Université de Paris
Ecole doctorale Step'Up, ED560
Laboratoire Astroparticule et Cosmologie

Performance studies of the dual-phase Liquid Argon TPC for the DUNE experiment and analysis of the 4-tonne prototype detector data

Par Andrea Scarpelli

Thèse de doctorat de Physique de l'Univers
dirigée par Alessandra Tonazzo et Thomas Patzak
présentée et soutenue publiquement le 30/09/2019

Devant un jury composé de:

PRÉSIDENT	Antoine KOUCHNER Professeur (Université de Paris)
RAPPORTEUR	Cécile JOLLET Maitre des Conférences HDR (Université de Bordeaux)
RAPPORTEUR	Jeffrey HARTNELL Professeur (University of Sussex)
EXAMINATEUR	Thomas MUELLER Chargé de Recherche (CNRS/LLR)
CO-DIRECTEUR DE THÈSE	Thomas PATZAK Professeur (Université de Paris)
DIRECTRICE DE THÈSE	Alessandra TONAZZO Professeur (Université de Paris)

Except where otherwise noted, this is work licensed under
https://creativecommons.org/licenses/by-nc-nd/3.0/fr/

Titre : Étude des performances d'une chambre a projection temporelle phase double à Argon liquide pour l'experience DUNE et analyse des données du prototype 4-tonne

Résumé : Grâce à leur excellente imagerie 3D couplée à des données calorimétriques de haute résolution, les 4 chambres à projection temporelle à argon liquide, de 10 kt chacune, de la future expérience DUNE vont pouvoir mesurer avec une grande précision les paramètres d'oscillation mais également observer pour la première fois la violation CP dans le secteur leptonique. Un ou plusieurs de ces 4 modules serait susceptible d'utiliser la technologique double-phase, s'appuyant sur l'extraction de la charge produite dans le volume d'argon liquide et sa multiplication dans la partie gazeuse, qui permettra d'avoir une meilleures granularité ainsi qu'une meilleure résolution en énergie. Cette thèse présente les stratégies de simulation et reconstruction de particules chargées dans le cadre de la chambre à projection temporelle double-phase de l'expérience DUNE avec sa validation préliminaire à l'aide des données de rayons cosmiques recueillies par le prototype 4 t du CERN.

Mots clefs : Neutrinos, DUNE, LArTPC, simulation, reconstruction

Title : Performances studies of the dual-phase Liquid Argon TPC for the DUNE experiment and analysis of the 4-tonne prototype detector data

Abstract : The 4 10 kt Liquid Argon Time Projection Chambers (LAr-TPCs) of the future DUNE experiment will enable precise measurements of the oscillation parameters and the discovery of CP violation for leptons, thanks to their excellent 3D imaging capabilities coupled with a high resolution calorimeter. One or more modules of the DUNE detector may exploit a dual phase (DP) LAr-TPC that, relying on the extraction of the charge produced in the liquid volume and its subsequent multiplication in argon gas, will increase the expected granularity and energy resolution. This thesis work present the simulation and reconstruction strategies of charged particles in a dual-phase LArTPC in the context of the DUNE experiment and its preliminary validation using cosmic ray data from the CERN 4 t demonstrator.

Keywords : Neutrino, DUNE, LArTPC, Dual-phase, simulation, reconstruction

Resumé de la thèse en Français

Les deux dernières décennies ont approfondi notre compréhension de la nature et du comportement des neutrinos. Après le prix Nobel 2015, décerné à Takaaki Kajita et Arthur McDonald pour *la découverte des oscillations de neutrinos montrant que les neutrinos ont une masse*, une nouvelle génération d'expériences a été proposée pour obtenir des mesures précises des paramètres d'oscillation des neutrinos et pour rechercher preuves de violation de parité et de charge (CPV) dans l'Univers.

L'expérience DUNE (Deep Underground Neutrino Experiment) est un projet ambitieux capable d'atteindre la plupart de ces objectifs avec une précision sans précédent. DUNE a pour but d'étudier les neutrinos et antineutrinos produits avec un large spectre d'énergies sur une distance de 1280 km, ce qui permet d'augmenter la sensibilité aux paramètres d'oscillation et de découvrir d'éventuelles sources de CPV pour le secteur léptonique. DUNE sera en mesure de détecter les neutrinos de toutes les saveurs grâce à 4 chambres à projection temporelles à Argon liquide (LArTPC) d'une masse active de 10 kt chacune, situées profondément sous terre au centre de recherche de Sandford, États-Unis. La technologie LArTPC est bien adaptée à l'étude des interactions des neutrinos, grâce à ses exceptionnelles capacités de suivi millimétrique couplées à un calorimètre totalement homogène. De plus, l'argon est un élément bon marché et abondant, et donc un excellent support pour ce type des détecteurs de taille géant.

Dans une LArTPC, le signal est donné par les électrons produits par l'ionisation des atomes d'Argon, dérivés par un fort champ électrique sur une anode constituée de plans d'électrodes. Les impuretés présentes dans le volume d'argon peuvent capter les électrons qui dérivent et atténuer le rapport signal et bruit. Cela limite la distance maximale de dérive active de la TPC et pose un défi à la faisabilité des détecteurs LArTPC géants. Un autre concept de LArTPC repose sur la multiplication Townsend des électrons dans une mince couche d'Argon gazeux située pour compenser toute atténuation survenue pendant la dérive et découpler le signal du bruit. Comme l'argon est utilisé à l'état liquide et gazeux, cette configuration est souvent appelée double-phase (DP), contrairement à la mono-phase (SP), qui ne déploie l'argon que dans son état liquide. Avec une multiplication d'électrons suffisamment importante, une LArTPC à phase-double permet en principe d'obtenir un volume de dérive de plus de dix mètres en pleine activité. La possibilité de réduire le nombre de canaux de lecture et une meilleure résolution du signal de charge formé sur les électrodes constituent d'autres avantages dans le cas d'un grand détecteur de neutrinos.

La faisabilité des TPC à grande échelle et la caractérisation complète des performances des technologies mono-phase et double-phase pour l'experience DUNE est en cours de validation. Une LArTPC mono-phase avec 700 tonnes de masse totale a été exploité avec succès au CERN pendant l'automne 2018, collectant les events produit d'un faisceau de particules chargées et des rayons cosmiques. Un prototype double-phase, appelé prototype 3x1x1x1m^3, avec 4 tonnes de volume actif et 1 mètre de dérive, a également été opérationnel au CERN entre Juin et novembre 2017 et a recueilli plus de 500 000 events des rayons cosmiques, démontrant la faisabilité d'une LArTPC à phase double. Un nouveau prototype phase double de plus grande taille, ProtoDUNE-DP, avec une dérive de 6 mètres et d'un volume actif de 300 t, vise à confirmer la fiabilité de la technologie dual-phase pour l'experience DUNE pendant l'hiver 2019. Les résultats des prototypes et leur interprétation dans le cadre des performances de DUNE font partie du rapport

technique de l'experience DUNE qui sera rendu public en 2019 pour le design mono phase et en 2020 pour le double-phase.

Les travaux présentés dans cette thèse représentent la première étude sur l'intégration de la conception phase double au sein de DUNE et l'évaluation de son potentiel et de son impact sur l'expérience. La première étape a été le développement d'une simulation pour la TPC à phase double et d'une reconstruction dédiée dans le cadre de LArSoft partagée par d'autres expériences sur l'argon liquide et la conception à phase single de DUNE:

- Pour la simulation, les travaux se sont concentrés sur l'introduction dans le cadre de certaines spécificités de la technologie double-phase, en particulier en ce qui concerne la formation du signal et la simulation du bruit sur la base de données d'entrée. Grâce à la flexibilité du cadre LArSoft aux différentes géométries, le travail effectué est facile à exporter vers d'autres détecteurs double phase.

- Une chaîne de reconstruction presque entièrement automatisée capable de reconstruire les neutrinos en DUNE est maintenant en place. Avec cette thèse, nous avons contribué à l'élaboration d'algorithmes spécifiques d'interface de traitement du signal et de recherche de réponse conçus pour la phase double , avec des algorithmes de reconstruction de niveau supérieur de clustering bidimensionnel et recostrution tridimensionnel.

- Une tentative de séparer localement les gerbes électromagnétiques des traces a été faite à l'aide d'un réseau neuronal convolutionnel capable de classer chaque réponse de l'événement en fonction du contexte dans sa proximité immédiate. La précision de la classification pour chaque classe est supérieure à 92% avec une erreur inférieure à 8%. La robustesse du modèle a été testée par rapport aux variations du niveau de bruit, de l'échelle de charge et du nombre de fils morts et les variations observées dans les scores de classification ont été comprises. Le modèle de classification préparé avec des pions et des protons a été introduit dans toute la chaîne de reconstruction et testé sur des événements de neutrinos, avec des performances globalement satisfaisantes.

Le prototype double phase, exposé aux rayons cosmiques fin 2017, a été le premier DP-LArTPC à l'échelle de la tonne. La contribution importante apportée par cette thèse sur l'évaluation des performances du détecteur provient de l'étude du bruit observé et de l'analyse du gain du détecteur. Il convient de rappeler que l'analyse a été réalisée en utilisant pour la première fois les outils de reconstruction automatisée à deux phases développés pour cette thèse.

- L'analyse du bruit a révélé une contribution significative du bruit cohérent apporté à l'intérieur du détecteur par les câbles de commande lents, en particulier les connecteurs aux sondes de température et à la carte haute tension. Des modèles cohérents ont été identifiés ainsi que les fréquences pertinentes et un filtre de bruit cohérent s'est avéré efficace pour réduire les fluctuations de bruit observées à une valeur très proche des attentes par rapport à la conception.

- L'analyse des performances du détecteur est fortement influencée par les réglages sous-optimaux de la haute tension des 3x1x1x1 m^3 causés par un connecteur défectueux sur la grille d'extraction. Néanmoins, il a été possible d'étudier le gain du détecteur, la résolution énergétique et l'uniformité du gain sur le plan de lecture de la charge. Malheureusement, il n'a pas été possible d'étudier les fluctuations du gain sur une longue période de temps.

- Un balayage de l'extraction et de l'amplification du détecteur a été effectué, montrant que les résultats concordent assez bien avec nos prévisions des effets du détecteur, mais seules des tendances qualitatives peuvent être comparées et le prototype 6x6x6m³ devrait traiter un balayage similaire dans les conditions nominales de haute tension.

Le potentiel de la conception à phase double pour l'analyse des oscillations à longue distance dans DUNE a été étudié pour la première fois. Un point important du test est d'utiliser une simulation complète pour prendre en compte correctement tous les effets de la phase double liés à la formation du signal sur les deux vues de collecte. Les événements Neutrino sont reconstruits à l'aide des outils développés dans cette thèse. Dans ce chapitre, la simulation LArSoft à deux phases a été utilisée pour étudier pour la première fois les neutrinos à faisceau au détecteur de loin à deux phases DUNE. De plus, une reconstruction entièrement automatique des hits et des pistes a été testée pour la première fois sur l'échantillon de neutrinos et appliquée à la reconstruction de l'énergie des neutrinos :

- L'efficacité de reconstruction pour les muons a été trouvée à 95%, similaire à celle observée avec une reconstruction similaire dans la simulation à phase single de DUNE. La qualité de la piste reconstruite est cependant moins bonne qu'en monophasé en raison de la direction des muons qui tend à être parallèle à une des vues de lecture. L'efficacité de reconstruction est plus faible pour les protons et les pions que dans le cas d'une phase unique.

- Une méthode pour la reconstruction énergétique des événements neutrino Monique et électroniques est proposée. La résolution obtenue est de 17% pour les événements de saveur muonique avec un muon contenu, 23% pour les événements avec un muon sortant et 12% pour les événements avec un neutrino électronique. Le biais est toujours inférieur à 1%. Les résultats obtenus sont comparables à ceux obtenus pour la phase unique avec une approche similaire.

Les perspectives possibles de ces travaux sont principalement liées à la reconstruction. Bien que les traces puissent être assez bien reconstruites, les topologies plus complexes nécessitent d'autres algorithmes développés de manière expressive pour le cas biphasé. ProtoDUNE biphasé est actuellement en service au CERN et offrira la possibilité de mieux comprendre la technologie, d'améliorer la simulation et d'explorer des stratégies de reconstruction également avec des données.

Le manuscrit est organisé comme suit. Le chapitre 1 donne un bref aperçu des connaissances actuelles sur les neutrinos et une introduction à la physique de DUNE. Le chapitre 2 illustre le principe de fonctionnement du LArTPC phase single et double, compare les avantages et les limites des deux technologies et décrit la conception proposée pour le LArTPC double phase de DUNE. La contribution originale apportée par cette thèse commence au chapitre 3, où sont présentés les outils logiciels de simulation et de reconstruction d'événements développés spécifiquement pour la mise en place du LArTPC à phase double. Le chapitre 3 décrit le fonctionnement des prototypes 3x1x1x1m³ et l'analyse des performances des détecteurs. Dans le chapitre 5 nous évaluons la qualité de la reconstruction des neutrinos au détecteur loin de DUNE et proposons une première méthode pour la reconstruction de leur énergie, qui peut être utilisée comme entrée pour l'analyse des oscillations des neutrinos.

Table of contents

List of figures v

List of tables xiii

List of acronyms xv

Introduction 1

1 Physics of Neutrino Oscillations 3
1.1 Neutrinos in the Standard Model . . . 3
1.1.1 History of neutrinos . . . 3
1.1.2 The Standard Model of particles . . . 4
1.2 Neutrino oscillations . . . 5
1.2.1 Oscillations Probabilities . . . 5
1.2.2 Mixing matrix and oscillation parameters . . . 7
1.3 LBL accelerator experiments . . . 9
1.3.1 Basic phenomenology . . . 11
1.3.2 Current experiments . . . 12
1.4 Next generation of LBL accelerator experiments . . . 13
1.4.1 The HyperKamiokande experiment . . . 13
1.4.2 The DUNE experiment . . . 14

2 Liquid Argon Time Projection Chamber technology 19
2.1 Liquid Argon Time Projection Chamber, concept of a new detector technology 19
2.2 LArTPC working principle . . . 20
2.2.1 Energy loss by charged particles in LAr . . . 21
2.2.2 Ionization and recombination . . . 23
2.2.3 Scintillation . . . 24
2.2.4 Electron transport . . . 26
2.3 Dual-Phase LArTPC . . . 27
2.3.1 Electrons emission from liquid to gaseous Argon . . . 28
2.3.2 Amplification in Gas Argon . . . 29
2.3.3 Secondary scintillation in gas . . . 31
2.4 Charge readout and signal formation . . . 32

2.5 Advantages and Challenges of LArTPCs for large neutrino detectors 33
2.6 Overview of the dual-phase design for DUNE 36
2.6.1 The charge readout system 37
2.6.2 Data readout system 40
2.7 The DUNE dual-phase prototyping strategy 43
2.7.1 The 3×1×1 m^3 prototype 44
2.7.2 ProtoDUNE-DP (6×6×6 m^3 prototype) 44
2.7.3 Developing software tools 44

3 Dual-phase simulation and reconstruction in LArSoft 45
3.1 LArSoft framework 46
3.2 Detector geometries 47
3.3 Event generation and propagation 48
3.3.1 The neutrino event generator 48
3.3.2 The cosmic ray event generator 49
3.3.3 Particle propagation in LAr and ionisation electron drift 49
3.4 Detector response 50
3.4.1 Signal convolution and digitalization 50
3.4.2 Data-driven noise simulation 51
3.5 Event reconstruction 53
3.5.1 Signal pre-processing 54
3.5.2 Hit finder 56
3.5.3 2D clustering 57
3.5.4 3D matching 60
3.5.5 Pandora 61
3.6 Convolutional Neural Network for track/shower separation 62
3.6.1 CNN basic principles 63
3.6.2 Data preparation 63
3.6.3 Network architecture 66
3.6.4 Training 68
3.6.5 Performances 69
3.6.6 Robustness of the model 70
3.6.7 CNN in the reconstruction chain 73
3.7 Summary and outlook for reconstruction 74

4 Operation and data analysis of the 3×1×1 m^3 dual-phase prototype 77
4.1 Experimental setup 78
4.1.1 CRP arrangement 79
4.1.2 CRP instrumentation 79
4.1.3 Detector operation and collected data 80
4.2 Detector operation stability 81
4.2.1 LAr level, pressure and temperature stability, and electron lifetime . 81
4.2.2 High-voltage stability and maximal field achievable 82
4.2.3 Noise level 83

4.2.4 Detector response with pulsing measurement . . . 86
4.3 Data-driven simulation . . . 87
4.3.1 Simulation of the electrons from the liquid to the gas phase . . . 87
4.3.2 Simulation of transmission and collection in gas . . . 88
4.4 Data reconstruction and selection . . . 88
4.4.1 Event reconstruction . . . 90
4.4.2 MIPs selection . . . 90
4.4.3 Distribution of the selected track and Monte Carlo sample . . . 92
4.5 Results from the $3\times1\times1$ m^3 prototype . . . 93
4.5.1 dQ/ds benchmark . . . 93
4.5.2 Charge sharing between views . . . 94
4.5.3 Effective gain . . . 95
4.5.4 Gain uniformity . . . 96
4.5.5 Impact of systematic uncertainties . . . 97
4.5.6 Field scans . . . 98
4.5.7 Resolution of the detector . . . 101
4.5.8 Comparison with the expected distribution . . . 102
4.6 Summary of the performances of the $3\times1\times1$ m^3 prototype . . . 103

5 Neutrino reconstruction and energy resolution with DUNE-DP 105
5.1 DUNE-DP far detector neutrino samples . . . 105
5.1.1 Far detector geometry . . . 106
5.1.2 Monte Carlo settings . . . 106
5.1.3 Sample phase space and detector acceptance . . . 108
5.2 Track reconstruction and study of the performances . . . 109
5.2.1 Definition of the quality metrics . . . 110
5.2.2 Performances on individual reconstructed tracks in ν_μCC events . . . 111
5.2.3 Global performance on ν_μ events . . . 115
5.2.4 Reconstruction performances with a rotated readout surface . . . 116
5.2.5 Summary of the reconstruction performances and outlook . . . 118
5.3 Shower reconstruction . . . 119
5.4 Energy reconstruction . . . 119
5.4.1 ν_μCC energy reconstruction using track momentum and hadronic energy . . . 120
5.4.2 ν_eCC events . . . 127
5.4.3 Impact of the gain . . . 128
5.4.4 Towards the oscillation analysis: smearing matrices . . . 129
5.4.5 Summary on the energy resolution and comparison with single-phase 129
5.5 Conclusions and outlook . . . 131

Conclusions 133

References 135

List of figures

1.1 Left: Sketch with the various elements involved in the production of a neutrino beam: 1) Target 2) Magnetic *horn* 3) Decay pipe 4) muon filter [36]. Right: Expected flux at the far detector for different *off-axis* configurations [29]. 10

1.2 Oscillation probability as a function of the neutrino energy for a baseline of 300 km (left) and 1300 km (right). The shaded area represents the phase space covered by different value of δ_{CP} in both Normal (red) and Inverted (blue) hierarchies. 12

1.3 Left: Marginal likelihood ratio as function of δ_{CP} for the T2K experiment for both NO and IO. Vertical band shows the allowed 2σ intervals calculated using the Feldman and Cousins method [25]. Right: Significance in at which each value of δ_{CP} is disfavored by NOvA for both NO and IO and upper (solid line) and lower (dotted line) octant of θ_{23} [23]. Note that NOvA reports δ_{CP} in a $[0, 2\pi)$ range, while T2K uses the $[-\pi, \pi)$ range. 13

