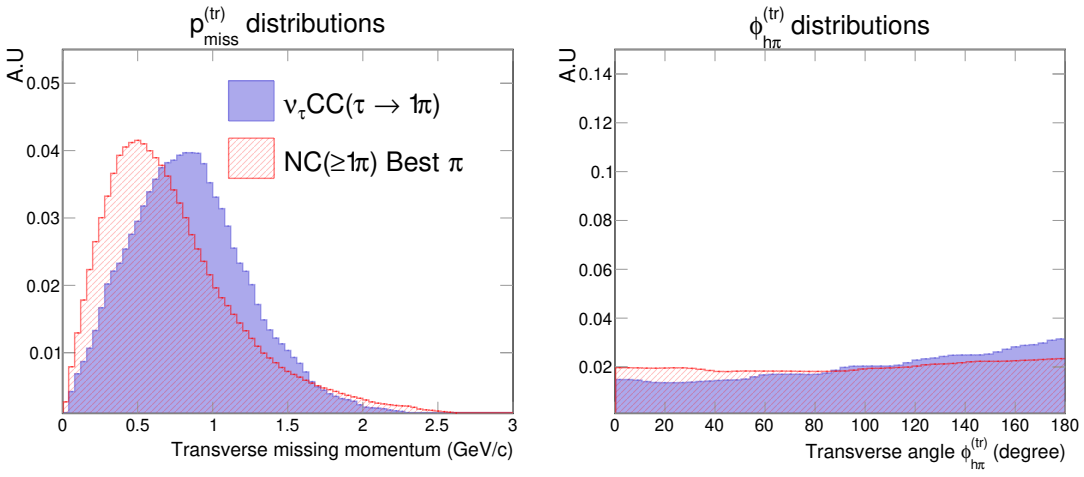
Figure 5.3: 2-dimensional distributions of correlated variables which show great signal/background separation power. Top (bottom) plot corresponds to the signal (background) distributions. Left: correlation between the pion kinetic energy and the pion kinetic energy normalized to the total reconstructed energy ($\alpha_{\pi^\pm}^{(had)}$). Right: correlation between the two laboratory angles $\theta_{\pi tot}$ and $\theta_{\pi h}$. A discussion of the plots is provided in the text (5.3.2).

opposite directions in the transverse plane, the missing energy is dominated by the neutrino so that the angle between the transverse missing momentum and transverse hadronic momentum favours high angles close to 180°.

In figure 5.3 are shown two 2-dimensional distributions that will be used in the following. The first (left) is the correlation between the pion kinetic energy (x-axis) and the pion kinetic energy normalized to the total reconstructed energy of the neutrino interaction. The background distribution (bottom left) is tight to low values for both variables while the corresponding signal (top left) distribution occupies a much wider region of the plane. A strong correlation between the two variables is observed. The second (right) 2-dimensional distribution describes the angles in the laboratory frame $\theta_{\pi h}$ (x-axis) and $\theta_{\pi tot}$ (y-axis). The background distribution is more spread in the plane which indicates large tails of the corresponding 1-dimensional distributions.

5.3.3 Illustration of non-discriminating variables

Some kinematical variables useful for the $\tau \to e$ decay mode analysis are no longer showing discriminating power for the present analysis. Figure 5.4 shows two such variables: the transverse missing momentum and the angle between the transverse pion momentum and the transverse hadronic momentum.



(a) (b)

Figure 5.4: Kinematical distributions of variables deployed for the signal(blue)/background(red) separation analysis which have no evident discriminating power. Left: transverse missing momentum. Right: angle between the transverse pion and hadronic momenta.

Indeed the transverse missing momenta of the signal and background are comparable. It can be understood by the fact that both have one undetected final state neutrino carrying a substantial missing energy in the transverse plane. On the other hand the two distributions of $\phi_{h\pi}^{(tr)}$ are flat. This means that there is no preferred orientation in the transverse plane between the pion momentum and the hadronic momentum. Again the presence of the final state neutrino allows having this uniformly distributed configuration.

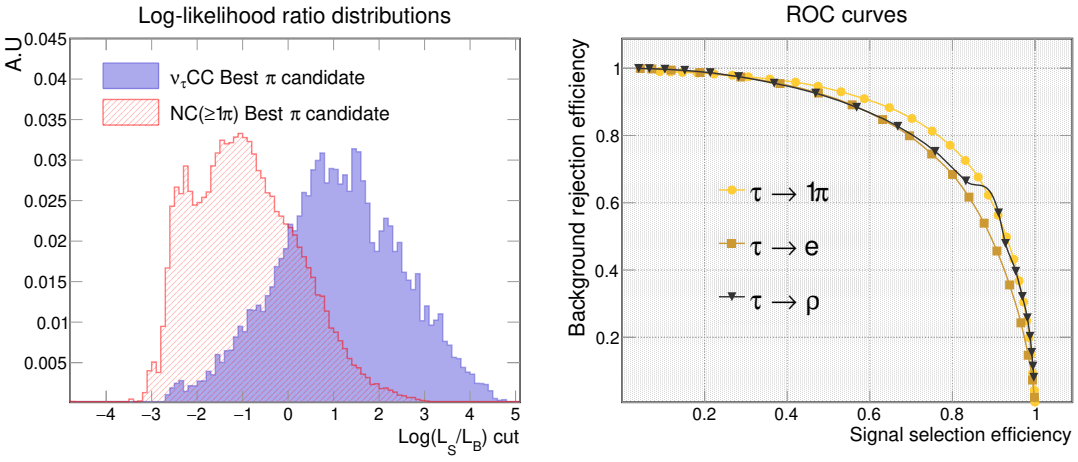
5.4 Likelihood based analysis

5.4.1 Signal selection and background rejection efficiencies

An optimal separation power combining the likelihoods of the two 2-dimensional variables shown in figure 5.3 and adding the likelihood of the transverse momentum of the pion is found. The corresponding log-likelihood ratio distributions of the signal (blue) and the background (red) are shown in figure 5.5a. The Medal Game is now used to reconstruct the pion τ daughter both for the signal and background events. A previous study in this chapter assessed that this method introduces a 6% bias in the pion reconstruction of ν_τ interactions.

Figure 5.5b shows the corresponding ROC curve (background rejection efficiency against signal selection efficiency). As for comparison the ROC curves obtained for the two other decay mode analysis are displayed as well. For the $\tau \to e$ decay mode only the ROC curve corresponding to the oscillated ν_e background is shown.

This plot suggests that the signal/background separation associated to the $\tau \to 1\pi$ decay mode is better than the other two. In addition in the $\tau \to \rho$ analysis there was a large 18% bias due to the ρ misreconstruction in $\nu_\tau(\tau \to \rho)$ events. A second type of background also affected the $\tau \to e$ analysis, namely the beam ν_e which were associated to electron neutrinos contaminating the produced muon neutrino beam for which we had a less good signal/background ROC curve (see figure 3.14b).