1.4 Setup of the DUNE experiment [46]. 14

1.5 Left: ν_e appearance spectra for selected ν_eCC-like events assuming 3.5 years of staged exposure (see text) in neutrino mode. Right:ν_μ surviving spectra for selected ν_μCC-like events assuming 3.5 years of staged exposure (see text) in neutrino mode. Both plots assumes normal ordering. Equivalent spectra for the antineutrino case are available in [50]. 15

2.1 Sketches of a single-phase (left) and dual-phase (right) LArTPCs [65, 66]. . 20

2.2 Left: dE/dx as function of the momentum p for muons, protons, and pions in liquid Argon. For muons, it is shown also the *restricted* Bethe-Block formula at T_{cut} equal to 4 and 10 MeV and the most probable value of the Landau-Vavilov distribution (see text). Right: Landau-Vavilov distribution for muons at 1 and 10 GeV and for a detector thickness of 1 and 1.5 cm. . . 22

2.3 Comparison of the recombination factor for the Birks law, and the canonical and modified Box models as function of the density ionization expressed as dE/dx (left) and as function of the drift field (right). 25

2.4 Left: transverse and longitudinal electron cloud size σ as function of the drift distance for a drift field of 0.5 $\mathrm{kV\,cm^{-1}}$ (solid line) and 1 $\mathrm{kV\,cm^{-1}}$ (dotted line). Right: charge attenuation as function of the drift distance for two different concentration of impurities calculated with a drift field of 0.5 $\mathrm{kV\,cm^{-1}}$ (solid line) and 1 $\mathrm{kV\,cm^{-1}}$ (dotted line). 27
2.5 Left: Extraction time of electrons in liquid argon as function of the electric field [83]. Right: extraction efficiency of electrons as function of the applied electric field at $T = 90$ K [84]. 28
2.6 Left: first Townsend coefficient α as function of the electric field for Argon at different value of Temperature around the cryogenic condensation point at 87 K. Right: probability distribution of the final avalanche size n for different values of initial electrons n_0 at a fixed gain of 10. For $n_0 = 20$ the expected distribution is compared with a normal distribution with mean Gn_0 and standard deviation $G\sqrt{n_0}$ (see text). 30
2.7 Left: simulation of the trajectories of the drift electrons in the electric fields produced by the three readout planes of the MicroBooNE detector (yellow) compared with the expected *weighted potentials* of the collection plane (green) [65]. Right: field responses (induced-currents) from various paths of one drifting ionization electrons for the three wire planes of the MicroBooNE detector. U, V, and Y are respectively the two induction and the collection plane. See the reference for more details about the arrangement of the electrodes and the simulation strategy [65]. 32
2.8 A ν_μCC interaction in the ICARUS T600 LArTPC (top). Individual particles well recognizable are: the primary muon (1), a secondary electron cascade from a π^0 decay (2), a $\pi^\pm$ (3), a $\mu + e$ (5) from the decay of K (6), a proton (7), and two not well identified tracks (4) and (8). A close out of the interaction vertex shows the differences between the collection (left) and induction (right) view due to the different nature (monopolar or bipolar) of the signal recorded. [92]. 34
2.9 **Left**: Transverse section of the DUNE single-phase LArTPC argon volume, with the four drifting regions and the anode-cathode arrangement (APA-CPA). **Right**: Open view inside the cryostat and the field cage of the dual-phase DUNE 10 kt far detector module. The location of cathode, anode, and PMTs is also highlighted. SFT chimneys and Slow control feedthroughs are located on top of the detector. The active drift region is 12 meters [56]. 35
2.10 A schematic representation of the charge readout system of a dual-phase LArTPC. The simulated field lines (in dark blue) indicate the path followed by the charge (without diffusion or space charge effects). [53] 37
2.11 The different elements composing the LEM-Anode sandwich: the 50x50 $\mathrm{cm^2}$ full LEM PCB and a zoom on the hole structure (a), the 50x50 $\mathrm{cm^2}$ Anode PCB with a zoom on the strip arrangement (b). A closeout image of the HV lem connectors and suspension pillars (c), and the location of the 32-pins connectors at the border of the LAS (d). [53] 38

2.12 The full 9 m^2 CRP. The Invar structure is drawn in gray and the G10 frame is shown in green. The yellow surface represents the surface covered by the assembly of LAS units. The extraction grid is located beneath the LAS surface and thus not visible here. 40
2.13 Left: Expected Voltage returned by the ASIC as function of the detector input charge. Right: Expected ASIC noise both at cold and room temperature as a function of the detector input capacitance[102]. 41
2.14 Pictures with details of a SFT layout and of a microTCA crate taken during the construction and commissioning of the $3\times1\times1$ m^3 prototype [102]. . . 42
2.15 The two DUNE dual-phase prototypes: the $3\times1\times1$ m^3 prototype (left) [102] and the $6\times6\times6$ m^3 or ProtoDUNE-DP (right) [66]. The two sketches are not in scale one with respect to the other (note the two different drift lengths highlighted with the red arrow). 43

3.1 Description of the LArSoft layered architecture. It is particularly highlighted how the specific LArSoft core algorithms and utilities bridges together the *art* framework, third part products, and experiment specific configuration. . 46
3.2 The dual-phase 10 kt DUNE far detector geometry description in LArSoft. . 48
3.3 Qualitative comparison between the data-driven noise simulation (a) and a real event (b) collected with the $3\times1\times1$ m^3 prototype. The correlation patterns found in data (see text and Chapter 4) is well reproduced by the simulation, along with the noise differences between the 3-meter strips long (left) and the 1-meter strips long (right) readout planes. 52
3.4 The average frequency spectra before and after the use of a coherent noise filter (see Section 3.5.1) for one strip taken from the run used as noise reference for the $3\times1\times1$ m^3 detector. The coherent noise spectra in blue is used for the noise simulation in the $3\times1\times1$ m^3 prototype, while incoherent spectra the red is used for the DUNE far detector. 53
3.5 Schematic example of the different reconstruction stage applied on a cosmic muon track collected with the $3\times1\times1$ m^3 detector. The raw data (a) are first preprocessed in order to remove the noise (b) and identify the signal hits (c). Hits are grouped into 2D clusters (d), and clusters are finally associated to form a 3D image of the event (e) [112]. 54
3.6 A cosmic muon event collected with the $3\times1\times1$ m^3 detector before (left) and after (center) the pre-processing strategy described in Section 3.5.1. Pre-processing impacts minimally the expected signal defined within the ROI (right). [112]. 55
3.7 Results of the hit finder in the dual-phase DUNE LArTPC. The ADC signal is in blue, while the function obtained from the fit is in orange. Three examples are shown: the fit of groups of peaks (a), the result of the refit (b), the hit fitter failure (c). 58

3.8 The two 2D projections of a ν_μCC event at the dual-phase DUNE far detector after the use of PMA for 3D view matching. The hits assigned to the same tridimensional trajectory are colored in the same way on both views. When no match is found, hits are colored in black 59
3.9 The performance of Pandora on a ν_μCC (left) and a ν_eCC (right) event at the dual-phase DUNE far detector. The event on the left was fully reconstructed in 3D, while for the event on the right it was not possible to find any good match between views. 61
3.10 Example of the ADC map of a simulated ν_μCC events. For a human eye, *Track*-like energy deposits are well distinguishable from *shower*-like deposits. 62
3.11 A sketch of a typical CNN architecture for the classification of types of vehicles. The 2D input image is first processed by kernel layers which extrapolate the features necessary for the classification which are successively aggregate by classification layers. [127] 64
3.12 Right: example of the data preparation on a simulated 2 GeV pion interaction. A patch (identified by the white box) is build around a pixel of the event. The label for the patch is given by the true information associated to the central pixel within the patch. Left: an example of patches of the four possible labels. 65
3.13 The network architecture implemented for the track/shower separation in a dual-phase LArTPC. 67
3.14 The evolution of the total *loss* (left) and accuracy (right) as function of the training epochs. The dotted gray line is the difference (Δ) between the score on the training and the validation sets. 69
3.15 The *confusion matrix* calculated on the validation sample of Table 3.1. . . 70
3.16 Example of how random noise (top), charge uncertainties δQ (center), and dead wires (bottom) transform one of the patches from the validation sample. 71
3.17 Impact of random noise (a), charge calibration uncertainties (b) and local density of dead wires (c) on the classification performances of the neural network model expressed in terms of *Purity* and *efficiency*. 73
3.18 Performances of the CNN model on a ν_μCC event in the DUNE-DP detector. Color scale represents the classification probability: 0 is a perfect shower, 1 is a perfect track. The truth case (left) is compared with the CNN output for the single hits (center) and for clusters of hits (right). 74

4.1 Transverse (left) and top (right) sections of the $3\times1\times1$ m^3 TPC. The drift direction is along x. The CRP is composed by 12 50×50 cm^2 LAS, electrically bridged to form an uniform readout surface of 3×1 m^2. The three meters long strips are parallel to the z direction and constitute View 0; the one meters long strips are parallel to y and constitutes View 1. The polar angle ϕ is defined on the $y-z$ plane. [102]. 78

4.2 Top: Pressure (left) and temperature (right) variations over one week of detector operation. The fluctuations of temperature are related to 4 different sensors located at different points on the CRP, 2.6 cm above the LEMs. Bottom: fluctuations of the liquid Argon level for a 12 days period (left) and visual inspection of the flatness of the level using a camera located inside the cryostat (right) [102]. 82
4.3 Left: comparison of the noise RMS collected in both *cold* conditions (at 140 K) and at *room temperature* (near 300 K). Right: Noise RMS at cold measured for different slow control configurations [102]. 84
4.4 The channel maps of the $3\times1\times1$ m^3 detector for both View 0 (left) and View 1 (right). 85
4.5 The evolution of the noise RMS in ADC counts over the entire operations of the $3\times1\times1$ m^3 prototype for both the coherent and incoherent noise. The dotted red lines represent the shut-down periods, while the yellow box identify when the reference run for noise was taken. 87
4.6 Top: Simulated trajectories for electrons in gaseous argon. The scatter plots show the areas where part of the charge is collected (for this specific simulation settings) and thus does not contribute to the signal. Bottom: The electron transmission probability map from the liquid surface to the LEM holes (left) and the collection probability from the hole to the anode (right). [112] 89
4.7 Top: Raw event from *Run 840* showing two cosmic muons candidate. Bottom: Superposition of the waveforms found on the strips between the dotted white lines of the event display above for both View 0 (left) and View 1 (right). Each color denotes the signal collected on a different strip before noise removal. 90
4.8 Left: example of the CBR method applied to one *mip* candidate. The relative charge between the green and the red box is computed along the track trajectory. Right: distribution of the track length and cbr ratio for the reconstructed track in data (top) and in Monte Carlo (bottom) for pions, electrons, and muons. [112] 91
4.9 Distribution of the coordinates of the hits in the selected tracks of run *Run 840* for the drift direction x and the three meters-long projection z (left). Distribution of the direction of the reconstructed tracks of run *Run 840* with respect to the azimuthal (θ) and polar (*phi*) angles, as defined in Figure 4.1 (right). The numerical value associated to each box is the number of reconstructed tracks for that particular bin. 92
4.10 Reconstructed dQ/ds distribution of *Run 840* for both View 0 (left and View 1 (right). The convolution of a Landau with a Gaussian fits the data: the parameters used for the fit are listed on the figure. The reported uncertainty is entirely statistical. 94

4.11 Left: Fitted MPVs for both View 0 and View 1 and asymmetry coefficient (in red) as function of the polar angle ϕ. Right: Effective gain calculated for every LAS. The solid line and the blue bands indicate respectively the average effective gain and $\pm 5\%$ design tolerance (see text). 96
4.12 **Left**: results of the extraction field scan. The extraction probability (blue) is convoluted with the probability of Figure 4.6 (bottom left) to obtain the transmission probability for electrons with an amplification of 28 $\mathrm{kV\,cm^{-1}}$ (yellow). Two different normalization are applied to the data (see text) to match blue (open circles) and yellow (full circles) curves. A χ^2 is calculated to show the relative agreement of data with the two expected probabilities. **Right**: results of the scan of the amplification field. The data collected are normalized to the high-voltage settings of the 3L setup [99] and fitted with an equivalent parameterization (see text). The dotted line shows the best fit of the data, while the gray area includes the contribution of the uncertainties on temperature and pressure. 99
4.13 Left: Resolution of *Run 840* as function of the polar angle ϕ. Right: Resolution as function of the amplification field for the runs of the amplification scan of Table 4.2. 101
4.14 Comparison of the distribution of the charge collected per unit of length in *Run 840* and the equivalent Monte Carlo sample without (left) and with (right) a Gaussian smearing of the charge (see text). 103

5.1 The dual-phase DUNE far detector *workspace* geometry with a simulated ν_μCC event. The upper display shows the transverse drift x-z plane, while the lower display is the y-z readout surface. The y-z is divided into 8 3x3 $\mathrm{cm^2}$ CRP regions separated by a 1 cm thick readout area with no readout, in order to simulate the frame between CRPs. The two readout strips samples respectively the directions along to y (View 0) and z (View 1). The neutrino beam direction is coming parallel to z. In the simulated ν_μCC event shown here, the red dots represent the energy deposits associated to electrons and photons, while the blue dots represent the deposits associated to muons. . . 107
5.2 Left: *inelasticity* distribution for ν_μCC events with respectively contained and not-contained primary muons in both the *workspace* and full detector geometry. The histograms are normalized to the total number of ν_μCC events. Right: distribution of the angle between the particle trajectory and the z axis for primary muons, pions, and protons emitted in ν_μCC events. 109
5.3 A sketch of a track showing bad *Completeness* (left) and bad *Purity* (right). Dotted lines represent MCParticles and solid lines the reconstructed tracks. Different colors indicate different particles. 110
5.4 The distribution *Completeness* (left) and *Purity* (right) for primary muons in ν_μCC interactions 111
5.5 Scatter plots with the *Completeness* calculated for each reconstructed track matched with a primary muons as function of $\cos(\theta_y)$ and the average signal intensity (left) or the residuals of the track length (right). 112

5.6 The reconstruction efficiency of tracks associated to the primary muon as function of the number of hits found in CRP with or without the interaction vertex (left) and of the momentum (right) in ν_μCC events. 113
5.7 The reconstruction efficiency of tracks associated with primary pions and protons as function of the number of reconstructed hits of the track (left). The distribution of the reconstructed hits associated to primary pions and protons on both views (right). 114
5.8 The reconstruction efficiency of tracks associated with primary pions and protons as function of the particle momentum (left). The distribution of the residuals of the reconstructed track associated to protons and pions (right). . 115
5.9 The number of reconstructed particles matched with a primary particle in each event as function of the number of primary true particles in the event for ν_μCC (left) and νNC (right) interactions. 115
5.10 The distribution of *Completeness* (left) and reconstructed length residuals (right) for the reconstructed track matched with the primary muons in ν_μCC events for both the standard and rotated *workspace* geometry. 116
5.11 The distribution of *Completeness* for the reconstructed track matched with the primary pions (left) and protons (right) in ν_μCC events for both the standard and rotated *workspace* geometry. 117
5.12 Distribution of the *Completeness* of each reconstructed track associated to a primary muon as function of the reconstructed length and the number of hits on view 0. The straight vertical lines pattern corresponds to tracks crossing the boundaries of the CRPs. 120
5.13 Calibration of the momentum in ν_μCC events using the reconstructed range (left) and the MCS (right). Fit errors are negligible for this analysis. 122
5.14 Distribution of the fractional residuals of the reconstructed muon momentum using the reconstructed range (left) and the MCS (right). Fit errors are negligible for this analysis 123
5.15 Left: reconstructed hadronic energy in ν_μCC events. Fit errors are negligible for this analysis. Right: fractional residuals for the energy estimation of the hadronic system. 123
5.16 Left: distribution of the fractional residuals between the true energy and the estimated energy in ν_μCC events with a contained muon tracks. Fit errors are negligible for this analysis. Right: estimate of the fractional bias and resolution for bins of true ν_μCC events with a contained muon. 125
5.17 Left: distribution for the fractional residuals between the true energy and the estimated energy in ν_μCC events with an exiting muon tracks. Fit errors are negligible for this analysis. Right: fractional bias and resolution for bins of true ν_μCC events with an exiting muon track. 125
5.18 Left: distribution for the fractional residuals between the true energy and the estimated energy in ν_eCC events. Fit errors are negligible for this analysis . Right: fractional bias and resolution for bins of true ν_eCC events. 127

5.19 Energy resolution of electrons simulated at fixed values of energy as function of the true energy. The samples are generated for different values of gain from 5 to 50. 129
5.20 The smearing matrices for the ν_μCC (left), ν_eCC (right) samples. Entries are not normalized. 130

List of tables

1.1 Most recent results for the global fit of the oscillation parameters in the three flavor paradigm for normal (inverted) hierarchy (see later in text) [29]. . . . 9

2.1 The principal chemical and physical properties of liquid Argon [67] 21
2.2 The nominal voltages, distance, electric field, and tolerance settings, sorted from bottom to top as proposed for the dual-phase DUNE CRP design [53]. 40

3.1 The number of events and the corresponding number of patches used to composed the training, validation, and testing samples. The fraction of classes present in each sample is also reported. Patches with Michel electrons are also counted as showers. 66

4.1 List of resolution and operative range of the relevant devices used during the regular operation of the $3\times1\times1$ m^3 detector 79
4.2 The electric field settings explored during the operations of $3\times1\times1$ m^3 LArTPC. Nominal values are taken from Figure 2.10 and are the settings necessary to reach an absolute detector gain of 20 in a small dual-phase LArTPC with a 10x10 cm^2 readout area [99]. 81
4.3 The parameterization of Equation 4.5 chosen for both the 3L and the $3\times1\times1$ m^3 amplification fit. Values relative to the 3L are taken from [99] and references therein, with the exceptions of $A\rho$ and $B\rho$ which are taken from [87]. For the $3\times1\times1$ m^3 the value of induction field $\mathscr{E}_{\text{ind}}$, extraction field $\mathscr{E}_{\text{ext}}$, temperature and pressure are relative to the running conditions during the scans, l and $\mathscr{T}$ are taken from [99]. k is found thanks to a simulation for the 3L and measured with a fit in the $3\times1\times1$ m^3 (see text). 100

5.1 The reconstruction efficiency for different primary particles in the DUNE dual-phase LArTPC with the standard and rotated *workspace* geometry compared with the results obtained for the single-phase DUNE LArTPC with a comparable reconstruction [141]. 118
5.2 Definition of the sample used for the ν_μCC energy studies, including also the percentage of true particles associated to the longest reconstructed track and the size of the total sample after the quality cut (see text). 121
5.3 Summary of the reconstructed energy resolution and biases for the ν_μCC and ν_eCC samples for both single (SP) and dual-phase (DP). 130

List of acronyms

SM	Standard Model
AMC	Advanced Mezzanine Card
ASIC	Application Specific Integrated Circuit
CBR	Charge Box Ratio
CC	Charge Current
CNN	Convolutional Neural Network
CPV	Charge Parity Violation
CSA	Charge Sensitive Amplifier
DAQ	Data acquisition
DP	Dual-phase
DUNE	Deep Underground Neutrino Experiment
EM	Electromagnetic
ENC	Equivalent Noise Charge
FE	Front-end
FFT	Fast Fourier Transform
FWHM	Full width at half maximum
GAr	Gaseous argon
HK	HyperKamiokande
HV	High-voltage
InvFFT	Inverse Fast Fourier Transform
LAr	Liquid argon