(a) (b)

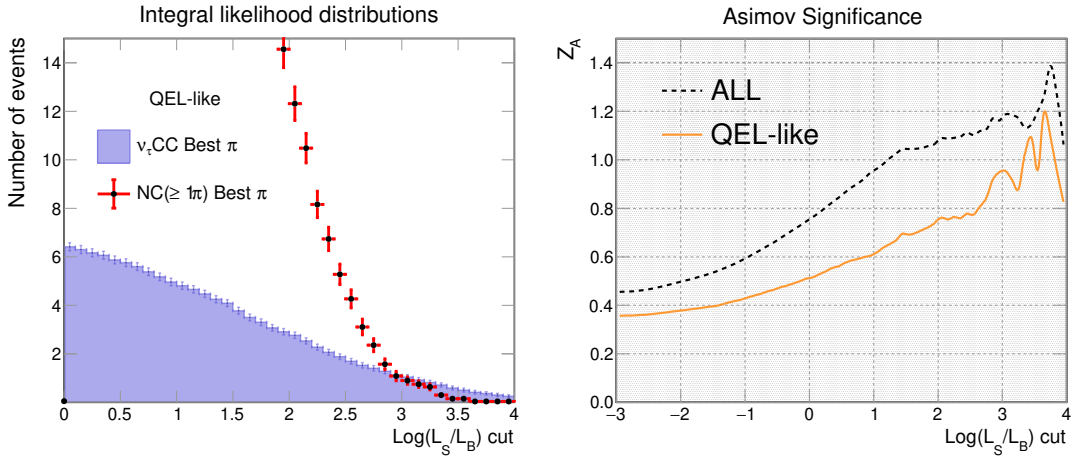
Figure 5.5: Left: Log-likelihood ratio distributions of ν_τCC (blue) and NC($\geq 1\pi^\pm$) (red) for which the Medal Game is used to reconstruct the τ daughter pion. Right: corresponding ROC curve (background rejection efficiency against signal selection efficiency) in yellow with circle markers. As for comparison, the previous ROC curves obtained for the $\tau \to e$ analysis (dark orange with square markers) and $\tau \to \rho$ analysis (black with inverted triangle markers) are shown. The bump of the black spline is an artefact from the ROOT smoothing algorithm.

5.4.2 Normalization to the 3.5 years staged deployment plan

The previous efficiencies are now used to predict the event rates assuming the normalization to the 3.5 years staged deployment plan of DUNE which is explained in 2.4.2. The expected number of events are taken from table 2.2: $270 \times 10.83\% \simeq 29$ signal events are expected against 8832 NC events. Few signal events are discarded either because there was no reconstructed pion in the final state or because the hadronic system was empty. The fraction of NC($\geq 1\pi^\pm$) (NC which has the minimal topology required to constitute a background) is assessed to be 47.2%. The initial signal to background ratio is then 29/4169.

An additional requirement imposes the final state of the interactions to be QEL-like (*i.e* has the topology $1p1\pi^\pm$ where p stands for proton). The objective of this chapter is to study very exclusive final state neutrino events. This condition keeps 27% of the remaining ν_τ (while there are 46% true QEL interactions, see figure 2.7) and 12% for the neutral currents. It is found that the large majority (98%) of QEL-like ν_τ interactions are true QEL interactions at the MC truth level. The QEL-like requirement increases the initial signal/background by a factor of 2. However it is found that the likelihood separation is less performing on this specific exclusive topology, which compensate to some extent the gain of 2 achieved.

Figure 5.6 shows the two figures of merit obtained with the same 3.5 years staged normalization. The first on the left is merely the signal and background expected number of events as a function of the log-likelihood ratio cut. Error bars are statistical only. This figure describes the difficulty to isolate a clean signal region even with the satisfying efficiencies exposed in the previous section. A cut at 3 (necessary to obtain a S/B ratio of 1) would indeed remove more than 99% of the background, however for this value the signal efficiency is 13% only. The Asimov significance as a function of the cut value used is shown on the right. It reveals that with this decay mode, DUNE would only be sensitive to the ν_τ appearance at the level of about 1σ. It also shows that study of exclusive QEL-like events makes DUNE a bit less sensitive at the level of 0.15σ.



(a) (b)

Figure 5.6: Optimized likelihood search normalized with respect to the 3.5 years staged deployment plan. Left: number of QEL-like events expected as a function of the log-likelihood cut value. Errors are statistical only. Right: corresponding Asimov significance of the ν_τ appearance (orange curve), and comparison with the case where we no longer restrict the analysis to QEL-like final state topologies (black dashed).

True π ranking (%)	-1	0	1	2	3	>3
No $\pi^\pm$ identification	0.4	63.5	30.5	4.7	0.7	0.2
$\pi^+/\pi-$ identification	0.4	87.6	10.3	1.6	0.1	< 0.1

Table 5.2: True pion rank in ν_τ events where the Medal Game is used to resolve the pion reconstruction ambiguity when the multiplicity is greater or equal to 2. The results obtained without the charged pion identification (second raw) are also shown for comparison.

With the standard LBNF neutrino beam the other decay mode analysis of $\tau \to e$ and $\tau \to \rho$ allowed achieving a sensitivity at the level of 1.8σ and 2.8σ respectively. Though the $\tau \to 1\pi$ is the most promising decay mode in terms of mere signal/background discrimination, it is in fact the less sensitive of the three decay modes studied. It is penalized both by a low branching ratio (11% against 18% for the $\tau \to e$ and 25% for the $\tau \to \rho$) and an unfavourable large amount of background events. Indeed the initial signal to background ratio with the QEL-like criteria is at the level of 2% to be compared with 3% (4%) for the $\tau \to e$ ($\tau \to \rho$) decay mode.

5.4.3 Charged pion identification

The possibility of the LArTPC technology to identify the charge of the pions via specific final state topologies allows evaluating the effect of such a capacity on the previous results. This capabilty has an efficiency that can be studied (see for instance [105]). Here, a 100% distinction between π^+ and π^- is assumed in order to assess the maximal impact that such a capacity would have on DUNE sensitivity to the beam ν_τ appearance. The exclusive QEL-like final state topology is no longer required (otherwise associated statistics is too low).

First the evolution of the fake pion multiplicity is shown in figure 5.7. The black (dashed gray) histogram gives the multiplicity of π^- ($\pi^\pm$) from the hadronic system. The fraction of events with at least one charged pion is 36% against 12% for the fraction with at least one negatively charged pion. It indicates that the charge of the pion for events containing only one charged

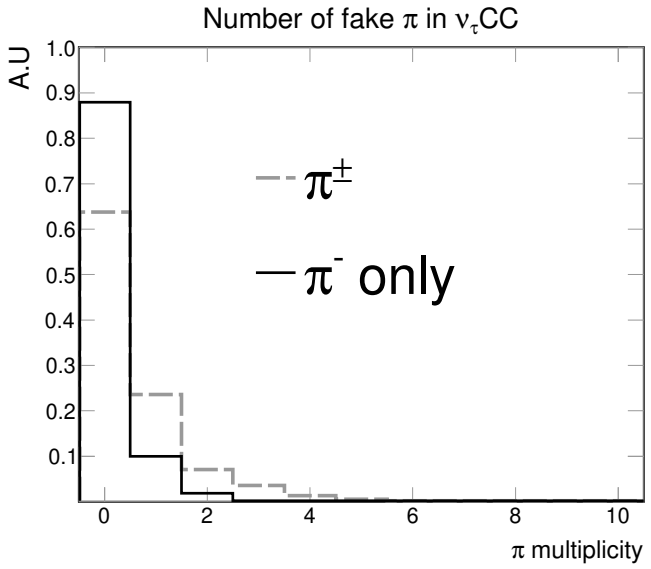


Figure 5.7: Fake charged pion multiplicity (gray dashed) and fake negatively charged pion multiplicity (black full line) in ν_τCC events.

pion in the final state is more likely to be positive. Because of charge conservation, the hadronic system of a ν_τCC event must be positively charged, so it is indeed likely to observe less π^- than π^+ in the hadronic system.