LArSoft Liquid Argon Software

LArTPC Liquid Argon Time Projection Chamber

LAS Charge Readout Plane

LAS LEM Anode Sandwich

LBL Long baseline

LEM Large Electron Multiplier

LVD Landau-Vavilov distribution

MCS Multiple Coulomb Scattering

ML Machine Learning

MO Mass Ordering

NC Neutral Current

PCB Printed Circuit Board

PMA Projection matching algorithm

PMT Photomultiplier Tube

POT Protons on target

ReLU Rectifier Linear Unit

ROI Region of Interest

SFT Signal feedthrough

SGD Stochastic Gradient Descent

SP Single-phase

SSM Standard solar model

TPC Time Projection Chamber

VUV Vacuum Ultraviolet

Introduction

The past two decades have set important milestones in our understanding of the nature and behavior of neutrinos. After the 2015 Nobel Prize, awarded to Takaaki Kajita and Arthur McDonald for the *discovery of neutrino oscillations which shows that neutrinos have mass*, a new generation of experiments has been proposed with the objective to pursue precise measures of the neutrino oscillation parameters and the search of hints of charge-parity violation (CPV) in the Universe.
The Deep Underground Neutrino Experiment is an ambitious project able to address most of these goals with unprecedented precision. DUNE aims to study the neutrinos and antineutrinos produced with a broad spectrum of energies over a distance of 1280 km, which allows to increase the sensitivity to the oscillation parameters and discover eventual sources of CPV for the leptonic sector. DUNE will be able to detect neutrinos of all flavors thanks to 4 Liquid Argon Time Projection chambers (LArTPC) with an active mass of 10 kt each located deep underground at the Sandford Underground Research Facility in the United States. The LArTPC technology is well-suited for the study of the neutrino interactions, thanks to its excellent millimetric tracking capabilities coupled with a fully homogeneous calorimeter. Moreover, Argon is a cheap and abundant element, and therefore an excellent media for giant detectors.
In a LArTPC, the signal is given by the electrons produced by the ionization of the atoms of Argon, drifted by a strong electric field onto an anode made by planes of electrodes. Impurities present in the argon volume may capture the drifting electrons and attenuate the signal-to-noise ratio. This limits the maximal active drift distance of the TPC and pose a challenge to the feasibility of giant LArTPC detectors. An alternative concept of LArTPCs relies on the Townsend multiplication of the electrons in a thin layer of gaseous Argon located to compensate for any attenuation occurred during the drift and decouple the signal from the noise. Since Argon is used in both liquid and gaseous state, this setup is often called dual-phase (DP), in contrast with the single-phase (SP), which deploys Argon only in its liquid state. Provided a sufficiently large electron multiplication, DP-LArTPCs allow in principle for a fully active drift volume of more than ten meters. Other advantages for the purpose of large neutrino detector is the possibility to reduce the number of readout channels, and a better resolution of the charge signal formed on the electrodes.
The feasibility of large scale TPCs and the complete characterization of the performances of both single and dual-phase technology is currently paramount for the DUNE strategy. A large (700 tons total mass) single-phase LArTPC has been successfully operated at CERN in Fall 2018, collecting triggers from both a beam of charged particles and cosmic rays. A

smaller dual-phase prototype, also called $3\times1\times1$ m^3 prototype, with 4 tons of active volume and 1 meter of drift was also operative at CERN between June and November 2017 and has collected more than 500k cosmic ray triggers, demonstrating the feasibility of the dual-phase LArTPC of few tons. A new and larger dual-phase prototype, ProtoDUNE-DP, with 6 meters of drift and 300 t of active volume, is currently being commissioned at CERN and aims to confirm the reliability of the dual-phase technology for neutrino detectors.
The results of the prototypes and their interpretation in the context of the performances of DUNE are part of the experiment Technical Design Report (TDR) which will be made public in 2019 for single-phase and in 2020 for dual-phase.
The work performed for this thesis represents one of the very first studies of the dual-phase concept applied to DUNE. The manuscript is organized as follows. Chapter 1 is a brief overview of the current knowledge about neutrinos and an introduction to the physics of DUNE. Chapter 2 illustrates the working principle of single and dual-phase LArTPC, compares benefits and limits of both technologies, and describes the design proposed for the DUNE dual-phase LArTPC. The original contribution brought by this thesis begins in Chapter 3, where are presented the software tools for the simulation and the reconstruction of events developed specifically for the dual-phase LArTPC setup. Chapter 4 describes the $3\times1\times1$ m^3 prototypes operation and the analysis of the detector performances. In Chapter 5, we asses the quality of the reconstruction of beam neutrinos at the DUNE dual-phase detector and propose a first method for the reconstruction of their energy, which can be used as input for the analysis of the neutrino oscillations.

Chapter 1

Physics of Neutrino Oscillations

This chapter provides the phenomenological and experimental context for this thesis work. Section 1.1 and Section 1.2 contain a brief introduction on the role of neutrinos in the Standard Model of particle physics, the theoretical background behind the neutrino oscillation of flavor and the present knowledge of the parameters of the mixing matrix. Section 1.3 analyzes the current status of the Long Baseline (LBL) accelerator neutrino experiments and Section 1.4 describes the planned future experimental efforts with particular attention to the Deep Underground Neutrino Experiment (DUNE).

1.1 Neutrinos in the Standard Model

1.1.1 History of neutrinos

The existence of the neutrino was postulated by Pauli in 1930 as an effort to explain the continuous spectrum observed in some nuclear decays [1]. The particle he proposed had to be a light neutral fermion, weakly interacting with matter. The name *neutrino* for this new particle was forged by Enrico Fermi in 1933, when he formalized the theory of β-decays [2]. In 1956 Cowans and Reines detected for the first time the antineutrinos emitted in a nuclear reactor using the inverse β-reaction [3]:

$$p + \bar{\nu}_e \rightarrow n + e^+.$$

The detector was a 200 l tank with water and cadmium chlorine segmented with liquid scintillator layers equipped with photomultiplier tubes. Neutrino events were tagged using the time coincidence between the positron scintillation signals and the neutron capture on cadmium. By comparing data with the nuclear reactor on and off, it was possible to provide compelling evidence of the existence of electron antineutrino and measure an interaction cross section of $\approx 1.2^{+0.7}_{-0.4} \times 10^{-43}$ cm^2 compatible with the theorized value of $(1.0 \pm 0.17) \times 10^{-43}$ cm^2 [4].
The muon neutrino neutrino was discovered in 1962 at Brookhaven National Laboratory [5]. A beam of neutrinos was produced from the decay of pions and detected in a spark chamber.

Neutrino charge current interaction were observed with muons in the final states, suggesting the existence of a second family of neutrinos different from those observed by Cowan and Reines.
After the discovery of the τ lepton in 1975 [6], the existence of a third type of neutrino was postulated and finally observed in 2001 by the DONUT experiment at Fermilab [7]. DONUT discovery of the tau neutrino confirmed also the prediction made in the '90s by the LEP experiments at CERN with the study of the decay channels of the Z_0 bosons that set to three the total number of lepton-neutrino families [8]. The observation made at LEP cannot however exclude the existence of a fourth family of neutrinos with masses larger than $m_Z/2 \simeq 45.5$ GeV. Such speculative fourth family of neutrinos has never been observed so far.

1.1.2 The Standard Model of particles

The experimental evidences previously described are interpreted in the context of a large theoretical framework known as the Standard Model of particle physics (*SM*). The *SM* is formalized as a renormalisable gauge field based on the symmetry groups $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. Each group represents a different interaction (or force), mediated by spin 1 bosons which are the generators of each symmetry group: eight massless *gluons* for the strong interaction, one massless *photon* for the electromagnetic, and three massive bosons, two charged $W^{\pm}$ and a neutral Z^0, for the weak one. The other known elementary particles are fermions organized in two types: quarks and leptons. Quarks can interact via all three forces, while leptons undergo only the electromagnetic and/or weak forces. Both quark and leptons exist in three different *generations* or *flavors* characterized by identical properties from the point of view of the forces, but of increased masses. The fermions are introduced in the Standard Model as a representation of the symmetry groups. For the $SU(2)_L \otimes U(1)_Y$ groups, we have the left-handed fields being $SU(2)_L$ doublets, while right-handed ones are $SU(2)_L$ singlets. In particular, for leptons we have that left-handed neutrino are in doublet with the left-handed charged lepton of corresponding flavor. *SM* interaction of neutrinos are driven by the exchange of the $W^{\pm}$ and Z^0 bosons. Interactions via the $W^{\pm}$ are called *charge current* (CC) interactions and involve also a lepton of the same flavor of the interacting neutrinos for charge conservation. *Neutral current* (NC) interactions occurs by the exchange of Z^0 and are not flavor-specific, therefore neutrinos of any flavor can interact with other quarks or leptons. A right-handed neutrino singolet has never been experimentally observed and, according to the *SM*, it would not couple with any of the three forces.
The *SM* provides also a mechanisms for the generation of the observed masses by introducing a new scalar field and breaking the gauge symmetry by choosing an expectation value for its vacuum state. This process in known as *spontaneous symmetry breaking* and it is mediated by a spin 0 Higgs boson. In this framework, fermions are allowed to acquire mass through the coupling of the right-handed singlets with the left-handed doublets via the Yukawa coupling constant, m. For a generic fermion field ψ the form of the mass term of the *SM* Lagrangian is:

$$\mathscr{L}_Y = m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L), \tag{1.1}$$

which explicitly depends on both left and right-handed fermion fields. Being only left-handed, neutrinos are not allowed to obtain mass through the Higgs mechanisms. Despite the impressive phenomenological success of the *SM*, compelling evidences of the existence of massive neutrinos has emerged in the past twenty years, in particular through the observation of the oscillations among neutrino flavors. For this reason, the precise study of the properties of neutrinos constitutes nowadays one of the most promising hints for new type of physics beyond the *SM* predictions.

1.2 Neutrino oscillations

Possible indication of a flavor oscillation in neutrinos propagating from a source to a detection point have been observed since the '60s as an anomaly in the flux of solar neutrinos. In the Standard Solar Model (SSM), electron neutrinos are mainly produced from the proton fission in deuterium with energies below 0.4 GeV. Several experiment such as Homestake [9], GALLEX, SAGE, and SuperKamiokande [10] reported an average flux that was a factor three smaller compared to the prediction. A solution to the solar neutrino anomaly came in 2001 with the SNO experiment [11], which was sensitive to all three flavors. The explicit observation of the CC interactions of electron neutrinos confirmed the solar flux deficit, whilst the rate of NC current interactions (possible for all flavor) was found in agreement with the prediction.
A second confirmation of neutrino flavor oscillation came from the study of atmospheric neutrinos almost in parallel to the solar neutrino anomaly. Atmospheric neutrinos and antineutrinos are mainly produced in pions and muons decays with an expected flavor ratio ν_μ:ν_e of 2:1. Discrepancies in the expected ratio have been observed at Kamiokande as function of the direction and the deficit found in agreement with the hypothesis of a flavor transition from muon to tau neutrinos [12]. Explicit evidence of the tau neutrino appearance was observed by the OPERA collaboration in 2014 [13].

1.2.1 Oscillations Probabilities

The experimental evidences previously described are in agreement with the hypothesis of a neutrino flavor transition during their propagation from a source. The theoretical formalism describing this phenomenon has been developed in 1957 by Bruno Potecorvo [14] in analogy with the $K^0 \leftrightarrow \bar{K}^0$ oscillations. After the discovery of the second neutrino flavor, Maki, Nakagawa and Sakata proposed the possibility for neutrinos to oscillate between different families [15].
The formalism requires the Hamiltonian dictating the free propagation of neutrinos mass eigenstates not to be diagonal in the basis of the neutrino flavor eigenstates, which are a base of the Hamiltonian of the weak interaction. A neutrino of flavor α with momentum p, created in a charged current weak interaction process, can be described as a superposition of $|\nu_k\rangle$ mass states:

$$|\nu_\alpha\rangle = \sum_k U^*_{\alpha k} |\nu_k\rangle , \tag{1.2}$$

with $U^*_{\alpha k}$ corresponding to an element of a rotation mixing-matrix between the neutrino flavor-eigenstate $|\nu_\alpha\rangle$ and the mass-eigenstate $|\nu_k\rangle$. Neutrinos are produced and successively detected as flavor-states, however the propagation between the source and the detector is dictated by the free particle dynamics affecting the mass-eigenstates. In vacuum, the time evolution of the mass eigenstates is described by the Shrödinger Equation as function of the energy E_k of the k-th mass state:

$$|\nu_\alpha(t)\rangle = \sum_k e^{-iE_k t} U^*_{\alpha k} \, |\nu_k\rangle \,, \tag{1.3}$$

Different mass states propagates with different energies causing a change in the relative phase between them. Therefore, in the condition of time coherence between the mass wave packets, a different flavor eigenstate could be detected after a time $t > 0$. The probability for a neutrino emitted with flavor α to be detected with a flavor β is:

$$P(\nu_\alpha \to \nu_\beta) = \langle \nu_\beta | \nu_\alpha(t) \rangle = \sum_{i,k} U^*_{\alpha k} U_{\beta k} U_{\alpha i} U^*_{\beta i} e^{-i(E_k - E_i)t}. \tag{1.4}$$

The energy is approximated at first order, assuming the three mass eigenstates propagate with the same momentum and relativistic energies ($E \simeq p \gg m$):

$$E_k = \sqrt{p^2 + m_k^2} \approx E + m_k^2/2E.$$

The transition probability can be now expressed [1] as a function of the energy, the distance L, and the squared mass-eigenstate splitting $\Delta m_{ki} = m_k^2 - m_i^2$:

$$P(\nu_\alpha \to \nu_\beta) = \sum_{i,k} U^*_{\alpha k} U_{\beta k} U_{\alpha i} U^*_{\beta i} \exp\left(-i \frac{\Delta m^2_{ki} L}{2E}\right). \tag{1.5}$$

$P(\nu_\alpha \to \nu_\beta)$ is called *appearance* probability when $\alpha \neq \beta$, while if $\alpha = \beta$ the probability is called *survival* probability and its complementary *disappearance* probability.
The previous formalism can also be applied also to antineutrinos. The flavor state α is decomposed in a linear combination of mass-eigenstates k:

$$|\bar{\nu}_\alpha\rangle = \sum_k U_{\alpha k} \, |\bar{\nu}_k\rangle \,, \tag{1.6}$$

where $U_{\alpha k}$ is just the hermitian conjugate of the mixing matrix element that appears in Equation 1.2. The transition probability is therefore:

$$P(\bar{\nu}_\alpha \to \bar{\nu}_\beta) = \sum_{i,k} U_{\alpha k} U^*_{\beta k} U^*_{\alpha i} U_{\beta i} \exp\left(-i \frac{\Delta m^2_{ki} L}{2E}\right). \tag{1.7}$$

Interesting considerations come from Equation 1.5 and Equation 1.7:

[1] In natural units

- **Massive neutrinos:** the oscillation formalism is derived on the assumptions that neutrinos have mass eigenstates different from zero. The observation of oscillations is therefore in contrast with the *SM* prediction of a massless neutrino. However, oscillations are not sensitive to the absolute mass eigenstates, but only to the squared mass splitting Δm^2_{ij}.
- **Oscillation phase modulated by L/E:** the amplitude of the oscillation probability does not depend solely on the elements of the mixing matrix but it can be modulated by an oscillatory term with phase $\sim \Delta m^2_{ij} L/4E$. An experiment becomes more sensitive to certain transitions when in proximity of an oscillation node given by $k\pi\Delta m^2_{ij} \sim 4E/L$.
- **CP violation:** if the Charge-Parity transformation is applied to the terms of Equation 1.7, it is not possible to obtain the transition probability for neutrinos of Equation 1.5. Due to the inversed hermitian conjugated terms of the mixing matrix elements U in 1.5 and 1.7, $P(\nu_\alpha \to \nu_\beta)$ differs from $P(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$ in one sign in the imaginary term. The CPV term is found to be:

$$\mathscr{J}_{\mathrm{CP}} = 4 \sum_{k>i} \Im[U^*_{\alpha k} U_{\beta k} U_{\alpha i} U^*_{\beta k}] \sin\left(\frac{\Delta m^2_{ki} L}{2E}\right). \tag{1.8}$$

  CPV may be studied in neutrino experiments which are sensitive to the appearance of a different flavor compared to the one produced at the source. It is interesting to notice that $\mathscr{J}_{\mathrm{CP}}$ depends on the phase $\frac{\Delta m^2_{ki} L}{2E}$, which determines that CP violating effects strengthen in proximity of the oscillation nodes.

Oscillation probability in matter

When neutrinos propagate through matter, the Hamiltonian describing the propagation of the mass-eigenstates has to be modified to account for the interaction between the neutrino states and the particles composing the propagation medium. Since stable ordinary matter is composed by electrons and protons, electron neutrinos are consequently more affected by interaction in matter than the other flavors. This produce one further splitting between the flavor eigenstates and gives rise to the possibility of matter-enhanced oscillations [16]. Matter has also a different effect for neutrinos and antineutrinos. As it is described later, this aspect is of particular relevance for experiments aiming at measuring the CPV asymmetry for leptons.

1.2.2 Mixing matrix and oscillation parameters

Within the three flavor eigenstates (e, μ, τ) of the *SM* and three mass eigenstates (1,2,3), the mixing matrix is analogous to a 3×3 rotation matrix and it is parameterized using three angles θ_{ij} and a complex phase δ. In this parameterization, the total matrix can be further decomposed in a product of three matrices, each of them representing the rotation around a

different angle θ_{ij}:

$$U_{\text{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.9}$$

where $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. Written in this form, the mixing matrix is also called U_{PMNS}[2]. This formalism also assumes that the neutrino is a Dirac particle [3]. The complex phase δ appears in the probability oscillation between flavors only associated to the imaginary part, and appears in the CP asymmetry term of Equation 1.8. It is for this reason regarded as the CP violating term of the lepton mixing, $\delta = \delta_{\text{CP}}$, in analogy with the quark mixing formalism [17].

Both *appearance* and *survival* probabilities are entirely defined by the three oscillations angles, the phase δ_{CP}, and the tree mass differences squared Δm^2_{12}, Δm^2_{13}, and Δm^2_{23}. As explained before, the sensitivity of an experiment to certain oscillation parameters and mass splitting depends on the detectable flavor states and on the factor E/L, therefore certain types of experiments are more sensitive to certain oscillation parameters than others. This is the reason why it is common to combine the results of every type of experiment to compose the entries of U_{PMNS}. The results of the most recent global fit of the oscillation parameters experiments are reported in Table 1.1. The angle θ_{12} and Δm^2_{12} are very well known thanks to solar neutrino experiments (Homestake [9], GALLEX [18], SAGE [19], and SNO [20]) and the reactor experiment KamLAND [21]. θ_{23} and $|\Delta m^2_{32}|$ are accessible via ν_μ ($\bar{\nu}_\mu$) *disappearence* and recent measures are provided by atmospheric neutrinos experiments (mainly IceCube/DeepCore [22]) and accelerator experiments (NOvA [23], MINOS [24], T2K [25]). The value of θ_{13} has been measured from the $\bar{\nu}_e$ disappearance in reactor neutrino experiments (DayaBay [26], RENO [27], and DoubleCHOOZ [28]) and found consistent with later measurements from the accelerator experiment T2K [25].

In the three-flavor paradigm, the PMNS matrix must be unitary in order to ensure the consistency of the probability. Any departure from unitarity would imply the mixing to a fourth neutrino family, opening new perspectives for particle physics. Currently, the main unknowns of the PMNS matrix are:

- **Octant of θ_{23}:** the global fit of θ_{23} using both atmospheric and accelerator experiments deviates from the *maximal mixing* condition at $\theta_{23} = 45°$ with two degenerate solution one in the lower octant ($\theta_{23} < 45°$) and a second in the upper octant ($\theta_{23} > 45°$). No experiment has so far excluded maximal mixing and determined the octant with more than a 5σ significance. The presence of maximal mixing would imply that ν_μ and ν_τ states contributes in equal manner to the mass eigenstates, and may be clue for a new symmetry for the neutrino sector.