Table 5.2 illustrates that the Medal Game only acts on 12% of the events and overall the true pion is correctly reconstructed in 98% of the cases. A very low contamination of fake pions in the reconstruction of the τ daughter is achieved thanks to the hypothesis of the charged pion identification.

In terms of signal/background separation, the biggest change is the fraction of NC which participate to the background. Such events must have a minimal topology containing at least one π^- and a corresponding non-null hadronic system (noted NC($\geq 1\pi^-$)). They represent 35.2% of the total NC events while the NC($\geq 1\pi^\pm$) fraction is 47.2%.

The likelihood performance is found to be only slightly better. Figure 5.8 shows the Asimov significance with (without) the charged pion identification as a function of the log-likelihood cut ratio in orange (black dashed). For cut values up to 2, there is a significance increase of about 0.1. The new significance reaches a maximum at 1.2 for a cut value of 2, then rapidly decreases, a feature not observed without the charged pion identification. However the region at log-likelihood values above 3 has a low statistic and should be interpreted carefully.

To conclude this part, the charged pion identification has a substantial impact on the correct π identification in ν_τCC interactions (98% against 94% without), and allows for a limited significance increase of about 0.1 (assuming the DUNE 3.5 years staged deployment plan) thanks to a more favourable initial signal/background ratio.

5.4.4 τ optimized flux

We finally report the improvement allowed by the use of the alternative τ optimized neutrino flux which was presented in 2.2.2. This high energy neutrino flux would allow to compensate to

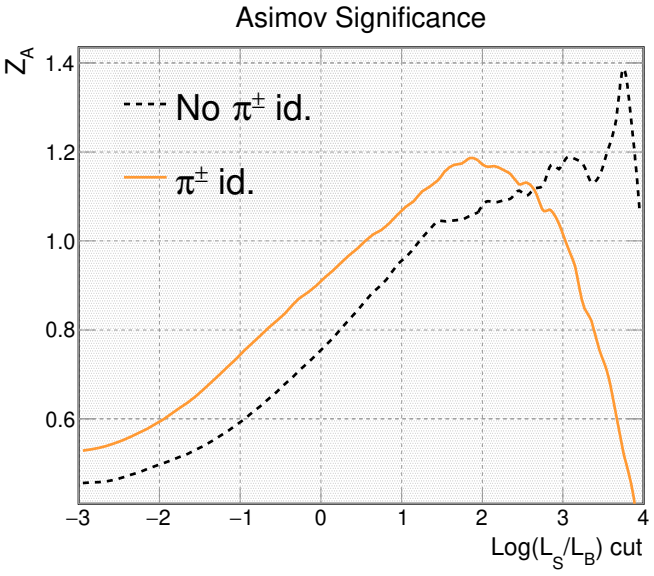


Figure 5.8: 3.5 years staged Asimov significance as a function of the log-likelihood cut ratio with (without) the charged pion identification assumption in black (dashed orange).

True π ranking	-1	0	1	2	3	>3
CP optimized flux	0.4	63.5	30.5	4.7	0.7	0.2
τ optimized flux	0.3	56.7	37.2	5.1	0.6	<0.1

Table 5.3: True pion rank in ν_τ events where the Medal Game is used to resolve the pion reconstruction ambiguity when the multiplicity is greater or equal to 2. A comparison of the results obtained with the two neutrino fluxes is shown.

some extent the kinematically suppressed ν_τCC cross section at energies of few GeV (and above all the critical ν_τCC energy threshold of 3.45 GeV). It has already been reported in this thesis that the use of this neutrino flux would multiply the ν_τCC statistics by a factor of 6. In this section the charged pion identification is no longer assumed.

The ν_τCC fake pion multiplicity and the Medal Game efficiency to reconstruct the correct pion are first evaluated. Figure 5.9 displays the fake pion multiplicity and reveals that there is a decrease of about 10% in the fraction of multiplicity 0 corresponding at first order to the increase of the multiplicity 1. The rest of the multiplicities are little affected.

Table 5.3 reports the true pion ranking as performed by the Medal Game. The sum of ranks 0 and 1 leads to a τ reconstruction efficiency of 93.9% with the τ optimized beam which is (almost) equal to the 94.0% efficiency obtained with the standard LBNF neutrino flux. As another indicator, the normalization of the rank 1 obtained without the ranks 0 and -1 leads to a correct τ reconstruction efficiency of 87%, slightly better than the 84% obtained with the standard LBNF neutrino flux.

The signal/background analysis is performed again. The first step is to reweight the kinematical distributions in the light of the alternative high energy neutrino beam. The NC($\geq 1\pi^{\pm}$) topology now represents 63.9% of the total NC events (against 47.2% with the standard neutrino beam optimized for CP violation studies). No likelihood combination was found to perform better than the one presented in 5.4.1 though others led to similar results. Signal and background events without reconstructed pions are discarded as well as events without reconstructed hadronic system in addition to the pion candidates (few events, at the level of 1%, are concerned). The

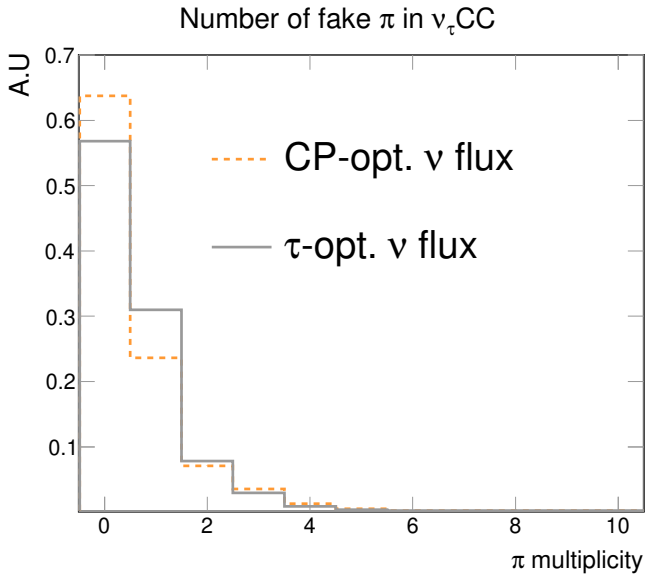
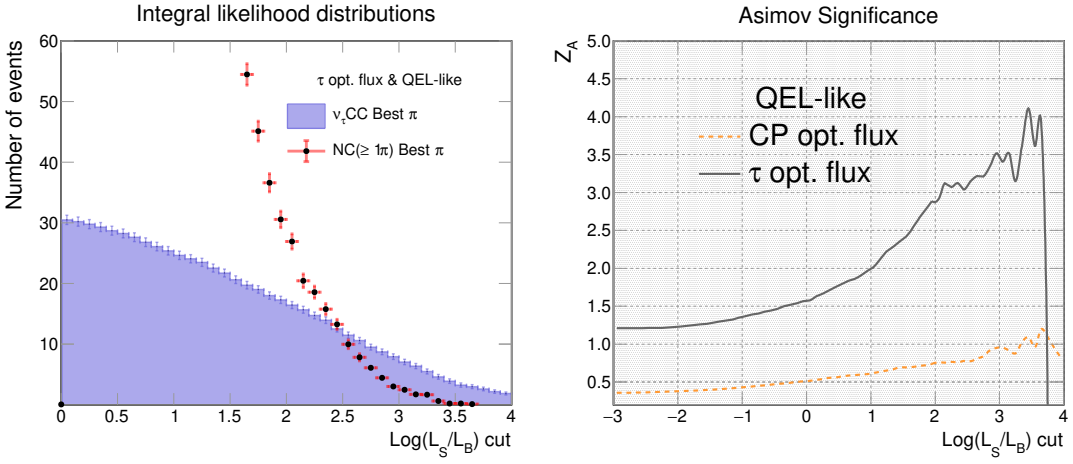


Figure 5.9: Fake charged pion multiplicity in ν_τCC interactions with CP optimized neutrino beam (τ optimized neutrino beam) in orange dashed (gray histogram).



(a) (b)

Figure 5.10: Optimized likelihood search normalized with respect to the 3.5 years staged deployment plan with the τ optimized flux, restricted to the QEL-like final state topologies. Left: number of signal and background events expected as a function of the log-likelihood cut value. Errors are statistical only. Right: corresponding Asimov significance of the ν_τ appearance.

raw discriminating power of the analysis is barely affected by the use of the alternative neutrino flux and is not displayed here.

The requirement of the QEL-final state topology allows the selection of 20% of the remaining signal and 8% of the remaining background. A normalization corresponding to the DUNE 3.5 years staged deployment plan is performed and results are shown in figure 5.10. The left plot shows the expected number of signal and background events where the error bars are statistical only. The plot on the right shows the evolution of the Asimov significance between the standard LBNF neutrino flux and the τ optimized flux. For a cut value of 2, the significance increases from 0.8 to 2.9, corresponding to a S/B ratio of 16/27.

5.5 Summary

This chapter suggested to extend the analysis deployed for the $\tau^- \to \rho^-\nu_\tau \to \pi_0\pi^-\nu_\tau$ decay mode to the more exclusive $\tau^- \to \pi^-\nu_\tau$ decay mode (10.83% branching ratio) and take advantage of the associated low ambiguity in the τ^- decay reconstruction. It was found that for ν_τCC ($\tau \to 1\pi$) events a correct reconstruction was achieved in 94% of the cases in comparison to the 82% ρ reconstruction of the $\tau \to \rho$ analysis.

The main backgrounds associated to the $\nu_\tau(\tau \to 1\pi)$ signal are the neutral currents with a final state charged pion multiplicity of at least one (NC($\geq 1\pi^\pm$)) which represents 47.2% of the total NC events assuming the standard LBNF neutrino beam optimized for CP violation studies. Though the raw signal/background discrimination achieved in this analysis is found to be the best of the three decay mode studied, this decay mode suffers from the highest level of background and the lowest τ decay branching ratio. The restriction to the QEL-like final state topology ($1p1\pi^\pm$) allows to gain a factor of 2 in the S/B ratio though it reduces the signal statistics by a factor of 4. The best significance achievable assuming the 3.5 years staged normalization is 0.8σ with the QEL-like final state topology requirement and it corresponds to a S/B ratio of 3/12.

Finally the use of the high energy τ optimized neutrino beam which enhances the τ neutrino event rate by a factor of 6 allows to achieve a significance of 3σ (assuming the 3.5 years staged normalization and requiring a QEL-like topology) corresponding to a S/B ratio of 16/27. The raw likelihood performance is found to be similar for the two beam configurations (CP optimized and τ optimized).

⋆ ⋆ ⋆

Chapter 6

Final Remarks

6.1 τ decay channels combined sensitivity

We propose in this short section to combine the sensitivities obtained with the three decay mode analysis. The restriction to the QEL-like topology in the $\tau \to 1\pi$ analysis allows disentangling the NC backgrounds of the $\tau \to \rho$ and $\tau \to 1\pi$ analysis. Indeed, the minimal topology required for the neutral currents in the ρ analysis is NC($\geq 1\pi_0 1\pi^{\pm}$) while for the QEL-like of the $\tau \to 1\pi$ analysis it is NC($1p1\pi^{\pm}$).

For each decay mode analysis, a log-likelihood cut value is chosen so that it optimizes the sensitivity of this particular analysis. When such a maximum is not found (typically the sensitivity increases until the signal becomes very low), the cut value is chosen in order to avoid a too low signal. The CVN pre-selection is not included neither the charged pion identification assumption. The detector effects are taken into account via the smearing process and the oscillation parameters are taken from table 2.1.

Table 6.1 gathers the results of the number of signal/background events for each decay mode corresponding to the 3.5 years staged normalization. It is assumed that DUNE operates in the neutrino beam configuration (forward horn current) and both the standard LBNF neutrino beam (optimized for CP violation studies) and the τ optimized neutrino beam are presented. The charged current interactions of $\bar{\nu}_e$ and $\bar{\nu}_\tau$ are not included in this table. Table 2.2 indicates the $\bar{\nu}_e$CC represent about 5% of the ν_eCC interactions (assuming a constant density, standard LBNF neutrino flux and normal ordering) and the $\bar{\nu}_\tau$CC 10% of the ν_τCC. A more complete analysis would require to take them into account in the signal/background definitions.

The sensitivities associated to each decay mode analysis are recalled and the combined sensitivities are shown at the bottom of table 6.1. Errors are statistical only, they depend on the sample size of the simulated events available. The message here is that the $\tau \to e$ and $\tau \to \rho$ analysis dominate the total sensitivity, and the use of the alternative τ optimized neutrino beam allows for a gigantic boost of the sensitivity to the $\nu_\mu \to \nu_\tau$ beam oscillations ($3.0\sigma \to 13.2\sigma$).

6.2 Conclusion

By the end of the decade, DUNE (Deep Underground Neutrino Experiment) is expected to start operate the study of the world's most intense (1.2 MW) neutrino beam produced at Fermilab.

	Standard LBNF ν beam	τ optimized beam
$\tau \to$ e		
$\nu_\tau(\tau \to e)$	22.4 ± 0.2	151.6 ± 1.2
ν_e osc.	87.0 ± 0.5	143.6 ± 0.5
ν_e beam	63.6 ± 1.5	82.3 ± 2.0
ν_e total	150.6 ± 1.5	225.9 ± 2.1
Significance	1.8 ± 0.0	9.2 ± 0.1
$\tau \to \rho$		
$\nu_\tau(\tau \to \rho)$	18.8 ± 0.2	116.2 ± 0.9
NC($\geq 1\pi^{\pm}1\pi_0$)	40.0 ± 1.2	122.5 ± 3.3
Significance	2.8 ± 0.0	9.3 ± 0.1
$\tau \to 1\pi$ (QEL-like)		
$\nu_\tau(\tau \to 1\pi)$	2.8 ± 0.1	16.4 ± 0.6
NC($\geq 1\pi^{\pm}$)	12.3 ± 0.7	26.9 ± 1.3
Significance	0.8 ± 0.0	2.9 ± 0.1
3 channels combined		
ν_τ	44.0 ± 0.3	284.2 ± 1.6
Backgrounds	202.9 ± 2.1	375.4 ± 4.1
Significance	3.0 ± 0.0	13.2 ± 0.1

Table 6.1: Optimal results of the three τ decay mode analysis for the two beam neutrino fluxes. Each subsection presents the expected signal and background assuming a 3.5 years staged normalization. At the bottom a combination of the three decay modes is shown.