- **Mass ordering (MO):** evidences from solar and atmospheric data have shown a distinct hierarchy between the mass splitting with $\Delta m^2_{12} \ll |\Delta m^2_{31}| \sim |\Delta m^2_{32}|$. Current

[2] from Pontecorvo, Maki, Nakagawa, and Sakata.

[3] If the neutrino was a Majorana particle, it would be necessary to include 2 more phases to U_{PMNS}. This two phases are however neglected here since they cancel while computing the transition probability.

Table 1.1 Most recent results for the global fit of the oscillation parameters in the three flavor paradigm for normal (inverted) hierarchy (see later in text) [29].

Parameter	Best fit ($\pm 1\sigma$)
$\sin^2\theta_{12}$	$0.310^{+0.013}_{-0.012}$
$\sin^2\theta_{23}$	$0.580^{+0.017}_{-0.021}$ ($0.584^{+0.016}_{-0.020}$)
$\sin^2\theta_{13}$	$0.02241^{+0.00065}_{-0.00065}$ ($0.02264^{+0.00066}_{-0.00066}$)
Δm^2_{12} [10^{-5} eV ^{2}]	$7.39^{+0.21}_{-0.20}$
Δm^2_{32} [10^{-3} eV ^{2}]	$2.525^{+0.033}_{-0.032}$ ($2.512^{+0.034}_{-0.032}$)
δ_{CP} [°]	215^{+40}_{-29}

experimental results are compatible with two non-equivalent ordering for the neutrino masses: a normal ordering (NO) with $m_1 < m_2 < m_3$ and an inverted ordering (IO) with $m_3 < m_1 < m_2$. In both ordering the smallest absolute mass splitting is always Δm^2_{12}, while the largest one can be either $\Delta m^2_{31} > 0$ for NO or $\Delta m^2_{32} < 0$ for the inverted ordering. The determination of the neutrino hierarchy with more than 5σ significance is not within the reach of the current generation of experiments. Future efforts are driven by atmospheric and accelerator experiments such as ORCA [30], PINGU [31], HyperKamiokande and DUNE (see later in text) in presence of matter and in vacuum by the reactor experiment JUNO [32].

- **Discovery of δ_{CP}:** its discovery and precise measurement is primary in the physics scope of the next generation of neutrino experiment that are discussed in Section 1.4 and it would assess the existence of CPV in the leptonic sector. This discovery as strong cosmological implications: CP symmetry violation is one of the Sakharov's conditions necessary to explain the baryon asymmetry which causes our Universe to be composed by matter and not anti-matter [33]. Analyses of kaons and B-mesons decay resulted in an accurate knowledge of the mixing matrix for the quark sector, however the measured CPV is insufficient to explain the observed baryon asymmetry [34]. Recent theoretical models consistent with the *SM* invoke the generation of the baryon asymmetry via *leptogenesis* calling for a rupture of the lepton number conservation and the existence of large CPV in the leptonic sector [35]. The discovery and measure of this further source CPV is one or the primary objective of the next generation of accelerator neutrino experiments: DUNE and HyperKamiokande.

1.3 LBL accelerator experiments

In an accelerator experiment it is possible to control the type of the neutrino produced, their energy spectrum and the oscillation baseline. This allows to probe the entire three

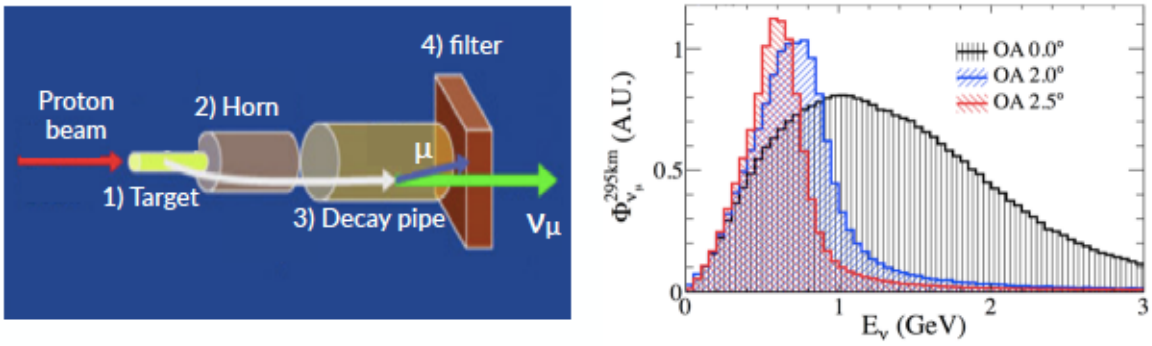


Fig. 1.1 Left: Sketch with the various elements involved in the production of a neutrino beam: 1) Target 2) Magnetic *horn* 3) Decay pipe 4) muon filter [36]. Right: Expected flux at the far detector for different *off-axis* configurations [29].

flavor-paradigm in a single experiment and for this reason are excellent to perform precise measurements of the PMNS parameters and addressing the current unknowns.

The technique used to produce an artificial neutrino beam using a particle accelerator is illustrated in Figure 1.1 (left). High-energy protons are let colliding on a fixed target and the particles produced in the collisions are selected using a device called *horn*, which uses a magnetic field to focus Pions and Kaons downstream from the target and isolate them from the other mesons. The direction of the electrical current traversing the horn allows to select Pions and Kaos of different charge. Then, the particles selected by the horn decay producing mainly muon neutrinos. Other decaying products, in particular secondary muons, are filtered, while the neutrino beam continues its travel to a detector located at certain baseline L from the production point. Selecting the charge of the mesons from the proton collisions allows to produce beam composed mainly composed of either neutrinos or antineutrinos.

The study the properties of the unoscillated neutrino flux is carried out in a *near* detector located very close to the source of the neutrino beam, while the oscillated beam is collected at a *far* detector, located at the required experiment baseline. Thanks to the near detector it is possible to improve the knowledge of the flux normalization and quantify the degree of contamination of the beam with neutrinos of different flavors or types.

The production mechanism using collisions on target allows to produce a flux with neutrinos covering a *wide band* of energies from hundreds of MeV up to several GeV. The average energy of the neutrino beam can be increased by increasing the energy of the primary protons. On the contrary, It is not straightforward to reduce the width of the energy spectrum of the beam. One possible strategy is illustrated in Figure 1.1 (right) and consists of moving the detector *off-axis* with respect to the primary direction of the beam: this naturally selects neutrinos at a certain energy due to the particular kinematics of the decay of the pions. Off-axis beams are used by some experiments, such as NOvA and T2K, to reduce the systematic uncertainties correlated with the neutrino energy and maximize the oscillation probability at their respective baselines. The overall exposure is however reduced, since only a small part of the total flux is used by the experiment.

1.3.1 Basic phenomenology

The baseline of accelerator experiments is in general of the order of several hundreds of km, therefore matter effects must be taken into account in the study of the oscillation. The energy of the neutrino beam lies in general in the GeV region. Replacing the parameterization of the mixing matrix of Equation 1.9 in the flavor transition probability of Equation 1.5, we obtain of the ν_μ disappearance probability:

$$P(\nu_\mu \to \nu_\mu) = 1 - [\cos^4\theta_{13}\sin^2 2\theta_{23} + \sin^2 2\theta_{13}\sin^2\theta_{23}]\sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right) \tag{1.10}$$

The ν_e appearance probability from a ν_μ beam is in general more complicated to workout since it requires adding the contribution of matter effects and can be approximated by [37]:

$$\begin{aligned} P(\nu_\mu \to \nu_e) \approx & \sin^2\theta_{23}\sin^2 2\theta_{13}\frac{\sin^2(\Delta_{31}-aL)}{(\Delta_{31}-aL)^2}\Delta^2_{31} \\ & + \sin 2\theta_{23}\sin 2\theta_{13}\sin 2\theta_{12}\frac{\sin(\Delta_{31}-aL)}{(\Delta_{31}-aL)}\Delta_{31}\frac{\sin(aL)}{(aL)}\Delta_{21}\cos(\Delta_{31}+\delta_{\rm cp}) \\ & + \cos^2\theta_{23}\sin^2 2\theta_{12}\frac{\sin^2(aL)}{(aL)^2}\Delta^2_{21}, \end{aligned} \tag{1.11}$$

where $\Delta_{ij} = \Delta^2_{ij}L/4E$, $a = G_F N_e/\sqrt{2}$, G_F is the Fermi constant, N_e is the number density of electrons in the Earth, L is the baseline in km and E the neutrino energy in GeV. The corresponding probabilities for antineutrinos is the same except that $a \to -a$ and $\delta_{\rm CP} \to$ -$\delta_{\rm CP}$. The first term of Equation 1.11 is the dominant *atmospheric* term dominated by Δm^2_{31} and θ_{23}, the last term is the *solar* term regulated by Δm^2_{21} and θ_{12}, while the middle term represent the interference between the two and contains the *CP-violating* contribution. The CPV asymmetry can be quantified using both neutrino and antineutrino appearance probabilities as follows:

$$\mathscr{A}_{\rm CP} = \frac{P(\nu_\mu \to \nu_e) - P(\bar\nu_\mu \to \bar\nu_e)}{P(\nu_\mu \to \nu_e) + P(\bar\nu_\mu \to \bar\nu_e)}, \tag{1.12}$$

which at first order in three flavor paradigm is expressed as [38]:

$$\mathscr{A}_{\rm CP} \approx \frac{\cos\theta_{23}\sin 2\theta_{12}\sin\delta_{\rm cp}}{\sin\theta_{23}\sin\theta_{13}}\left(\frac{\Delta m^2_{21} L}{4E}\right) + \text{Matter effects.} \tag{1.13}$$

From Equations 1.10 and 1.11 it can be observed that the same transition probability is obtained in certain conditions for different values of Mass Ordering, $\delta_{\rm CP}$ and the octant of θ_{23}. The main reason for this degeneracy is related to matter effects which alters the interaction probability for neutrinos and antineutrinos, causing a phony asymmetry that may be confused with CPV. It has been pointed out [39] that at very long baselines ($>$ 1000 km) it is possible to disentangle the effects of the ν-$\bar\nu$ asymmetry introduced by matter and CPV. This happens because the CP asymmetry is modulated by the phase L/E, while the

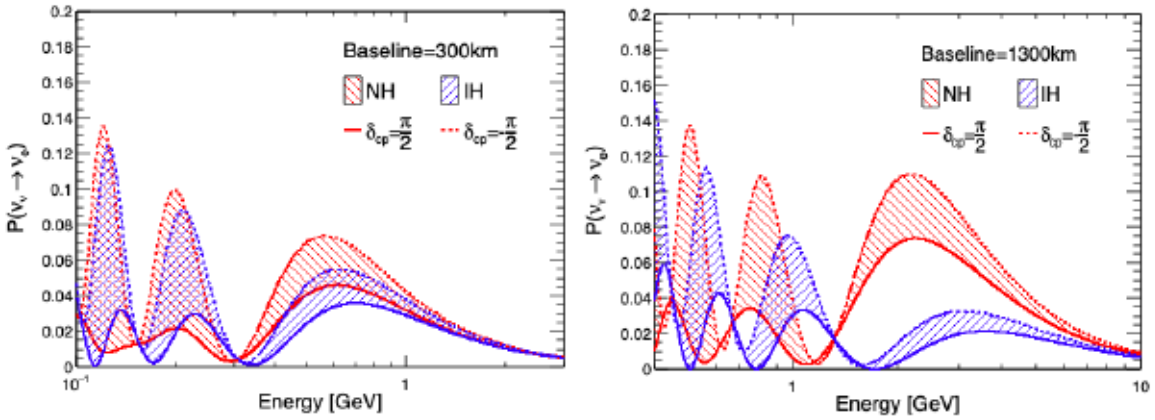


Fig. 1.2 Oscillation probability as a function of the neutrino energy for a baseline of 300 km (left) and 1300 km (right). The shaded area represents the phase space covered by different value of δ_{CP} in both Normal (red) and Inverted (blue) hierarchies.

matter effect grows approximately as $\sim E$. Figure 1.2 (left) and Figure 1.2 (right) show the oscillation probability for a baseline of respectively 300 km and 1300 km. The red bands and the blue bands indicates the effect of all possible values of δ_{CP} for Normal and Inverted Ordering. Overlapping bands represent degenerate values of the oscillations probability for different combination of ordering and δ. It appears evident that an experiment with energy close to the first oscillation maximum and 1300 km of baseline would be sensitive to the Mass Ordering, independently from the value of δ_{CP}.

1.3.2 Current experiments

Two accelerator experiments are currently carrying out the simultaneous measurement of θ_{23}, δ_{cp} and Mass Ordering:

- **T2K:** is an experiment based in Japan which studies the oscillation of a muon neutrino beam produced at J-PARC over a baseline of 292 km. Both near and far detectors are 2.5° off-axis with respect to the beam direction and are sensitive to a very narrow band energy beam with average around 600 MeV. The far detector is the 50 kt Super-kamiokande water Cherenkov detector. Since the beginning of the operations, T2K has collected 1.51×10^{21} protons on target (POT) in neutrino mode and 1.65×10^{21} in antineutrino mode.

- **NOvA:** is a long baseline experiment located in the United States. It studies the oscillation from an off-axis ($\sim 0.84°$) beam with a narrow energy band around 2 GeV from Fermilab over a long baseline of about 810 km. The far detector is a 14 kt tracking calorimeter made up of cells of highly reflective plastic PVC filled with liquid scintillator. The peculiarity of NOvA is that both near and far detector share the same technology. NOvA has collected about 8.85×10^{20} and 6.91×10^{20} POT respectively in neutrino and antineutrino mode.

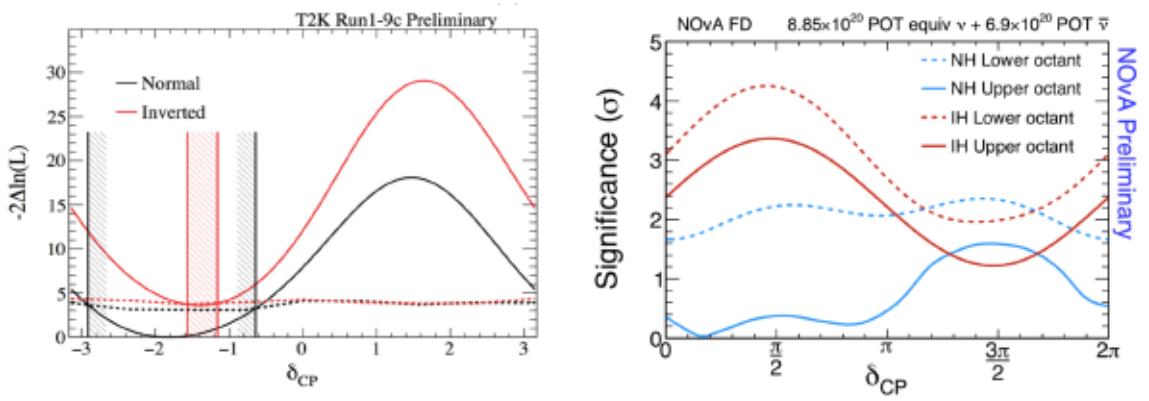


Fig. 1.3 Left: Marginal likelihood ratio as function of δ_{CP} for the T2K experiment for both NO and IO. Vertical band shows the allowed 2σ intervals calculated using the Feldman and Cousins method [25]. Right: Significance in at which each value of δ_{CP} is disfavored by NOvA for both NO and IO and upper (solid line) and lower (dotted line) octant of θ_{23} [23]. Note that NOvA reports δ_{CP} in a $[0, 2\pi)$ range, while T2K uses the $[-\pi, \pi)$ range.

T2K was the fist accelerator experiment to provide evidences of ν_e appearance from a ν_μ beam [40], and in 2016 it gave also the first hint of non-zero δ_{CP} [41]. The latest results of T2K and NOvA have been presented in 2018 and can be found respectively in [25] and [23]. Both experiments agrees on a slight preference for the upper octant of θ_{23} and, in Normal Hierarchy, for a rather large negative value for δ_{CP} (Figure 1.3). A joint T2K-NOvA fit of the experimental results is currently ongoing and planned to be released by 2021.

1.4 Next generation of LBL accelerator experiments

Due to the combination of baseline, energy, and exposure limits, neither T2K nor NOvA are able to independently resolve the current degeneracy between the oscillation parameters. For the next generation of accelerator long baseline neutrino experiments, it is paramount to carry out precise study of the neutrino mixing matrix and unitarity tests. This calls for an experimental setup which can have a sufficient exposure and control over the systematic uncertainties to determine the octant for θ_{23}, resolve the mass ordering, and discovery of δ_{CP}. The currently proposed experiments, HyperKamiokande and DUNE, are both discussed in this section, with particular attention on DUNE.

1.4.1 The HyperKamiokande experiment

HyperKamiokande (HK) [42] is a proposed new Water Cherenkov detector with a total active volume of 740 kt [43]. The experimental setup is identical to the existing SuperKamiokande detector with larger active volume and improved photo-detection system. The oscillations from the J-PARC beam are studied over the same baseline of T2K. Improvements on the

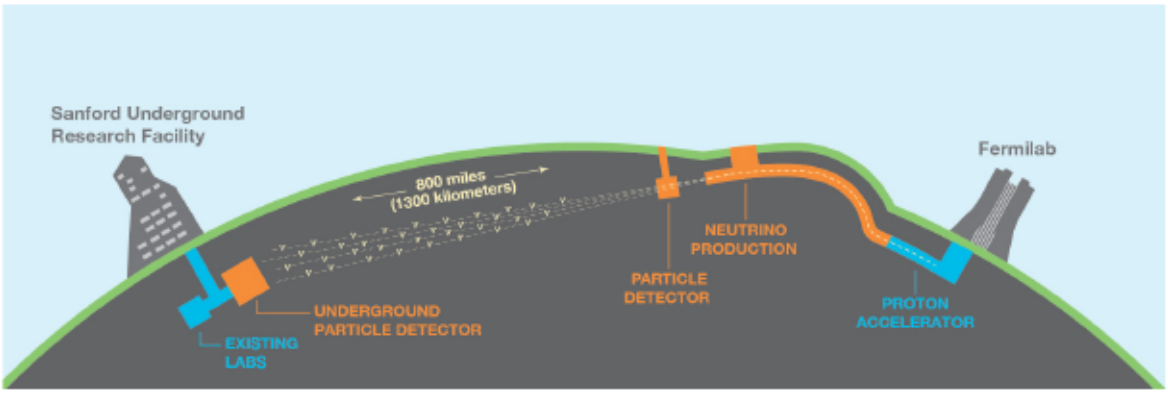


Fig. 1.4 Setup of the DUNE experiment [46].

existing T2K near detector are also foreseen in order to reduce the systematic uncertainties on beam and neutrino cross-section at the level of few % [44].
Thanks to its very large mass, HK can reach a sensitivity to $\delta_{CP} > 5\sigma$ for 56% of the parameters space in less than 10 years. Due to the short baseline of the experiment, the degeneracy between hierarchy and δ_{CP} can be solved only by combining accelerator and atmospheric neutrino measurements. A extension of HK with a second detector in South Korea at a baseline of ∼1300 km has also been proposed [45]. The second HK detector would allow the experiment to access the second oscillation maximum and therefore perform better tests of the three-flavor paradigm.