A combined near detector site and far detector complex will allow the experiment to probe the possible CP violation in the leptonic sector via the study of the $\nu_\mu \to \nu_e$ oscillation channel. The establishment of non-0 and non-π values of the still unconstrained phase parameter δ_{CP} of the PMNS matrix, which would rime with a CP violation discovery, is the core of the primary scientific program of DUNE.

In addition the experiment will have an unprecedented sensitivity to the ν_τ appearance thanks to the large $\nu_\mu \to \nu_\tau$ oscillation channel along its 1285 km baseline. The study of this poorly explored neutrino flavour will allow performing a unique 3-flavour oscillation phenomenology and offer a test for the unitarity of the PMNS matrix and for the 3-flavour paradigm. In addition DUNE will have the opportunity to operate with an alternative beam configuration tuned to enhance the ν_τ detection.

The prompt decay of the charged leptons $\tau^{\pm}$ will enjoin DUNE to deploy a kinematical analysis to search for $\tau^{\pm}$ decay signatures like the pioneering NOMAD experiment at the end of the 90's. This thesis explored three τ^- decay modes and developed for each of them a dedicated signal/background analysis.

In a first place the $\tau^- \to e^-\nu_\tau\bar{\nu}_e$ was explored for its large branching ratio (17.83%) and the final state topology containing an electron. A likelihood based analysis in the transverse plane of the beam neutrino interaction assuming a 3.5 years staged normalization allows reaching a statistical significance of 1.8σ with the standard LBNF neutrino beam and 9σ with the alternative τ optimized neutrino beam in regards of the $\nu_\mu \to \nu_\tau$ oscillation channel. The $\tau^- \to \mu^-\nu_\tau\bar{\nu}_\mu$ decay mode was rapidly reviewed but it was found to be largely unfavourable because of the large associated ν_μ background.

In a second movement the $\tau^- \to \rho^-\nu_\tau \to \pi_0\pi^-\nu_\tau$ was explored for its large branching ratio (25.52%) and the possibility to exploit the kinematic signature associated to the ρ resonance. The ρ reconstruction of the associated ν_τ had an efficiency of 82%. The dominant background are the NC($\geq 1\pi_0 1\pi^\pm$) for which the final state topology can mimic a ρ decay. A large significance of 3.4σ was achieved with the standard LBNF neutrino beam which increases to 9σ with the τ optimized neutrino beam. These significances assume the 3.5 years staged deployment plan of DUNE.

Finally the more exclusive semi-leptonic decay mode $\tau^- \to \pi^-\nu_\tau$ was studied to exploit the low level of ambiguity in the τ^- reconstruction. The latter was found to be the less sensitive (0.8σ with the standard LBNF neutrino beam and 3.0σ with the τ optimized neutrino beam) decay mode analysis though it allowed reaching the best signal/background separation power. The background considered were the NC($\geq 1\pi^\pm$).

Several features in addition to the likelihood analysis were explored such as the use of the CVN as a pre-selection tool which at the time is found to lower the significance of all three decay modes. However it should be pointed out that a combined CVN-likelihood analysis could be powerful in the sense that it would combine the visual discriminating power of the CVN with the fine kinematical analysis of the likelihood approach necessary to identify the τ neutrino flavour in DUNE. The charged pion identification assumption was studied in the semi-leptonic decay modes and was found to have a moderate impact on the significance. In parallel, the use of machine learning techniques (Boosted Decision Trees and Artificial Neural Networks) were found to perform similarly as the likelihood tools on the various signal/background separation problems.

A combined sensitivity assuming the 3.5 years staged normalization of the three decay modes was given in a last section. With the CP optimized neutrino beam, the $\tau \to \rho$ decay mode (2.8σ) dominates the overall sensitivity (3.0σ). However with the τ optimized neutrino beam, the $\tau \to e$ and the $\tau \to \rho$ decay modes are found to have a similar sensitivity at the level of 9σ. The total statistical sensitivity reached is at the level of 13σ which shows the major improvement permitted by this alternative neutrino beam configuration in regards of the $\nu_\mu \to \nu_\tau$ appearance phenomenon.

$\star$ $\star$ $\star$

Bibliography

[1] J. Chadwick. Intensitätsverteilung im magnetischen spektren der β-strahlen von radium b + c. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 1914. ***Cited on page 16***

[2] C. D. Ellis and W. A. Wooster. The average energy of disintegration of radium e. *The Royal Society Publishing*, 1927. ***Cited on page 16***

[3] W. Pauli. Open letter to the group of radioactive people at the Gauverein meeting in Tübingen, December 1930. ***Cited on page 16***

[4] E. Fermi. Tentativo di una teoria dei raggi β. *Il Nuovo Cimento*, 11, 1933. doi: https://doi.org/10.1007/BF02959820. ***Cited on page 17***

[5] F. L. Wilson. Fermi's theory of beta decay. *American Journal of Physics*, 36:1150–1160, 1968. ***Cited on page 17***

[6] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: a confirmation. *Science, New Series*, 124(3212):103–104, 1956. ***Cited on pages 17 and 20***

[7] G. Danby, J-M Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Physical Review Letters*, 9:36–44, July 1962. doi: 10.1103/PhysRevLett.9.36. ***Cited on pages 17, 21, and 50***

[8] The ALEPH Collaboration. A precise determination of the number of families with light neutrinos and of the Z boson partial widths. *Physics Letters B*, 235(3):399–411, 1990. doi: 10.1016/0370-2693(90)91984-J. ***Cited on page 17***

[9] The DONuT Collaboration. Observation of tau neutrino interactions. *Physics Letters B*, 504:218–224, April 2001. ISSN 03702693. doi: 10.1016/S0370-2693(01)00307-0. ***Cited on pages 17 and 37***

[10] R. Davis. A review of the Homestake solar neutrino experiment. *Progress in Particle and Nuclear Physics*, 32:13–32, January 1994. doi: 10.1016/0146-6410(94)90004-3. ***Cited on pages 18 and 20***

[11] B. Pontecorvo. Neutrino Experiments and the Problem of Conservation of Leptonic Charge . *SOVIET PHYSICS JETP*, 26(5):984–988, 1968. ***Cited on page 19***

[12] The Kamiokande Collaboration. Experimental study of the atmospheric neutrino flux . *Physics Letters B*, 205(2):416–420, 1988. ISSN 03702693. doi: 10.1016/0370-2693(88)91690-5. ***Cited on page 19***

[13] The Kamiokande Collaboration. Observation of a small atmospheric ν_μ/ν_e ratio in Kamiokande. *Physics Letters B* , 280(1):7, 1992. ***Cited on page 19***

[14] The Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, August 1998. doi: 10.1103/PhysRevLett.81.1562. ***Cited on pages 19 and 21***