1.4.2 The DUNE experiment

The Deep Underground Neutrino Experiment (DUNE) is the second new generation long baseline neutrino oscillation experiment aiming at determining the mass hierarchy, discover CP violation for the leptonic sector. Figure 1.4 shows a sketch of the experimental setup of DUNE: an intense (∼1-2 MW) proton beam is used to produce a neutrino/antineutrino beam which travels for about 1285 km underground to the Sandford Underground Research Facility where at about 1.5 km deep underground is located the DUNE far detector. The far detector is composed by 4 Liquid Argon Time Projection Chamber with an active mass of 10 kt each. The characterization of the liquid Argon technology used in the context of DUNE and the study of one of the DUNE far detector designs is the main argument of this thesis and therefore it is not further explained here.
The most important characteristic of DUNE is the wide-band beam between 0 and 20 GeV coupled with a very long baseline. This setup has been optimize in several works [47, 48] and offers the possibility to:

- Resolve the parameter degeneracy with beam neutrinos only, without relying on other experiments or the atmospheric neutrino flux. In fact, as previously pointed out in Figure 1.2, the mass-δ_{CP} degeneracy is resolved at DUNE.
- Carry out a bin-to-bin comparison between the observed and expected oscillation spectra. This is particularly advantageous for the δ_{CP} which could be measured both

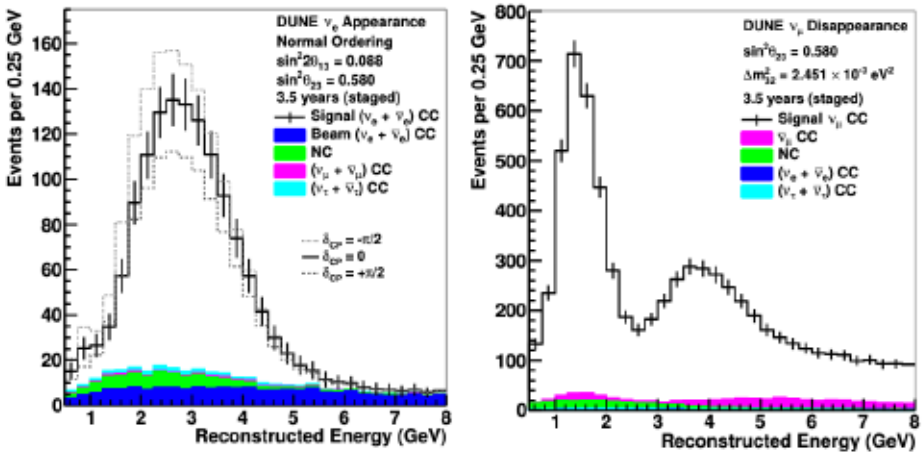


Fig. 1.5 Left: ν_e appearance spectra for selected ν_eCC-like events assuming 3.5 years of staged exposure (see text) in neutrino mode. Right: ν_μ surviving spectra for selected ν_μCC-like events assuming 3.5 years of staged exposure (see text) in neutrino mode. Both plots assumes normal ordering. Equivalent spectra for the antineutrino case are available in [50].

from the oscillation spectrum and from the asymmetry factor of Equation 1.13, obtained from the integral asymmetry of detected neutrinos and antineutrinos. The predicted electron neutrino spectrum at the DUNE far detector after 3.5 year of data taking is shown in Figure 1.5 (left) for different values of δ_{CP}. Furthermore, through the study of ν_μ *disapperance* spectrum (shown on the right of Figure 1.5), DUNE is also particularly sensitive to the value of θ_{23} and Δm^2_{23}. The wide energy beam up to 20 GeV offers also the possibility to observe a small ν_τ sample, which may improve the overall sensitivity of DUNE to the unitaritity tests of the three-flavor paradigm [49].

The first two of the four module of the DUNE far detector will be exposed to the 1.2 MW beam starting from 2026. After three years, the two remaining modules are added reaching the total detector mass of 40 kt. The beam will undergo an upgrade to 2.4 MW after six years. With this *staged* approach, and running for half of the time in neutrino mode and for the other half in antineutrino mode, the collaboration plans to reach:

- Sensitivity $> 5\sigma$ to the MH for 100% of δ_{cp} values in 2 years.
- Sensitivity $> 5\sigma$ to 50% of δ_{cp} values in 11 years.
- Resolution of 20 degrees with $\delta_{cp} = -\pi/2$ in 10 years.
- Sensitivity $> 3\sigma$ to 75% of δ_{cp} values in 14 years.
- Resolution of 0.004 to $\sin^2\theta_{13}$, comparable to reactor experiments, in 17 years.

More details about the primary and ancillary DUNE goals, the tools and methods and strategy developed are described in the available documents prepared by the DUNE collaboration [46,

51–56]. A Technical Design Report of the experiment is under preparation and will be made available in late 2019 [50].

Neutrino interactions at DUNE

The analysis of the oscillations at DUNE is carried out with a fit of the spectra of neutrinos of each flavors, collected in every bin of the energy spectrum. The flavor is reconstructed from the particles produced in the final state after the neutrino interaction. For the study of the oscillation signal at DUNE, only CC interactions are of interest since they allow to determine the flavor of the interacting neutrino. NC processes contributes instead to the background channels. The characteristics of the interactions and the produced in the final state depends also on the energy of the incoming neutrino. For energies $\lesssim 1$ GeV, neutrinos mainly experience a quasi-elastic (QE) scattering from the nucleon. This processes do not destroy the nucleon and produce a clean final state:

$$\nu_\alpha(\overline{\nu}_\alpha) + n(p) \rightarrow p(n) + l_\alpha^-(l_\alpha^+).$$

For intermediate energies between 1-10 GeV, the neutrino can excite the nucleon in an intermediate excited state $N*$ or resonance (RES) Δ which successively decays mostly often producing a lepton, a nucleon and a pion in the final state. Since both CC and NC processes are involved in the interaction, the NC π^0 resonant production is one of the main background to the ν_e *appearance* signal: the two photons produced by the decaying π^0 coverts to an electron-positron pair which mimics the expected electron signal from a ν_e CC interaction. Above 10 GeV, the neutrino has enough energy to penetrate within the nucleon and undergoes a deep-inelastic scattering (DIS) with the quarks. The final states produced in DIS interactions feature large hadronic showers and have in general very complicated topological structures. The picture of the neutrino interaction depicted above ignores possible re-scattering processes that occurs within nuclei with complex structures and may alter the angular and energy distribution, and the multiplicity of the observed final state.

Experimental requirements for DUNE

The ambitious physics goals of DUNE require an excellent control over systematic uncertainties which derive mainly from the neutrino interaction cross-sections, background rejection and calibration of the energy scale. A large exposure is also necessary to achieve a sufficient statistics across the whole energy spectrum within few years. All these requirements demand a detection technology with the following characteristics:

- **Fine particle tracking and identification capabilities:** the classification of the neutrino flavor and the type of interaction is determined by the configuration of the particles in the final state. In particular, the separation between electrons and photons is paramount for the correct determination of the neutrino CC signals and NC backgrounds.

- **Energy resolution over a wide spectrum beam:** the oscillation probability at DUNE is determined also with a bin-to-bin analysis over a wide spectrum. An unbiased and high-resolution determination of the neutrino energy is required in order to have a faithful representation of the correct oscillation probabilities and improve the precision to the oscillation parameters. Excellent resolution is required in particular for ν_e events over a really wide range of energy and for all the possible type of interactions.

- **Active mass of several kt:** a large active mass at the order of several kt is necessary to increase the detector exposure and increase the sensitivity to the oscillation parameters. The interaction medium and the readout channels must therefore be cost-effective for such large scales.

The technologies deployed by past and current accelerator experiments, Cherenkov light detectors and tracking calorimeters, are unable to provide all the required characteristics at the same time. Water-based *Cherenkov light detectors* are easily scalable up to many kt and provide excellent energy resolution and particle identification capabilities in the sub-GeV region. However, above few GeV this technique is not able to resolve with sufficient precision complicated final states produced by RES and DIS interactions, which are dominant in DUNE. On the other hands, *Tracking calorimeters* offer the possibility to reconstruct fine details of the topology of the neutrino interaction and the associated energy deposited for a wide range of energies, but the complexity, cost and large number of readout channels necessary makes this technology not very suitable for very large scale detectors.
Liquid Argon Time Projection Chamber technology has been identified by the DUNE collaboration able to provide both outstanding tracking and calorimetric performances with the possibility, at the same time, to have a sufficiently large active mass. The principal characteristics of a LArTPC, its advantages and limits in the context of DUNE are discussed in the next Chapter.

Chapter 2

Liquid Argon Time Projection Chamber technology

In the previous chapter, we have introduced DUNE, the next generation neutrino experiment which is expected to usher neutrinos towards the precision realm and to provide accurate measurements of the oscillation parameters, and the discovery of δ_{CP}. DUNE will study the oscillation probability at very long baseline with a neutrino beam able to cover several GeV in energy and therefore it must be able to identify and reconstruct with high resolution the neutrino interaction over a broad energy range. It is necessary also to provide a large active mass in order to attain a sufficiently high statistics and sensitivity despite the long baseline. The Liquid Argon Time projection chambers (LArTPCs) technology has been identified as able to provide all the necessary requirements to the physics goals of DUNE. An introduction to the technology and its basic working principle is provided for both single and dual-phase in Sections 2.1, 2.2, 2.3, and 2.4. Section 2.5 presents the main advantages and challenges in building large scale LArTPCs, whilst the dual-phase LArTPC design for DUNE and its prototypes, studied for this thesis, are illustrated in Section 2.6 and Section 2.7.

2.1 Liquid Argon Time Projection Chamber, concept of a new detector technology

The LArTPC is a novel concept of neutrino detector which holds the promise to cope with the requirements of precision and size of the next generation of neutrino experiments. The technology went through several year of successful R&D operations, and its application on the DUNE experiment is currently being studied.
The concept of a LArTPC for neutrino detector dates back to the late '70s where it was first put forward by C. Rubbia in a report [57]. Rubbia conceived an instrument inspired by the noble gas Time Projection Chambers (TPC), designed few yeas before by D. Nygren, that was capable to achieve the same fined-tracking performances of Bubble Chambers, with also high timing resolution and fast recovering after triggering [58]. The idea to use Argon in

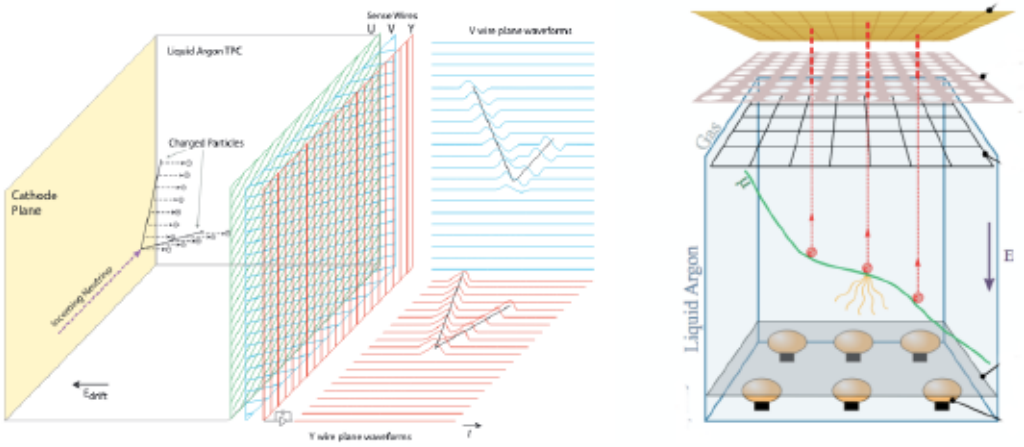


Fig. 2.1 Sketches of a single-phase (left) and dual-phase (right) LArTPCs [65, 66].

liquid state (LAr) came after the work of W. J. Willis and W. Radeka, which first studied its calorimetric properties for particle physics applications [59].
The ICARUS 50L was the first LArTPC to be exposed to a neutrino beam form the CERN west area about thirty years after the original idea was proposed [60]. The ICARUS 50L was followed by ICARUS T600 [61], which pioneered the use of this technology for long baseline neutrino oscillation studies. In the wake of the success of ICARUS, many more experiments have further developed the LArTPC concept such as ArgoNeuT [62], MicroBooNE [63], and the recent ProtoDUNE-SP [64].
The working principle of a basic LArTPC is illustrated in Figure 2.1 (left). Although the intense R&D work over the past decade, the arguments and setup put forward in its original proposal [57] have not changed. At the base of a LArTPC, there is the collection of the ionization electrons deposited by a charge particle crossing a volume of liquid Argon over anode planes composed by conductive wires. The signal current induced by the electron charge on each wire is used to form a tri-dimensional picture of the particle interactions, by matching the images formed on each plane. Scintillation photons are produced along with the ionization electrons and can be used as time origin of the interaction due to the different propagation velocities of light and charge in LAr. It exists also an alternative design for a LArTPC, which consists in multiplying the electron after the drift before the collection on the readout planes. Since it uses Argon in both liquid and gaseous state, this setup, shown in Figure 2.1 (right), is called dual-phase (DP), to be distinguished from the single-phase design of Figure 2.1 (left).

2.2 LArTPC working principle

The use of liquid Argon as detector medium is motivated by several advantages. Its relatively high density of 1.4 $\mathrm{g\,cm^{-3}}$ and low price makes of it a suitable choice for large active mass detectors for rare events searches. Furthermore, being a noble element, Argon offers excellent ionization and scintillation yields and a dielectric constant able to sustain very large electric

Table 2.1 The principal chemical and physical properties of liquid Argon [67]

Parameter	Value
Atomic number Z, atomic weight	18, 39.948 $g\,mol^{-1}$
Boiling point at 1atm	87.3 K
Triple point	83.81 K, 0.689 bar
Liquid density at boiling point	1.389 $g\,cm^{-3}$
Mean excitation energy I	188 eV
Average ionization energy W_{ion}	23.6 eV
Average energy for photon emission W_{ph}	19.55 eV
Average energy loss for *mip*	1,519 MeV $cm^2\,g^{-1}$
Radiation length X_0	19.55 $cm^2\,g^{-1}$
Moliere radius R_M	12.62 $cm^2\,g^{-1}$
Critical energy for electrons E_c	32.84 MeV
Scintillation wavelength	127 nm
Dielectric constant (liquid)	1.5

fields across its volume. The only drawback is the necessity to use cryogenic equipment since Argon maintains its liquid state when below 87 K at atmospheric pressure. The most relevant parameter for the liquid Argon as a detector medium are summarized in Table 2.1.

2.2.1 Energy loss by charged particles in LAr

A relativistic charged particle traveling through a medium loses its energy via single collisions with the electrons and atoms. The average loss per unit of thickness of the material dE/dx, also known as stopping power, is described by the Bethe-Block equation [17]:

$$\frac{dE}{dx} = Kz^2 \frac{Z}{A}\frac{1}{\beta^2}\left[log\left(\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - 2\beta^2 - \delta(\beta\gamma)\right)\right], \tag{2.1}$$

where β is velocity of the particle normalized to the speed of light c, γ is the relativistic Lorentz factor, z is he charge of the particle in unit of the electron charge, $K = 0.307$ MeV g^{-1} cm, and Z/A is the ratio of the Atomic number and mass of the absorber, m_e is the electron mass, T_{max} is the maximal energy which is transferred by a particle of mass M to an electron of mass m_e [1] , I is the mean excitation potential of LAr which can be found in Table 2.1, and $\delta(\beta\gamma)$ is the density effect correction which account for the medium polarization and the relativistic effects once the particle velocity increases. The density term is given by:

$$\delta(\beta\gamma) = \begin{cases} 0 & \text{if } x < x_0 \\ 2(ln10)x - \overline{C} + a(x_1 - x)^k & \text{if } x_0 < x < x_1 \\ 2(ln10)x - \overline{C} & \text{if } x > x_1 \end{cases}, \tag{2.2}$$

[1] Using the Bohr calculation [68] $T_{max} = \frac{2m_e c^2\beta^2\gamma}{1+2\gamma m_e/M+(M_e/M)}$

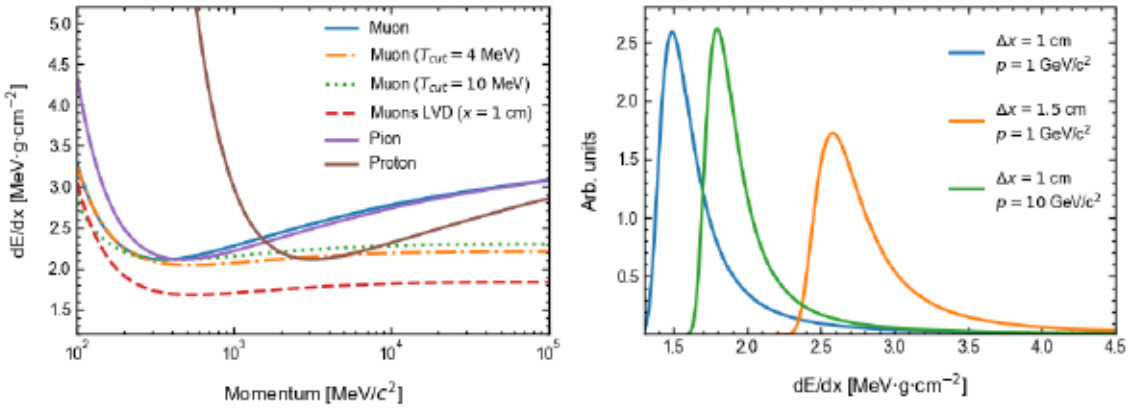


Fig. 2.2 Left: dE/dx as function of the momentum p for muons, protons, and pions in liquid Argon. For muons, it is shown also the *restricted* Bethe-Block formula at T_{cut} equal to 4 and 10 MeV and the most probable value of the Landau-Vavilov distribution (see text). Right: Landau-Vavilov distribution for muons at 1 and 10 GeV and for a detector thickness of 1 and 1.5 cm.

where in the case of liquid Argon we have: $x = log_{10}\beta\gamma$, $x_0 = 0.201$, $x_1 = 3$, $a = 0.196$, $m = 3$, and $\overline{C} = 5.217$. $-\frac{dE}{dx}$ is given in units of MeV g^{-1}cm and its formulation as in Equation 2.3 is valid within the region $0.1 \lesssim \beta\gamma \lesssim 1000$. Large energy transfers to a single electron are possible, but rare. When an electron receives enough energy to further ionize the medium is called δ-ray. In a LArTPC, the high-granularity allows to resolve δ-ray electrons down to a cut-off energy between 2 and 4 MeV. If δ-rays are discarded or not resolved by the detector, a modified (or *restricted*) version of Equation 2.3 is in general proposed, with the maximal transferred energy limited by a threshold value T_{cut} [17]. The *restricted Bethe-Block* formula reaches a plateau, known as *Fermi Plateau* for large value of $\beta\gamma$, which in general depends on T_{cut}. The value of dE/dx as function of the particle momentum p is shown in Figure 2.2 (left) for muons, protons, and pions. For muons it is also shown the effect of T_{cut} at 2 and 4 MeV.

Equation 2.3 and its modifications previously described represent the mean energy loss of a charged particle while crossing the detector medium. The energy loss probability distribution, measured for small detector thickness Δx is described by the Landau-Vavilov distribution (LVD) [69, 70], which is characterized by a prominent peak, followed by a long tail. Figure 2.2 (right) shows different LVD distributions calculated in LAr for muons at 1 and 10 GeV and with different detector thickness. As expected from the previous consideration and a comparison with Figure 2.2 (left), a variation in momentum has no huge impact on the mean value or shape of the distribution, whilst a change in Δx, however, has a much larger effect on the distribution. In a detector with very large thickness the LVD tends to become a normal distribution as consequence of the central limit theorem. The long tail of the LVD describes the large and rare energy transfers which may sometimes occur. It is physically equivalent to the δ-ray emission from the Bethe-Block formula. Applying an upper energy cut T_{cut} to the LVD means to reduce the weight of the events on the tail to the total mean,

therefore moving its value towards the distribution peak. The peak of the Landau-Valivov distribution identifies the most probable energy loss value Δ_p. Its value may be expressed with the relation [17]:

$$\Delta_p = \xi \left[log\left(\frac{2m_e c^2 \beta^2 \gamma^2}{I}\right) + log\left(\frac{\xi}{I}\right) + j - \beta^2 - \delta(\beta\gamma)\right], \quad \text{with: } \xi = K\frac{Z}{A}\frac{x}{\beta^2}, \quad (2.3)$$

where x represent the unit of surface crossed by the particle in unit of $\mathrm{g\,cm^{-2}}$ and $j = 0.2$. Δ_p depends on the detector thickness, unlikely the mean value expressed with the Bethe-Block equation. Above the ionization minimum It is also much less dependent on the energy of the particle compared to the mean of the LVD.
At very high energies, radiative processes such as *bremsstrahlung* photon emission, electron-pair production and photon-nuclear interactions become the most probable processes of energy loss with the medium. The transition, or critical energy, E_c, above which radiative processes become dominant is different for every particles and interaction media: for a muon in liquid Argon, the critical energy is of about 500 GeV, while for electrons it is of few tens of MeV.
Electrons produced from beam neutrinos are usually above their critical energy, thus they can trigger and electromagnetic induced cascade. When an electron above the critical energy interacts with the Argon medium, it loses its energy producing a *bremsstrahlung* photon which, above E_c, can convert into an electron-position pair. The mutual production of photons and electron-positron pairs starts an avalanche of electromagnetic particles which is self-sustained down to several hundreds of keV in LAr.

2.2.2 Ionization and recombination

The energy loss by a charged particles is shared between the excitation of the Argon atoms and the emission of ionization charge in form of electron-ion pairs. The average energy necessary to form an election-ion pair in Argon is $W = 23.6^{+0.5}_{-0.3}$ [71]. The presence of the TPC electric field, $\mathscr{E}$, counteracts the Coulomb attraction of electrons and ions and allows to a fraction $\mathscr{R}$ of ionized electrons to not recombine on the Argon ions and as consequence to be detectable.
The recombination process depends on both the applied electric field and on the local ion density. Electrons can be captured back by their parent ion (*germinate recombination*) and by other charged particles localized in their proximity (*columnar recombination*). Both the *germinate* or *columnar* recombination are accounted in the Jaffé formalism [72], which describes the initial ionized charge distributed as a column around the particle trajectory. Thomas and Imel noted that both the electron diffusion and ion mobility are negligible in liquid Argon during the few ns of the recombination time. By dropping those terms from the Jaffé equation, they found a more compact description of the recombination process, which they called *Box model* [73] :

$$\mathscr{R}_{\text{box}} = \frac{1}{\xi}\ln(\alpha + \xi), \qquad \text{where} \quad \xi = \beta\,(dE/dx)\,/\mathscr{E}, \quad (2.4)$$

with α and β are respectively the recombination coefficient and a fit parameter. The dependence on the charge density is introduced in Equation 2.4 through the dE/dx of the particle. α is commonly assumed to be 1 to have the convergence of $\mathscr{R} \to 1$ for very strong electric fields.
ICARUS explored for the first time recombination as function for various dE/dx and $\mathscr{E}$ in the range of interest of LArTPC experiments (hundreds of $\mathrm{V\,cm^{-1}}$) using stopping particles [74]. A modified version of Birks law [75] valid for liquid scintillators was found to be the best fit for the data:

$$\mathscr{R}_{\mathrm{ICARUS}} = \frac{A_B}{1 + k_B \cdot (dE/dx)/\mathscr{E}}, \tag{2.5}$$

with $A_B = 0.800 \pm 0.003$ and $k_B = 0.0486 \pm 0.006$ $(\mathrm{kV\,cm^{-1}})$ $(\mathrm{g\,cm^{-3}})$ / MeV. Equation 2.5 is purely phenomenological and it holds for a limited range of electrics field, between 0.1 and 5 $\mathrm{kV\,cm^{-1}}$, and of dE/dx between 0.1 and 30 $\mathrm{MeV/(g\,cm^2)}$. The main differences between the Birks and the Box models are related on the assumption that $\mathscr{R} \to 1$ in the limit of $\mathscr{E} \to$ inf.
A recent work from the ArgonNeuT collaboration [76] has put forward the existence of some practical difficulties in applying Equation 2.5 to the calorimetric reconstruction of highly ionizing particle, which are not present if 2.4 is used instead. In a LArTPC, the energy density deposited by a charged particle is calculated from the measured charge density dQ/dx using the relation:

$$dQ/dx = dE/dx \frac{\mathscr{R}}{W_{\mathrm{ion}}}. \tag{2.6}$$

Using the Birks coefficient $\mathscr{R}_{\mathrm{Birks}}$ provided in 2.5 results in:

$$dE/dx = \frac{dQ/dx}{A/W_{\mathrm{ion}} - k_B \cdot (dQ/dx)/\mathscr{E}}, \tag{2.7}$$

which gives spurious results for dE/dx for values of dQ/dx roughly equivalent to 55 $\mathrm{fC\,cm^{-1}}$. The same behavior is not found when $\mathscr{R}_{\mathrm{Box}}$ is used, since the inverse of Equation 2.4 follows an exponential law.
For this reason, the authors introduced a *modified Box model* which improves the similarities between the canonical Box and Birks models removing the assumption of $\alpha = 1$. The best agreement between the two models is achieved for $\alpha = 0.93$ and β at 0.30 cm/MeV. A comparison between the recombination coefficient calculated with the canonical Box model, the Birks approximation, and the modified Box model is shown in Figure 2.3. The modified Box model is set as recombination model for the Monte Carlo simulations described in Chapter 3.

2.2.3 Scintillation

As previously explained, a fraction of the energy shared by a charged particle with the Argon atoms contributed to form Ar^+ ions, as well as excited states atoms $Ar*$. As happens also for other noble gases such as the Xenon, Argon forms excited diatomic molecular states, Ar_2^*, which are called *dimers*. Dimers de-excite under the emission of photons in the Vacuum

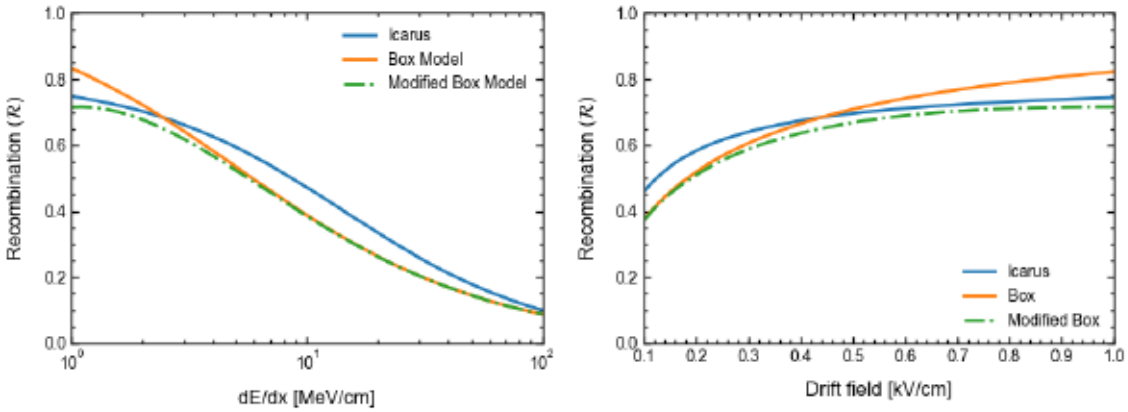


Fig. 2.3 Comparison of the recombination factor for the Birks law, and the canonical and modified Box models as function of the density ionization expressed as dE/dx (left) and as function of the drift field (right).

Ultraviolet (VUV) region [77] via the process:

$$Ar_2^* \rightarrow 2Ar + h\nu.$$

Dimers can be either created with *self trapped luminescence* or with *recobination luminescence*. The former process is realized by the collision between one Argon atom with an excited Argon atom:

$$Ar^* + Ar \rightarrow Ar_2^*,$$

while the latter is based on the recombination of an ionized argon dimer Ar_2^+ with an electron:

$$\begin{aligned} Ar^+ + Ar &\rightarrow Ar_2^+, \\ Ar_2^+ + e &\rightarrow Ar^{**} + Ar \\ Ar^{**} &\rightarrow Ar^* + \text{heat} \\ Ar^* + Ar &\rightarrow Ar_2^*. \end{aligned}$$

The Ar_2^* dimers have populated both the singlet and triped states. This produces two characteristic relaxation times in Argon, one with time constant $\tau_s \approx 6$ ns and a second with time constant $\tau_t \approx 1.6$ µs.
The fraction of the scintillation photons produced and the ion-electron pairs is $N_{ex}/N_i \approx 0.21$ [78]. This can be used to estimate the energy required to produce one photon:

$$W_{ph} = \frac{W_{ion}}{1 + N_{ex}/N_i} \approx 19.5\text{eV}.$$

The energy of the photons is not sufficient to be absorbed via photoelectric effect, therefore Argon has the characteristic to be transparent to its own scintillation light. Since the typical

scintillation times is more than three orders of magnitude smaller than the electron drift time, the photon signal can be used as time origin for the charge events.

2.2.4 Electron transport

The electrons and ions which survive recombination are then drifted under the effect of the TPC field $\mathscr{E}$. Electrons are usually drifted in clouds with a velocity depending on the drift field. The cloud is diffused by random elastic scattering between the electrons and the Argon atoms. A fraction of the electrons may be captured by electronegative impurities in the Argon volume. Those processes are important because they may impact severely on the final charge signal at the detector and the reconstruction of the coordinates of the ionization point.

Drift and Diffusion

Electron are drifted in liquid Argon with a drift velocity v_d, which depends on the component of the drift field directed towards the anode and the Argon temperature T. Also the presence of dopants may enhance v_d [79]. The drift velocity as a function of $\mathscr{E}$ and T is well described by an empirical formula found by Walkowiak [80] valid in the rage 0.5-4.0 kV cm^{-1} and between 87-97 K. This relation was found in good agreement with the data from the ICARUS collaboration taken at the fixed temperature of 89 K and drift fields below 0.5 kV cm^{-1} [81]. Random scattering between the electrons and the Argon atoms deviates the drift velocity from average. The result of this is small spread in time and space of the drifting electron cloud. This process in known as Diffusion. Classical theory for diffusion were developed for gases and follows the The Einstein formula, which describes ratio between the diffusion coefficient and the electron mobility being $D/\mu = kT/e$ where k, T, and e are the Boltzmann constant, the temperature and the electron elementary charge. Einstein formula describes well the data inside gas, but deviates from the measurements in liquid Argon. The most recent measurements of the transverse diffusion coefficient D_T are between 13 and 15 cm^2 s^{-1} [82]. The longitudinal diffusion coefficient D_L has been recently measured to be 3.8 cm^2 s^{-1} for an electric field of 500 kV cm^{-1} [80]. Figure 2.4 (left) shows the expected size of the electron cloud of the order of few mm for drift lengths below 15 meters. This makes possible to realize large LArTPC with drift distances of the order of several meters without significant distortion of the image.

Argon purity

The free movement of charge is granted inside a liquid Argon volume by its electrical neutrality. However, small electronegative contaminants can capture electrons. Due to their abundance in air, oxygen, water, nitrogen, and carbon dioxide are common contaminants present inside an0 Argon TPC. The effect of the impurities over the surviving drift electrons $N_e(t)$ after a drift time t is well described by the differential equation

$$\frac{dN_e(t)}{dt} = -\sum_i k_i[N_i] \cdot N_e(t)$$

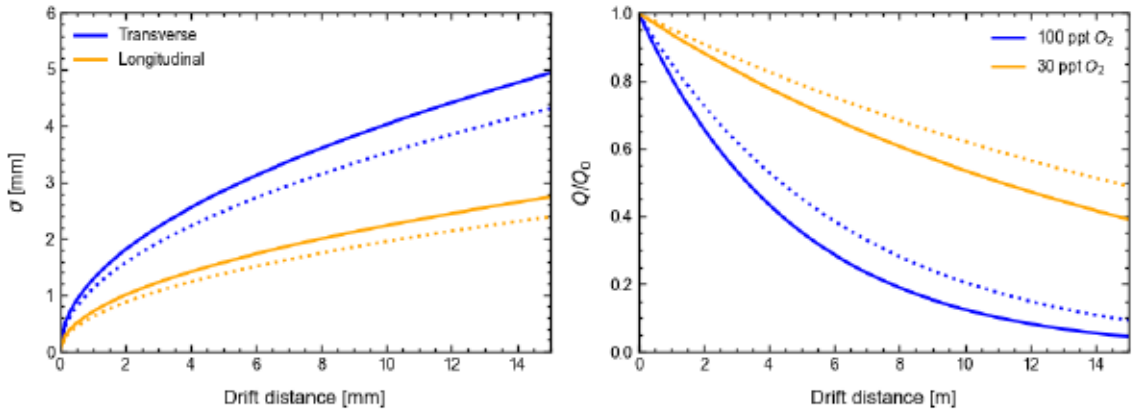


Fig. 2.4 Left: transverse and longitudinal electron cloud size σ as function of the drift distance for a drift field of 0.5 kV cm^{-1} (solid line) and 1 kV cm^{-1} (dotted line). Right: charge attenuation as function of the drift distance for two different concentration of impurities calculated with a drift field of 0.5 kV cm^{-1} (solid line) and 1 kV cm^{-1} (dotted line).

with trivial solution

$$N_e(t) = N_e(0) \cdot e^{-(t-t_0)\sum_i k_i[N_i]}, \tag{2.8}$$

where k_I and $[N_i]$ are the attachment rate and the molar concentration for the different contaminants respectively. t_0 is the ionization, or trigger time, of the charge considered. In general, LArTPC data allows to measure the charge attenuation at the operating directing fields, through the simple expression:

$$Q(t) = Q_0 \cdot e^{-(t-t_0)/\tau_e}, \tag{2.9}$$

where Q_0 and $Q(t)$ represents respectively the charge observed before and after the drift time t. τ_e acts exactly as a free electron lifetime which is accessible without prior knowledge of $\sum_i k_i[N_i]$. Between τ_e and the molar concentration of impurities (expressed in oxygen equivalent) there is an empirical relation:

$$[O_2]_{eq}[ppb] \approx \frac{300\,\mu s}{\tau_e}. \tag{2.10}$$

The electron capture on impurities inside the Argon volume is one of the main obstacle towards the operation of giant LArTPCs. Figure 2.4 (right) shows the attenuation of the charge signal for different values of τ_e and drift field.

2.3 Dual-Phase LArTPC

Dual Phase-LArTPCs rely on the multiplication of the drifted charge in a thin layer of gaseous Argon, which ensures a relatively high signal-to-noise ratio also for drift distances of several meters. The general setup of a DP-LArTPC is shown in Figure 2.1 (right). Although the

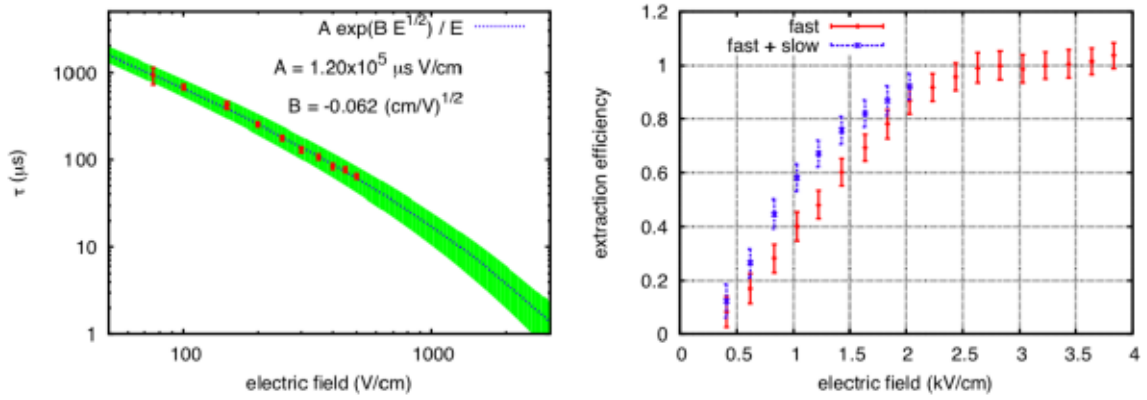


Fig. 2.5 Left: Extraction time of electrons in liquid argon as function of the electric field [83]. Right: extraction efficiency of electrons as function of the applied electric field at $T = 90$ K [84].

charge is drifted along the vertical direction, compared with a single-phase (SP) LArTPC in Figure 2.1 (left), all the processes occurring in LAr are identical to what previously described in Section 2.2. The specific processes which occurs in the gaseous layer and produce the charge amplification and collection on the Anode are discussed in the following part of this section.

2.3.1 Electrons emission from liquid to gaseous Argon

In a dual-phase LArTPC, an electric field $\mathscr{E}$ is used to extract electrons from liquid to gaseous Argon. The phenomenon has been investigated since longtime and it has been observed that the emissions process happens in two steps: a fraction of electrons are emitted *fast* within a time lower than 0.1 µs, while a second component is emitted *slow* with time longer than 0.1 ms at a temperature close to the triple point [83]. The fraction of *slow* electrons is reduced by applying an electric field, and it becomes negligible for fields $\geq 2\,\mathrm{kV\,cm^{-1}}$ [84], as illustrated in Figure 2.5. This behavior is understandable as the result of the combined effects of both a field-enhanced and a thermal emission.
At the interface between the liquid and gaseous Argon exists and energy band $V_0 \simeq 0.2$ eV [85]. In presence of an electric field, $\mathscr{E}$, the potential energy owned by a charge q along the direction x transverse to the surface between the two media can be written as [83]:

$$\begin{aligned} \Phi_l &= -A_l/x - e\mathscr{E}_x - V_0 && \text{in liquid} \\ \Phi_g &= -A_g/x - e\mathscr{E}_x && \text{in gas,} \end{aligned}$$

with:

$$A_v = \frac{e^2}{16\pi\varepsilon_0\varepsilon_v}\frac{\varepsilon_l - \varepsilon_v}{\varepsilon_l + \varepsilon_v},$$

and

$$A_l = A_v \frac{\varepsilon_v}{\varepsilon_l},$$

where ε_0, ε_l and ε_v are respectively the dielectric constants of vacuum, and liquid and gaseous Argon. Thanks to the action of $\mathscr{E}$ the potential barrier is now $V = V_0 - \Delta\Phi$, with:

$$\Delta\Phi = 2e\sqrt{A_l \mathscr{E}_x}(1 + \sqrt{A_l/A_v}).$$

An electron is emitted from the liquid if the condition $p_x^2/2m_e \leq V$ is satisfied, where p_x is the component of the electron momentum parallel to the electric field direction and m_e is the electron mass. In case the total kinetic energy of the electrons $T > V$, but the requirement on p_x is not satisfied, the electron is reflected towards the liquid. Several adiabatic reflections may occur until the momentum component p_x reaches a satisfactory large value. This process lasts less than 0.1 µs. Electrons extracted with this mechanism are defined *fast* or *hot* electrons, since their kinetic energy is larger than the average kinetic energy of electrons at the Argon triple point.
Electrons with kinetic energies $T < V$, are instead emitted according to the thermionic emission [83], whit a characteristic time that depends on the energy band. Those electrons are called *cold* or *slow*. With an electric field larger than 1 $\mathrm{kV\,cm^{-1}}$ the average kinetic energy of the electrons becomes of the order of 0.1 eV. This means that a larger fraction of the electron velocity distribution meets the requirements to be extracted as *fast*. Furthermore, the larger electric field contributes to lower the energy gap event further, making the *slow* emission more efficient.
Given a potential ΔV_{ext} applied between two electrodes, one immersed in liquid and the other one above the surface, it is possible to calculate the effective extraction field $\mathscr{E}_{\mathrm{ext}}$ in liquid:

$$\mathscr{E}_{\mathrm{ext}} = \frac{\Delta V_{\mathrm{ext}}}{d_l\left(1 - \frac{\varepsilon_l}{\varepsilon_g}\right) + D\frac{\varepsilon_l}{\varepsilon_g}}, \tag{2.11}$$

where D is the total distance between the two electrodes and d_l is the distance between the electrode located in liquid and the surface.