[15] The SNO Collaboration. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the sudbury neutrino observatory. *Phys. Rev. Lett.*, 87, jul 2001. doi: 10.1103/PhysRevLett.87.071301. ***Cited on page 19***

[16] F. Reines and C.L. Cowan. Detection of the Free Neutrino. *Physical Review*, 92(3): 830–831, 1953. ISSN 0031-899X. doi: 10.1103/PhysRev.92.830. ***Cited on page 19***

[17] The GALLEX Collaboration. GALLEX solar neutrino observations: results for GALLEX IV. *Physics Letters B*, 447:127–133, February 1999. doi: 10.1016/S0370-2693(98)01579-2. ***Cited on page 20***

[18] J. N. Bahcall, S. Basu, and M.H. Pinsonneault. How uncertain are solar neutrino predictions ? *Physics Letters B*, 433(1-2):1–8, August 1998. doi: 10.1016/S0370-2693(98)00657-1. ***Cited on page 20***

[19] The Kamiokande Collaboration . Observation in the Kamiokande-II detector of the neutrino burst from supernova SN1987A . *Physical Review D*, 38(2):448–458, 1988. ISSN 0556-2821. doi: 10.1103/PhysRevD.38.448. ***Cited on page 21***

[20] The NOMAD Collaboration. The NOMAD experiment at the CERN SPS. *Nuclear Instruments and Methods in Physics Research A*, 404(1):96–128, 1998. ISSN 01689002. doi: 10.1016/S0168-9002(97)01079-6. ***Cited on page 21***

[21] F. Vannucci. The NOMAD experiment at CERN. *Advances in High Energy Physics*, 2014: 1–20, 2014. doi: 10.1155/2014/129694. ***Cited on page 21***

[22] C. H. Albright and R. E. Shrock. Signals for tau neutrino interactions in a beam dump experiment. *Physics Letters B*, 84:123–127, June 1979. doi: 10.1016/0370-2693(79)90665-8. ***Cited on pages 22, 40, and 68***

[23] C. H. Albright, R. E. Shrock, and J. Smith. Tests for observing ν_τ interactions in a beam-dump experiment . *Physical Review D*, 20(9):2177–2197, 1979. ISSN 0556-2821. doi: 10.1103/PhysRevD.20.2177. ***Cited on pages 22 and 40***

[24] The NOMAD Collaboration. A search for $\nu_\mu \rightarrow \nu_\tau$ oscillations using the NOMAD detector. *Physical Letters B.*, 431:219–236, July 1998. ***Cited on page 22***

[25] The NOMAD Collaboration. A more sensitive search for $\nu_\mu \rightarrow \nu_\tau$ oscillations in NOMAD. *Physics Letter B*, 453:169–186, 1999. ***Cited on pages 22 and 134***

[26] The NOMAD Collaboration. Updated results from the ν_τ appearance search in NOMAD. *Physics Letters B*, 483:387–404, 2000. doi: 10.1016/S0370-2693(00)00611-0. ***Cited on pages 22 and 134***

[27] The NOMAD Collaboration. Final NOMAD results on $\nu_\mu \rightarrow \nu_\tau$ and $\nu_e \rightarrow \nu_\tau$ oscillations including a new search for ν_τ appearance using hadronic τ decays. *Nuclear Physics B*, 611(1):3–39, 2001. ISSN 05503213. doi: 10.1016/S0550-3213(01)00339-X. ***Cited on pages 22 and 40***

[28] The Daya Bay Collaboration. Observation of electron-antineutrino disappearance at Daya Bay. *Physical Review Letters*, 108(17):171803, 2012. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.108.171803. ***Cited on page 23***

[29] The Double Chooz Collaboration. Improved measurements of the neutrino mixing angle θ_{12} with the Double Chooz detector. *Journal of High Energy Physics*, 2014(10):86, 2014. ISSN 1029-8479. doi: 10.1007/JHEP10(2014)086. ***Cited on page 23***

[30] R.B. Patterson. The NOνA experiment: status and outlook . *Nuclear Physics B - Proceedings Supplements*, 235-236:151–157, 2013. ISSN 09205632. doi: 10.1016/j.nuclphysbps.2013.04.005. ***Cited on page 23***

[31] The NOνA Collaboration. First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOνA . *Physical Review Letters*, 123(15):151803, 2019-10-11. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.123.151803. ***Cited on page 23***

[32] Carlo Giunti and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 2007. ***Cited on pages 25, 26, 27, and 44***

[33] C.S Wu, E. Ambler, R.W Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Physical Review*, 105(4):1413–1415, 1957. ISSN 0031-899X. doi: 10.1103/PhysRev.105.1413. ***Cited on page 25***

[34] R.P. Feynman and M. Gell-Mann. Theory of the Fermi interaction. *Physical Review*, 109 (1):193–198, 1958. ISSN 0031-899X. doi: 10.1103/PhysRev.109.193. ***Cited on page 25***

[35] E. C. G. Sudarshan and R. E. Marshak. Chirality invariance and the universal Fermi interaction. *Physical Review*, 109(5):1860–1862, 1958. ISSN 0031-899X. doi: 10.1103/PhysRev.109.1860.2. ***Cited on page 25***

[36] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Physical Review*, 109(3):1015–1017, 1958. ISSN 0031-899X. doi: 10.1103/PhysRev.109.1015. ***Cited on page 25***

[37] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the K_2^0 meson. *Physical Review Letters*, 13(4):138–140, 1964. ISSN 0031-9007. doi: 10.1103/PhysRevLett.13.138. ***Cited on page 26***

[38] L. Wolfenstein. Neutrino oscillations in matter. *Physical Review D*, 17(9):2369–2374, 1978. ISSN 0556-2821. doi: 10.1103/PhysRevD.17.2369. ***Cited on page 27***

[39] H.A. Bethe. Possible explanation of the solar-neutrino puzzle. *Physical Review Letters*, 56(12):1305–1308, 1986. ISSN 0031-9007. doi: 10.1103/PhysRevLett.56.1305. ***Cited on page 28***

[40] P. Langacker and J. Liu. Standard Model contributions to the neutrino index of refraction in the early universe. *Physical Review D*, 46(10):4140–4160, 1992. ISSN 0556-2821. doi: 10.1103/PhysRevD.46.4140. ***Cited on page 29***

[41] The DUNE Collaboration. Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics. *TBD*, 2020. ***Cited on pages 30, 34, 36, 51, 52, 55, 58, 62, 63, and 103***

[42] The DUNE Collaboration. Long-baseline neutrino oscillation physics potential of the dune experiment. *The European Physical Journal C*, 80(10):978, 2020. doi: 10.1140/epjc/s10052-020-08456-z. ***Cited on pages 30, 52, 56, 58, 65, and 77***

[43] URL `https://www.dunescience.org/`. ***Cited on pages 30 and 31***

[44] The T2K Collaboration. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, April 2020. doi: 10.1038/s41586-020-2177-0. ***Cited on page 30***

[45] The NOνA Collaboration. An Improved Measurement of Neutrino Oscillation Parameters by the NOνA Experiment. *arXiv:2108.08219 [hep-ex]*, August 2021. URL `http://arxiv.org/abs/2108.08219`. arXiv: 2108.08219. ***Cited on page 30***