2.3.2 Amplification in Gas Argon

The dual-phase principle relies on the amplification of charge in gaseous Argon operated thanks to an intense amplification field. The electrons are accelerated by the field and, when their kinetic energy exceeds the ionization potential of the neutral Argon atoms, *secondary* ionizations may occur. The secondary electrons can further ionize the Argon forming an electron *avalanche*. This process is at the origin of signal amplification techniques in gas chambers.
The first Townsend coefficient (α) describes the number of ion-electron pairs produced per unit of length. α can be expressed following the empirical formulation of Rose and Koroff

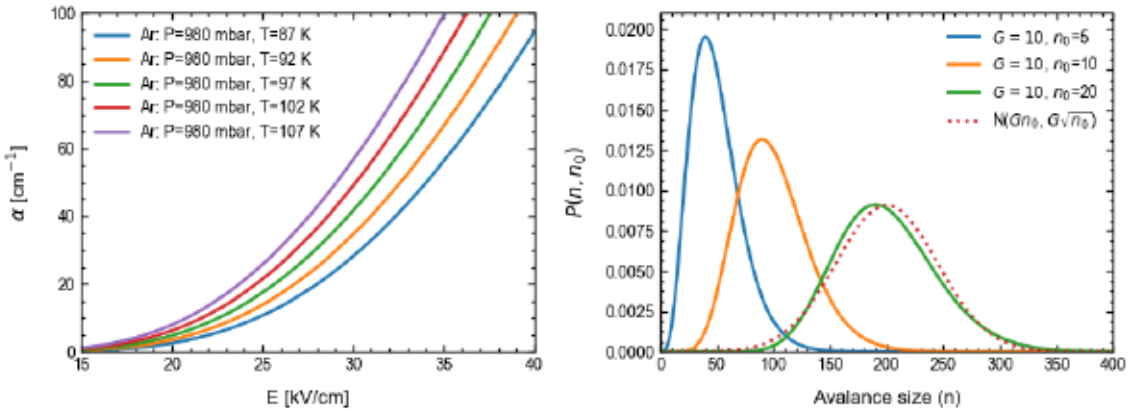


Fig. 2.6 Left: first Townsend coefficient α as function of the electric field for Argon at different value of Temperature around the cryogenic condensation point at 87 K. Right: probability distribution of the final avalanche size n for different values of initial electrons n_0 at a fixed gain of 10. For $n_0 = 20$ the expected distribution is compared with a normal distribution with mean Gn_0 and standard deviation $G\sqrt{n_0}$ (see text).

[86] as function of the *amplification* field, $\mathscr{E}$, and the gas density ρ:

$$\alpha = A\rho\, e^{-\frac{B\rho}{\mathscr{E}}}. \tag{2.12}$$

A and B are phenomenological constants that depends on the type of gas. The first Townsend coefficient for Argon at cryogenic temperature was estimated in a previous work using a MAGBOLTZ simulation ([87] and references therein). Values for $A\rho$ and $B\rho$ are found to be 3576 ± 87 cm^{-2} and 145 ± 1 kV cm^{-1} at 980 mbar and 87 K. The variation of α in gaseous Argon as function of $\mathscr{E}$ for different values of temperature is shown in Figure 2.6 (left). The energy gained by the electrons between two collisions is proportional to the electric field and inversely proportional to the mean free path. For this reason α tends to decrease with an increased gas density. For a given gas, in a condition where $\mathscr{E}$ and ρ are changed, but $\mathscr{E}/\rho$ is constant, the ionization probability per collision (proportional to α/ρ) is unchanged, because, in average, the kinetic energy gained by the electron in between two collisions remains the same.
The number electrons-ion pairs n produced over a certain distance l for the initial number of electrons n_0 is given by:

$$n = n_0\, e^{\int_l \alpha(\mathscr{E}(x))dx}, \tag{2.13}$$

where $\alpha(\mathscr{E}(x))$ represents the variations of the electric field as function of the distance. The gain G is defined as the ration n/n_0. For the case of a uniform electric field, such as the one produced by parallel plates with an infinite surface and gap length equal to x, Equation 2.13 becomes:

$$n = n_0\, e^{\alpha x}, \tag{2.14}$$

and the gain $G = e^{\alpha x}$.

All the aforementioned quantities must be regarded as average values. The Townsend avalanche is in reality a stochastic process since statistical variations occurs in the individual collisions during the avalanche development. The fluctuations of the total number of final electrons produced n determines the intrinsic energy resolution of the detector. The probability to have exactly n electrons after a path x in an avalanche initiated by a single electron is:

$$P(n,x) = (1/G)(1-1/G)^{(n-1)}. \tag{2.15}$$

Equation 2.15 can be seen as the first term in an exponential expansion, thus it becomes for $G \gg 1$:

$$P(n,x) \approx \frac{e^{-n/G}}{G}.$$

This relation is known as Furry's law [88]. For an avalanche starting with n_0 initial electrons, the corresponding size of the avalanche is obtained from the convolution of n_0 Furry's exponentials:

$$P(n,n_0,G) = \frac{1}{G}\left(\frac{n}{G}\right)^{n_0-1}\frac{e^{-n/G}}{(n_0-1)!}. \tag{2.16}$$

In the limit of $n_0 > 10$ the distribution of Equation 2.16 tends to have a Gaussian shape centered around the average gain n_0G and with standard deviation $G\sqrt{n_0}$. The relative width of the distribution is completely independent on the gain and it decreases for larger initial electrons n_0. Figure 2.6 (right) shows an example of the distribution of Equation 2.16 calculated for a fixed value of gain and different number of initial electron n_0. The comparison with a Normal distribution is also shown for $n_0 = 20$.
Electrons have also a not-negligible probability to simply excite the Argon instead of ionizing it. The excited Argon atoms return to the ground state by emitting VUV photons following a similar process to the one described in Section 2.2.3 for the scintillation in liquid. The energy of scintillation photons is not sufficient to further ionize the Argon atoms, however they can extracts electrons through photoelectric interactions with the electrodes. This effect is called *photon feedback* and has the consequence to increase the first Townsend since an electron-ion pair can be generated also from the excited states. The ions produced after the ionization are slowly drifted towards the cathode, where they can further release other electrons. This process is called *ion backflow*, and it is slower compared to the usual avalanche formation. Drifting ions, in particular produced by avalanches of large size, produces also local distortion of the electric fields and, along with the *photon backflow*, they may result in forming a continuous current, or *streamer*. The streamer persists until the amplification device is completely discharged and the fields are back to the working point. Frequent discharges may damage the amplification device, for this reason it is important to find an optimal working configuration of the field at a given pressure and temperature.

2.3.3 Secondary scintillation in gas

Secondary scintillation light is typically produced in dual-phase LArTPC in gas when the energy transferred from the electrons with the Argon molecules is not sufficiently large

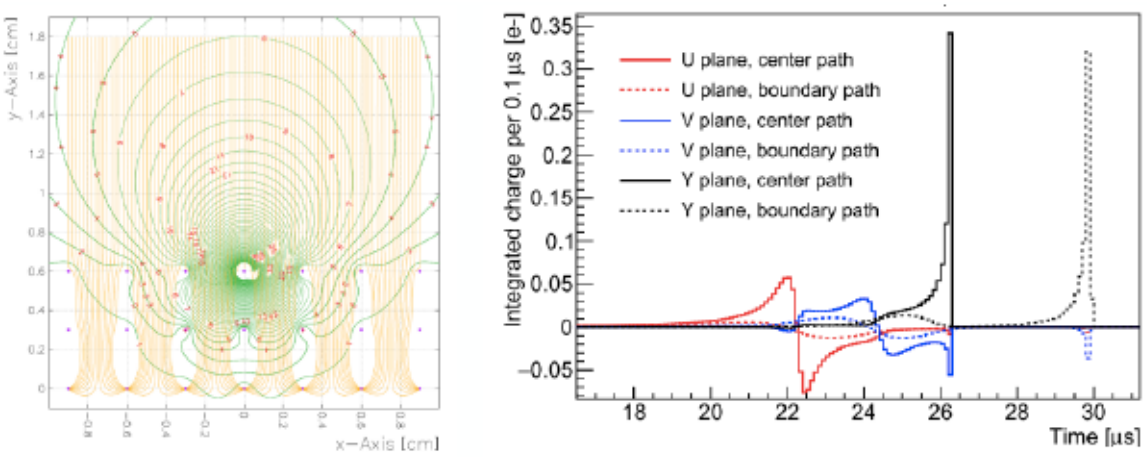


Fig. 2.7 Left: simulation of the trajectories of the drift electrons in the electric fields produced by the three readout planes of the MicroBooNE detector (yellow) compared with the expected *weighted potentials* of the collection plane (green) [65]. Right: field responses (induced-currents) from various paths of one drifting ionization electrons for the three wire planes of the MicroBooNE detector. U, V, and Y are respectively the two induction and the collection plane. See the reference for more details about the arrangement of the electrodes and the simulation strategy [65].

to produce a ionization. This typically happens in the region after the extraction. The extraction field is not sufficiently large to produce a Townsend avalanche, however in the right condition of gas density, the number of secondary photons produced is proportional to the charge extracted. Secondary scintillation light is exploited by some double-phase noble gas experiment for dark matter searches [89, 90].

2.4 Charge readout and signal formation

For both single and dual-phase LArTPC, the drifted charge is used to form a charge signal on a readout anode. The anode in LArTPCs is made of arrays of electrodes (wires) arranged in two or more planes. The passage of the charge q and velocity $\vec{v}$ induces a signal current i on the electrodes according to the Ramo-Shockley theorem [91]:

$$i = -q\vec{\mathscr{E}}_w \cdot \vec{v},$$

where $\vec{\mathscr{E}}_w$ is the so-called *weighting field* which depends on the geometry of the electrode and can be computed neglecting the free charges and setting the *weighted potentials*, V_w on the interested electrode to 1 and on the others to 0. Since the *weighted potentials* is a dimensionless quantity the *weighting field* has the dimension of the inverse of a length. In general, the effect of the induced signal is better understood through the integral of i:

$$Q = -q\left(V_w^{\text{end}} - V_w^{\text{start}}\right).$$

The total induced charge on the electrode Q is a fraction of the drifted charge q and depends on the difference of the weighted potential between the start and the end of the drift. According to the electrode arrangement and biases, the configuration of the *weighted potentials* is different and may produce different signal currents. Figure 2.7 (left) shows an example of both the electron trajectories following the drifting field lines and the simulated lines of the weighted potentials for the electrodes.
In a standard LArTPC, a minimum of two planes of electrodes is necessary to form an independent signal for each of the two coordinates perpendicular to the drift direction. We refer to the plane where the induced charge Q is collected as *collection plane*. On the other planes the collected charge $Q = 0$ and the signal current has a bipolar shape, since the drifting charge first approaches the electron and then it drifts away from it. Those planes are called *induction plane*. Single-phase LArTPCs in general features one or two induction and one collection planes. Figure 2.7 (right) shows for the MicroBooNE detector [65] the simulated signal current induced by a drifting charge on both the induction planes (called U and V in the figure) and the collection plane Y. The bipolar shape of the two induction signals are well noticeable in contrast to the collection signal. In general, the bipolar nature of the signal is not an ideal behavior when the charge distribution of the drifting charge extends in space along the drift direction. In those cases it may happen that, while part of the charge is drifting away, a similar amount is approaching, canceling therefore the signal previously produced. This issue does not occur in dual-phase LArTPCs, since charge amplification allows to use two collection planes.

2.5 Advantages and Challenges of LArTPCs for large neutrino detectors

LArTPC have been proposed for next generation of neutrino detectors. The main advantages of the technology relies on the possibility to build detectors with a kt active mass, as it is possible with Water Cherenkov, and to reconstruct neutrino events with great precision in the multi-GeV range, as it is usually done with segmented tracking calorimeters.
The ICARUS LArTPC constitutes a proof-of-principle of the processes behind the charge production, drift, and signal formation. It was also the first LArTPC detector ever exposed to a neutrino beam and it confirmed the expected performances of a LArTPC for the identification and reconstruction of the charged particles emitted from a neutrino interaction in the sub-GeV and GeV region [93]. Figure 2.8 (top) shows one ν_μCC event collected with the ICARUS detector. The sub-products of the neutrino interaction are easy recognizable and dedicated algorithms (operated with human supervision) are used to identify the various particles and reconstruct their momentum. One great advantage is the possibility to provide an excellent identification of electrons, thanks to the peculiar topology of the electromagnetic shower. Electrons can be separated from γ produced by the π^0 decays thanks to the study of the ionization at the beginning of the shower, which is in general twice larger for γ due to the electron-positron pair produced after the photon conversion. In some bi-dimensional projection of the event it is sometimes possible to even distinguish a small photon conversion

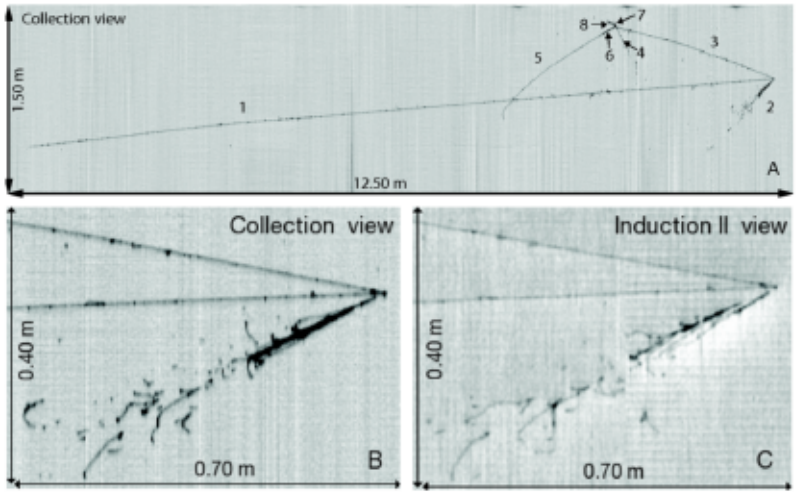


Fig. 2.8 A ν_μCC interaction in the ICARUS T600 LArTPC (top). Individual particles well recognizable are: the primary muon (1), a secondary electron cascade from a π^0 decay (2), a $\pi^\pm$ (3), a $\mu + e$ (5) from the decay of K (6), a proton (7), and two not well identified tracks (4) and (8). A close out of the interaction vertex shows the differences between the collection (left) and induction (right) view due to the different nature (monopolar or bipolar) of the signal recorded. [92].

gap between the neutrino vertex and the shower. The electromagnetic energy resolution is also measured with great precision. ICARUS provides an estimation of $\sigma(E)/E = 0.03\sqrt{E} \oplus 0.01$ for electron cascades the sub-GeV range using the $\pi^0 \to \gamma\gamma$ invariant mass [94].
Although the success of ICARUS, some challenges remain to safely scale the technology at the kilo-ton level demanded by the future generations of neutrino detectors:

1) The electron lifetime in liquid Argon represents a significative limit to the possible maximal drift distances. It can be computed that with an electron lifetime of 3 ms, a minimum ionizing particle, *mip*, in LAr is visible with a signal-to-noise ratio of 9:1 after a three meter drift. In absence of charge amplification, for an equivalent result with an drift distances of 12 meters (such as expected in giant LArTPCs), it is necessary a lifetime better than 8 ms.

2) One further challenge is represented by the signal formed on the *induction* view. Due to the bipolar nature of the signal, induction views have poorer resolution compared to collection. Figure 2.8 (bottom left) and (bottom right) show a ν_μCC interaction vertex recorder on respectively the collection and induction views. On the collection view, the signal appears more neat and with a larger signal-to-noise ratio (represented in this case by the gray-scale of the image) than on the induction counterpart. This effects make more complicate the tridimensional reconstruction of the event, which relies on both bi-dimensional projection. In recent times, the MicroBooNE collaboration has successfully applied a deconvolution technique on the induction signal, in order

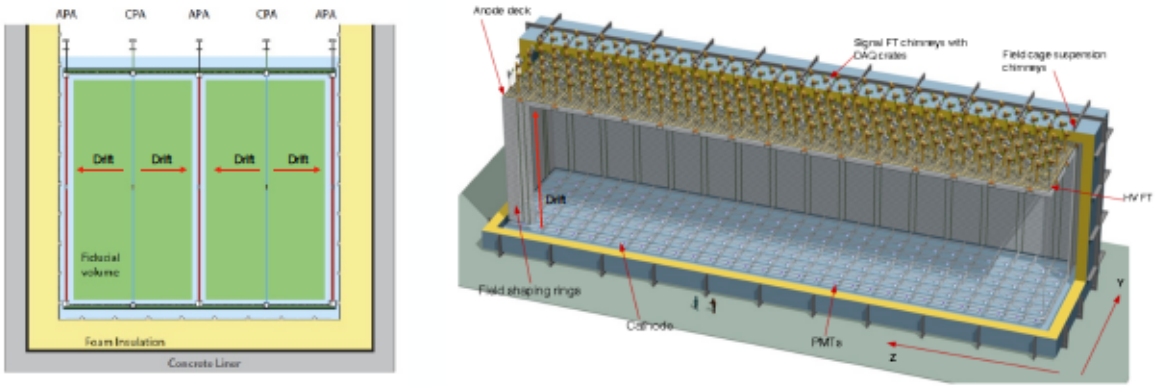


Fig. 2.9 **Left**: Transverse section of the DUNE single-phase LArTPC argon volume, with the four drifting regions and the anode-cathode arrangement (APA-CPA). **Right**: Open view inside the cryostat and the field cage of the dual-phase DUNE 10 kt far detector module. The location of cathode, anode, and PMTs is also highlighted. SFT chimneys and Slow control feedthroughs are located on top of the detector. The active drift region is 12 meters [56].

to mitigate the impact of the bipolar shape [65]. Although this techniques bear a noticeable improvement to the data quality, it relies on the *a priori* knowledge the electronic response of the electrode and demands large computing power to perform the convolution operation over complex waveforms.

3) Increasing the size of the detector demands also to increase the length and the number of readout wires. This impact on the cost, robustness and achievable signal-to-noise of the detector. Longer wires in general have a larger inter-strip capacitance which produces larger intrinsic noise. Reducing the length of wires is possible at the consequence of increasing the complexity of the detector and the number of readout channel, and thus, the cost of the experiment.

4) A strategy to maintain a very low level of noise is to operate the front end electronics at cryogenic temperatures. This means to include the front-end readout cards inside the cryostat, with the consequence to have them difficult to access and replace, making the entire detector more vulnerable to unpredictable malfunctioning.

Recent R&D developments for single-phase LArTPC have addressed part of the previous challenges. A proposed arrangement of a DUNE 10 kilo-ton module uses a modular structure in order to limit the drift length to only 3.6 meters. A sketch of the proposed setup is shown in Figure 2.9 (left). This has the major limitation to introduce dead material directly within the Argon volume. Furthermore, the number of channels to read out remains very large, and the cold front-end electronics is not accessible inside the cryostat.