[46] The DUNE Collaboration. Supernova Neutrino Burst Detection With The Deep Underground Neutrino Experiment . *The European Physical Journal C*, 81(5):423, 2021. ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-021-09166-w. ***Cited on page 30***

[47] URL `http://www.nu-fit.org/`. ***Cited on page 30***

[48] S. K. Agarwalla, S. Prakash, and S. U. Sankar. Resolving the octant of θ_{23} with T2K and NOνA. *Journal of High Energy Physics*, 2013:131, July 2013. ISSN 1029-8479. doi: 10.1007/JHEP07(2013)131. URL `http://arxiv.org/abs/1301.2574`. ***Cited on page 30***

[49] F. Capozzi, E. Di Valentino, E. Lisi, A. Marrone, A. Melchiorri, and A. Palazzo. Global constraints on absolute neutrino masses and their ordering. *Phys.Rev.D*, 95:096014, 2017. doi: 10.1103/PhysRevD.95.096014. ***Cited on page 31***

[50] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou. The fate of hints: updated global analysis of three-flavor neutrino oscillations. *Journal of High Energy Physics*, 2020(9):178, 2020. ISSN 1029-8479. doi: 10.1007/JHEP09(2020)178. ***Cited on page 31***

[51] C. Rubbia. The liquid-argon time projection chamber: A new concept for neutrino physics, 1977. ***Cited on page 32***

[52] W. J. Willis and V. Radeka. Liquid-argon ionization chambers as total-absorption detectors. *Nuclear Instruments and Methods*, 120:221–236, 1974. doi: https://doi.org/10.1016/0029-554X(74)90039-1. ***Cited on pages 32 and 34***

[53] E. Aprile, K. L. Giboni, and C. Rubbia. A study of ionization electrons drifting large distances in liquid and solid argon. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 241:62–71, 1985. ***Cited on page 32***

[54] The ICARUS Collaboration. Performance of a three-ton liquid argon time projection chamber. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 345(2):230–243, 1994. ISSN 01689002. doi: 10.1016/0168-9002(94)90996-2. ***Cited on page 32***

[55] The ICARUS Collaboration. Performance of the ICARUS liquid argon prototype. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 498(1):292–311, 2003. ISSN 01689002. doi: 10.1016/S0168-9002(02)01989-7. ***Cited on page 32***

[56] The ICARUS Collaboration. Design, construction and tests of the ICARUS T600 detector. *Nuclear Instruments and Methods in Physics Research Section A*, 527(3):329–410, 2004. ISSN 01689002. doi: 10.1016/j.nima.2004.02.044. ***Cited on page 32***

[57] The ICARUS-Milano Collaboration. Performance of a liquid argon time projection chamber exposed to the CERN West Area Neutrino Facility neutrino beam. *Physical Review D*, 74(11):112001, 2006. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.74.112001. ***Cited on pages 32 and 62***

[58] The ICARUS Collaboration . Observation of long ionizing tracks with the ICARUS T600 first half-module . *Nuclear Instruments and Methods in Physics Research A*, 508(3):287–294, 2003. ISSN 01689002. doi: 10.1016/S0168-9002(03)01508-0. ***Cited on page 34***

[59] The ICARUS Collaboration. Energy reconstruction of electromagnetic showers from π_0 decays with the ICARUS T600 Liquid Argon TPC. *Acta Physica Polonica B*, 41:103–125, 2010. ***Cited on pages 34 and 62***

[60] The MicroBooNE Collaboration. Reconstruction and measurement of $\mathcal{O}(100)$ mev energy electromagnetic activity from $\pi_0 \to \gamma\gamma$ decays in the MicroBooNE LArTPC. *J. Inst.*, 15(02):P02007, February 2020. doi: 10.1088/1748-0221/15/02/P02007. ***Cited on pages 34, 62, and 98***

[61] The ICARUS Collaboration. Measurement of the muon decay spectrum with the ICARUS Liquid Argon TPC. *Eur. Phys. J. C*, 33(2):233–241, March 2004. doi: 10.1140/epjc/s2004-01597-7. ***Cited on page 34***

[62] The MicroBooNE Collaboration. Michel electron reconstruction using cosmic-ray data from the MicroBooNE LArTPC. *J. Inst.*, 12, September 2017. doi: 10.1088/1748-0221/12/09/P09014. ***Cited on pages 34 and 62***

[63] The ArgoNeuT Collaboration. Demonstration of MeV-scale physics in liquid argon time projection chambers using ArgoNeuT. *Physical Review D*, 99(1):012002, 2019. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.99.012002. ***Cited on page 35***

[64] The DUNE Collaboration. The Single-Phase ProtoDUNE Technical Design Report. 2017. doi: http://arxiv.org/abs/1706.07081. ***Cited on page 35***

[65] The WA105 Collaboration. LBNO-DEMO: Large-scale neutrino detector demonstrators for phased performance assessment in view of a long-baseline oscillation experiment. *arXiv:1409.4405*, 2014. ***Cited on page 35***

[66] L. Zambelli, S. Murphy, and WA105 Collaboration. WA105: A large demonstrator of a Liquid Argon Dual Phase TPC. *Journal of Physics: Conference Series*, 888, 2017. doi: 10.1088/1742-6596/888/1/012202. ***Cited on pages 35 and 36***

[67] B. Aimard, C. Alt, J. Asaadi, et al. A 4 tonne demonstrator for large-scale dual-phase liquid argon time projection chambers. *Journal of Instrumentation*, 13(11):P11003–P11003, 2018. ISSN 1748-0221. doi: 10.1088/1748-0221/13/11/P11003. ***Cited on page 36***

[68] The DUNE Collaboration. First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN neutrino platform. *J. Inst.*, 15(12):P12004, December 2020. doi: 10.1088/1748-0221/15/12/P12004. ***Cited on pages 36, 62, and 123***

[69] The DUNE Collaboration. Volume I: Introduction to DUNE. *Journal of Instrumentation*, 15(8), 2020. ISSN 1748-0221. doi: 10.1088/1748-0221/15/08/T08008. ***Cited on page 36***

[70] The DUNE Collaboration. Volume III: DUNE far detector technical coordination . *Journal of Instrumentation*, 15(8), 2020. doi: 10.1088/1748-0221/15/08/T08009. ***Cited on page 36***

[71] The DUNE Collaboration. Volume IV: The DUNE far detector single-phase technology. *Journal of Instrumentation*, 15(8), 2020. doi: 10.1088/1748-0221/15/08/T08010. ***Cited on page 36***

[72] The DONuT Collaboration. Final tau-neutrino results from the DONuT experiment. *Phys. Rev. D*, 78:052002, September 2008. doi: 10.1103/PhysRevD.78.052002. ***Cited on page 37***

[73] The OPERA Collaboration. Final results of the OPERA experiment on ν_τ appearance in the CNGS neutrino beam. *Phys. Rev. Lett.*, 120(21), May 2018. doi: 10.1103/PhysRevLett.120.211801. ***Cited on pages 37 and 38***

[74] The Super-Kamiokande Collaboration. Measurement of atmospheric neutrino flux consistent with tau neutrino appearance. *Phys. Rev. Lett.*, 97(17):171801, October 2006. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.97.171801. ***Cited on page 37***

[75] The Super-Kamiokande Collaboration. Evidence for the appearance of atmospheric tau neutrinos in super-kamiokande. *Phys. Rev. Lett.*, 110(18):181802, May 2013. doi: 10.1103/PhysRevLett.110.181802. ***Cited on page 37***

[76] The Super-Kamiokande Collaboration. Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande . *Physical Review D*, 98(5):052006, 2018. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.98.052006. ***Cited on pages 37 and 38***

[77] The IceCube Collaboration. Measurement of atmospheric tau neutrino appearance with IceCube DeepCore . *Physical Review D*, 99(3):032007, 2019. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.99.032007. ***Cited on pages 37 and 38***

[78] The OPERA Collaboration. Observation of tau neutrino appearance in the CNGS beam with the OPERA experiment. *Progress of Theoretical and Experimental Physics*, 2014 (10), 2014. ISSN 2050-3911. doi: 10.1093/ptep/ptu132. ***Cited on page 37***

[79] A. de Gouvêa, K. J. Kelly, G. V. Stenico, and P. Pasquini. Physics with beam tau-neutrino appearance at DUNE. *Phys. Rev. D.*, July 2019. doi: 10.1103/PhysRevD.100.016004. ***Cited on pages 37, 38, and 39***

[80] A. Ghoshal, A. Giarnetti, and D. Meloni. On the role of the ν_τ appearance in DUNE in constraining standard neutrino physics and beyond. *Journal of High Energy Physics*, 2019(12):126, 2019. doi: 10.1007/JHEP12(2019)126. ***Cited on pages 37, 38, and 39***

[81] Y. S. Jeong and M. H. Reno. Tau neutrino and antineutrino cross sections. *Phys. Rev. D*, 82(3):033010, August 2010. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.82.033010. ***Cited on page 38***

[82] A. Rashed and A. Datta. Determination of mass hierarchy with $\nu_\mu \to \nu_\tau$ appearance and the effect of nonstandard interactions. *Int. J. Mod. Phys. A*, 32(11):1750060, April 2017. doi: 10.1142/S0217751X17500609. ***Cited on page 39***

[83] The Neutrino-4 Collaboration. Experiment Neutrino-4 search for sterile neutrino and results of measurements. *Physics of Atomic Nuclei*, 83:930–936, 2020. ***Cited on page 39***

[84] P. Machado, H. Schulz, and J. Turner. Tau neutrinos at DUNE: new strategies, new opportunities. *arXiv:2007.00015 [hep-ex, physics:hep-ph]*, June 2020. ***Cited on page 40***

[85] A. Bueno, A. Martínez de la Ossa, S. Navas-Concha, and A. Rubbia. Statistical Pattern Recognition: Application to $\nu_\mu \to \nu_\tau$ oscillation searches based on kinematic Criteria. *J. High Energy Phys.*, 2004(11):014–014, November 2004. ISSN 1029-8479. doi: 10.1088/1126-6708/2004/11/014. ***Cited on page 40***

[86] J. Conrad, A. de Gouvêa, S. Shalgar, and J. Spitz. Atmospheric tau neutrinos in a multikiloton liquid argon detector. *Phys. Rev. D*, 82(9):093012, November 2010. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.82.093012. ***Cited on page 40***

[87] C. Andreopoulos et al. The GENIE neutrino Monte Carlo generator. *Nucl. Instrum. Meth. A*, 614:87–104, 2010. doi: 10.1016/j.nima.2009.12.009. ***Cited on pages 42 and 58***

[88] Francis Halzen and Alan D. Martin. *Quarks and Leptons - An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984. ***Cited on page 44***

[89] The CHORUS Collaboration. Cross-section measurement for quasi-elastic production of charmed baryons in νn interactions. *Physics Letters B*, 575(3-4):198–207, November 2003. doi: 10.1016/j.physletb.2003.09.056. ***Cited on page 45***

[90] D. Rein and L. M. Seghal. Neutrino-excitation of baryon resonances and single pion production. *Annals of Physics*, 133:79–153, 1981. doi: https://doi.org/10.1016/0003-4916(81)90242-6. ***Cited on page 45***

[91] The Particle Data Group. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 2020. ISSN 2050-3911. doi: 10.1093/ptep/ptaa104. ***Cited on page 46***

[92] URL `https://home.fnal.gov/~ljf26/DUNEFluxes/`. ***Cited on pages 50 and 52***

[93] P. Huber, M. Lindner, and W. Winter. Simulation of long-baseline neutrino oscillation experiments with GLoBES. *Computer Physics Communications*, 167(3):195–202, May 2005. ISSN 00104655. doi: 10.1016/j.cpc.2005.01.003. arXiv: hep-ph/0407333. ***Cited on page 53***

[94] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter. New features in the simulation of neutrino oscillation experiments with GLoBES 3.0. *Computer Physics Communications*, 177(5):432–438, 2007. ISSN 00104655. doi: 10.1016/j.cpc.2007.05.004. ***Cited on page 53***

[95] A bunch of DUNE . Experiment simulation configurations approximating DUNE TDR. 2021. doi: 2103.04797. arXiv: 2103.04797. ***Cited on pages 53 and 54***

[96] B. Roe. Matter density versus distance for the neutrino beam from Fermilab to Lead, South Dakota, and comparison of oscillations with a variable and a constant density. *Phys. Rev. D*, 95(11), June 2017. doi: 10.1103/PhysRevD.95.113004. ***Cited on pages 54 and 55***

[97] S. Agostinelli et al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* , 506(3):250–303, 2003. ISSN 01689002. doi: 10.1016/S0168-9002(03)01368-8. ***Cited on page 58***

[98] URL `https://physics.nist.gov/PhysRefData/Star/Text/PSTAR.html`. ***Cited on page 60***

[99] The DUNE Collaboration. Long-baseline neutrino facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 2: the physics program for DUNE at LBNF. *arXiv:1512.06148 [hep-ex, physics:physics]*, jan 2016. ***Cited on pages 62 and 65***

[100] S. Jadach, J. H. Kühn, and Z. Was. TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons. *Computer Physics Communications*, 64:275–299, 1991. doi: https://doi.org/10.1016/0010-4655(91)90038-M. ***Cited on page 64***

[101] The BABAR and BELLE Collaborations. The physics of the B factories. *The European Physical Journal C*, 74, 2014. ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-014-3026-9. ***Cited on page 64***

[102] O. Shekhovtsova. Study of the tau meson decay modes with Monte Carlo generator TAUOLA. status and perspectives. *Nuclear Physics B Proceedings Supplement*, 260:52–55, March 2015. ***Cited on page 64***

[103] The DUNE Collaboration. Neutrino interaction classification with a Convolutional Neural Network in the dune far detector. *Phys. Rev. D*, 102, November 2020. doi: 10.1103/PhysRevD.102.092003. ***Cited on pages 64 and 65***

[104] Donald H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 1972. ISBN 0-521-62196-8. ***Cited on page 102***

[105] I. Nutini. *Study of charged particles interaction processes on Ar in the 0.2 - 2.0 GeV energy range through combined information from ionization free charge and scintillation light* . PhD thesis, Scuola di Scienze Matematiche Fisiche e Naturali , january 2015. URL `http://www.osti.gov/servlets/purl/1221325/`. ***Cited on pages 123 and 143***