2.6 Overview of the dual-phase design for DUNE

A dual-phase LArTPC is also being considered for at least one of the DUNE 10 kt modules. Dual-phase LArTPCs are less subject to most of the issues previously described. The possibility to amplify the charge allows to exploit a fully active drifting volume, without the necessity of the complicate arrangement of Figure 2.9 (left). The amplification allows also to compensate for the charge losses, enabling the use of two and not only one *collection* plane. Furthermore, the charge readout of a dual-phase LArTPC is located on the top surface of the detector. This has the advantage to allow access to the front-end electronics without unsealing the cryostat. The readout surface can also be segmented more easily, and the number and length of readout channel better controlled. The main limit of the dual-phase LArTPC is the necessity to sustain a sufficiently high gain with outstanding stability, in particular against fluctuations of temperature and pressure.
The design proposed for DUNE follows the concepts of GLACIER [95, 96], the 50 kt cryostat design proposed in the context of the Long Baseline Neutrino Oscillation (LBNO) experiment, and successively converged into DUNE. Compared to the equivalent single-phase design detector [55], potential technical advantages brought by the dual-phase design are:

- Larger fiducial volume with active drift path of $\approx$ 12 meters;
- Gain of the ionization electrons in gaseous phase provides signal entirely decoupled from noise ratio, offering the potential to lower the energy threshold of the detector, while compensating for electron losses over the long drifts;
- Two charge collection planes with 3.125 mm readout pitch, yielding an increased granularity with fewer readout channel at the same time (about 50% less compared to the single-phase reference design);
- Accessible and replaceable cold front-end (FE) electronics;

A sketch of the dual-phase DUNE far detector module is shown in Figure 2.9 (right). It is built as a single active volume 60 m long, 12 m wide and 12 m high for a total 12.1 kt active mass located inside a 2 m thick cryostat. The active area is surrounded by the field cage on the walls, with an anode on top and a cathode near the bottom with an array of 180 photomultiplier tubes (PMTs) sitting underneath. The anode covers a total surface 12×60 m^2 and it is segmented with 80 3×3 m^2 Charge readout planes (CRP). The signal formed on each CRP is passed to the cryogenic application-specific integrated circuits (ASIC) amplifiers, which are connected through 240 signal feedthrough (SFT) chimneys (640 channels/chimney) with a microTCA crate ensuring the signal digitalization and data acquisition. The crates are connected via an optical fiber link to the DAQ back-end. Dedicated feedthrough chimneys ensure also the connection between the CRPs and the slow control system, which is used to monitor the level, the temperature and pressure probes, to align the CRP orientation, and to apply the high voltage (HV) bias to the LAS and to the extraction grid.
For the purpose of this thesis, it is relevant to introduce the design of the charge readout system and the front end electronics. For more details about the technical specifications of

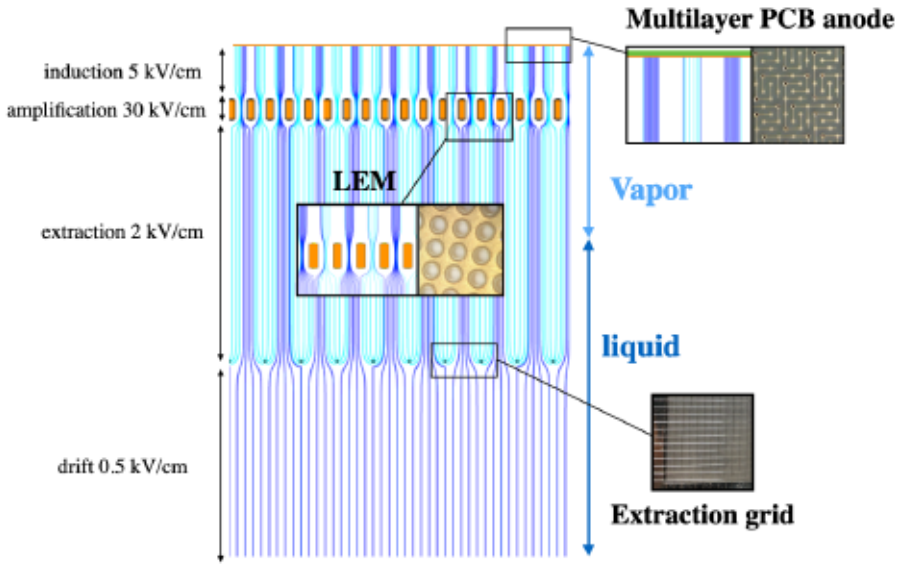


Fig. 2.10 A schematic representation of the charge readout system of a dual-phase LArTPC. The simulated field lines (in dark blue) indicate the path followed by the charge (without diffusion or space charge effects). [53]

the detector, the reader is pointed to the most recent public documentation of the DUNE collaboration [54–56].

2.6.1 The charge readout system

Implementing a successful extraction, amplification and subsequent collection of the charge is the main goal of the dual-phase charge readout system.The setup is illustrated in Figure 2.10; the extraction field to enhance the electron emission into gaseous Argon is provided by a wire grid immersed few millimeters below the liquid surface. The amplification of electrons is delivered by a Large Electron Multiplier (LEM), which is assembled together with the printed card-board (PCB) collection anode to form a LEM-Anode *sandwich* (LAS). The LAS can be assembled together to form a larger CRP.

The LEM anode sandwich (LAS)

The entire charge readout system is designed to be compact: the full extraction-amplification and collection operations are performed by in a 1.3 cm-thick layer with a surface of 50×50 cm^2. Due to the large number of components foreseen for the DUNE 10 kt module, each element of the LAS is designed to be manufactured as printed circuit board, which can be industrially produced at large scale and relatively low cost. The two PCB layers composing a single LAS are the LEM and the anode units. This components are shown respectively in Figure 2.11 (a) and Figure 2.11 (b) and described in more detail below:

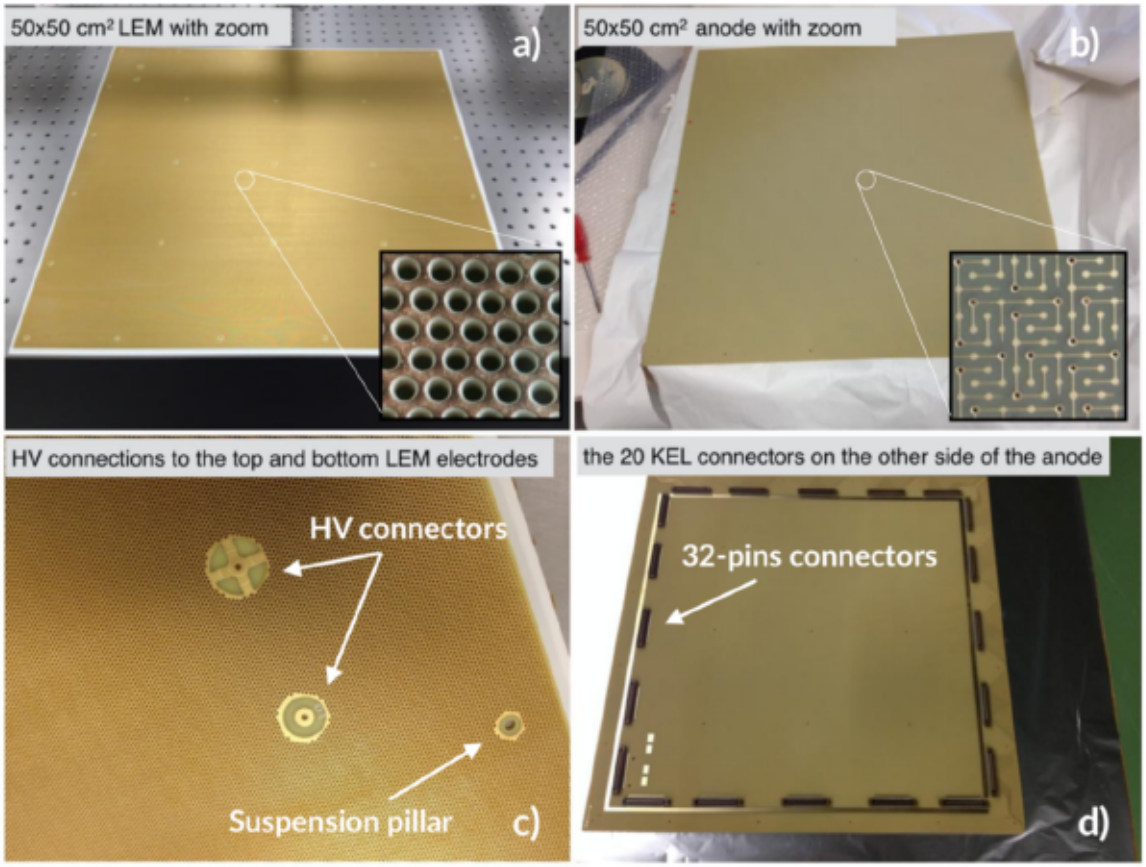


Fig. 2.11 The different elements composing the LEM-Anode sandwich: the 50x50 cm^2 full LEM PCB and a zoom on the hole structure (a), the 50x50 cm^2 Anode PCB with a zoom on the strip arrangement (b). A closeout image of the HV lem connectors and suspension pillars (c), and the location of the 32-pins connectors at the border of the LAS (d). [53]

- $50 \times 50\,\mathrm{cm}^2$ **LEMs**. Each LEM is produced from a 1 mm-thick copper-clay epoxy PCB with 500 µm-diameter holes drilled through. The holes are arranged in a honeycomb pattern with 800 µm pitch, resulting in a density of 200 holes per cm^2. Each hole is surrounded by a 40 µm dielectric rim. The rim has been demonstrated to be able to reduce the occurrence of discharges [97]. The diameter of the holes and the thickness of the LEM layer are engineered to maximize the electron amplification and provide containment for the UV photons produced during the avalanche process and thus acting as *mechanical quencher* to prevent discharge [98]. This is an important characteristic of the LEM design introduced specifically to cope with giant LArTPCs. In fact, the long drift distance imposes strict limits on the Argon purity and discourage the use of *chemical quencher* to mitigate the probability of discharges. The LEM design is the result of intense R&D and optimization work and the amplification of drifting charges with a stable gains of at least 15 after a *charging up* period of about 1 day has been demonstrated on a $10 \times 10\,\mathrm{cm}^2$ LEM [99].
- $50 \times 50\,\mathrm{cm}^2$ **PCB Anode**. The anode is manufactured from a single multilayer PCB. The readout consists of a pattern of gold-plated copper strips with a 3.125 mm readout pitch. The strips samples the two orthogonal projections (or views) along y and z and are electrically isolated from the others. The maze-like arrangement of the strips is visible in Figure 2.11 b) and it is the result of a design aiming at minimizing noise and inter-strip capacitance, while ensuring equal charge sharing between both views. Tests on $10 \times 10\,\mathrm{cm}^2$ anode have demonstrated a noise level of about ≈ 1000 equivalent noise charge (ENC) at cold, with an electrical capacitance of $150\,\mathrm{pF\,cm}^{-1}$ and charge sharing asymmetries at the level of 1% [100].

The LEM and Anode layers are assembled with a specific system to maintain a constant 1 mm distance [101]. Readout pins arranged in groups of 32 are located at the border of each LAS in order to create an electrical bridge between anodes from different LASs. The dead readout surface is lower than the 0.5 % of the total area and comes mostly from the suspensions pillars, the LEM high voltage connectors, and the LEM frames (Figure 2.11 (c) and Figure 2.11 (d)). The nominal high voltage settings and the distances between each pair of layers are sorted in Table 2.2.

The charge readout plane (CRP)

The CRP design for DUNE is shown in Figure 2.12. It is composed by 36 LAS electrically bridged together trough the anode pins to form an homogeneous readout surface of $3 \times 3\,\mathrm{m}^2$. At the bottom of each CRP, lies a wire grid composed of two orthogonal array of stainless steel wires with 100 µm of diameter, tensioned to cover the entire surface of the CRP. The wire pitch is 3.125 mm. The role of the grid is to provide the $2\,\mathrm{kV\,cm}^{-1}$ electric field defined across the liquid-gas interface. This field is applied between the grid submerged in liquid and the bottom side of the LEMs. Each LAS is kept in position by a G10 frame, which also surrounds the entire CRP and holds the extraction grid. The main structure is supported by an Invar nickel-iron alloy structure, which provides also the support for the suspensions cables. Special connectors allows to compensate for the different dilatation coefficients of

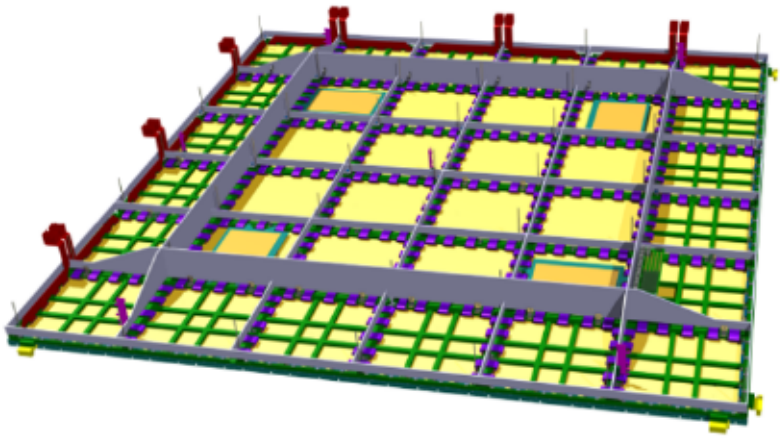


Fig. 2.12 The full 9 m^2 CRP. The Invar structure is drawn in gray and the G10 frame is shown in green. The yellow surface represents the surface covered by the assembly of LAS units. The extraction grid is located beneath the LAS surface and thus not visible here.

Table 2.2 The nominal voltages, distance, electric field, and tolerance settings, sorted from bottom to top as proposed for the dual-phase DUNE CRP design [53].

Component	Potential (kV)	Distance (mm)	Electric field ($kV\,cm^{-1}$)	Tolerance (mm)
LEM bottom electrode-grid	1.5 (in GAr)	0.5 (in GAr)	3 (in GAr)	1
LEM top-bottom electrode	3.2	1	30-35	0.01
Anode-LEM top electrode	1	2	5	0.1

Invar and G10. The CRP is suspended from the top of the cryostat with a system of stainless steel cables which can be electrically operated from outside in order to regulate the CRP alignment and vertical level with high precision.
The total effective gain of the detector it is provided by the electric field applied between the different layers of the CRP. With the nominal field settings shown in Table 2.2, the CRP is expected to be able to provide an average effective gain larger than 15, ensuring extraction and charge collection efficiencies of fast electrons larger than 90%. Table 2.2 reports also the maximal tolerance on the alignment of the different layers required to ensure a uniform gain on the whole CRP structure within $\approx 5\%$. A 1×3 m^2 CRP composed by the previously described LAS units was operated for the first time by the $3\times1\times1$ m^3 prototype. The results are described in Chapter 4. Two full scale 3×3 m^2 CRPs are expected to be tested with the $6\times6\times6$ m^3 prototype.

2.6.2 Data readout system

The data readout electronics is responsible to create a digital signal from the signal current collected at the anode. The design proposed is based on three main requirements:

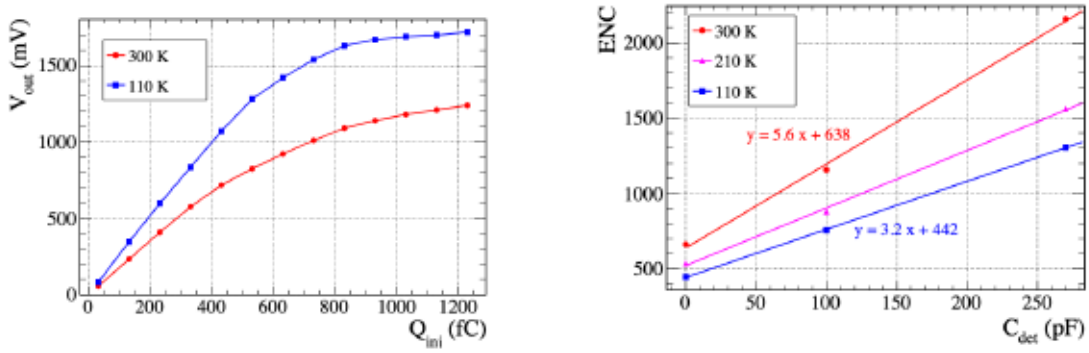


Fig. 2.13 Left: Expected Voltage returned by the ASIC as function of the detector input charge. Right: Expected ASIC noise both at cold and room temperature as a function of the detector input capacitance[102].

- A low noise cold front-end (FE) electronics with large dynamic range able to cover the charge variations due to the long detector drift and amplification.
- Possibility to access the FE electronic cards from outside with the cryostat filled.
- A low-cost and fast data acquisition system (DAQ) to readout, process and store the information from the 153600 channels of the detector.

Front End cryogenics cards and chimneys

The FE electronics implementation has to match with the high dynamic range of the signal coming from the two collection views of the PCB anode after the LEM amplification. In the nominal high-voltage settings the maximal signal expected from electrons and hadron showers corresponds to 1200 fC.
In order to optimize the resolution while preserving the large dynamic range of the detector, the dual-phase ASIC chips feature 16 input channels each consisting of a charge sensitive amplifier (CSA) with a linear gain for input charges of up to 400 fC and a logarithmic response in the 400-1200 fC range. The double-slope behavior is obtained with a MOSCAP capacitor in the amplifier feedback loop that changes the capacitance above a certain signal threshold of injected charge. Figure 2.13 (left) shows the response of the CSA for different values of injected charge measured at room and cryogenic temperature. Figure 2.13 (right) shows the expected equivalent noise charge (ENC) for different value of the detector input capacitance C_{det} at room and cryogenic temperatures. In cryogenic conditions, 1300 ENC is expected for the 250 pF design capacitance of the PCB anode.
In the current design, the ASICs cards are located in cold within the SFT chimneys. Each chimney is a 2 meter long stainless steel tube which provides an isolated volume from the gas Argon. A *cold flange* interconnects the PCB anode to the FE electronics located inside. A second flange at the top of the chimney, *warm flange* connects the analog signal from the FE electronics to the DAQ.

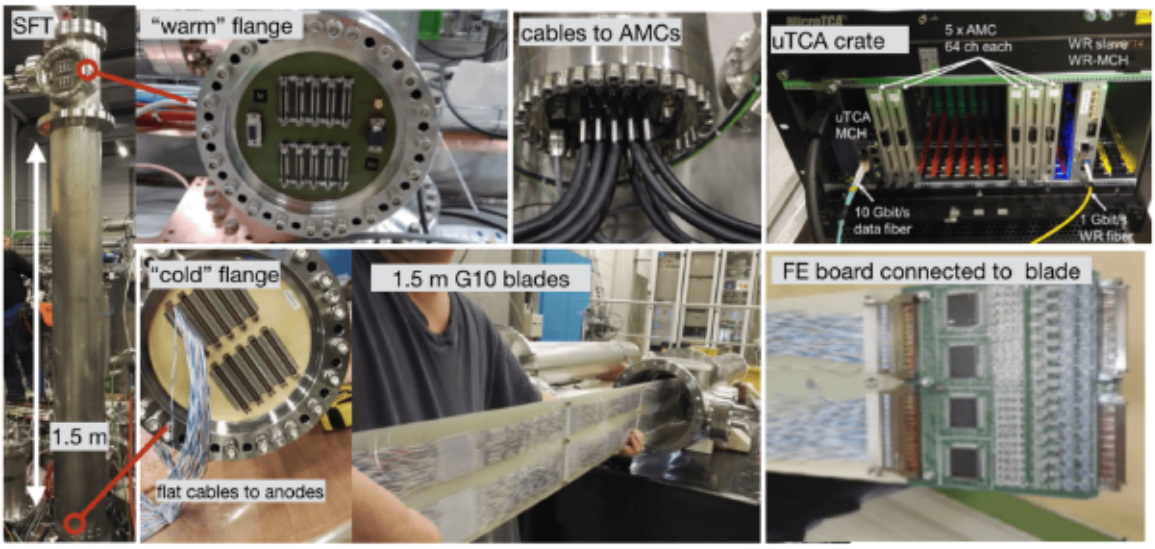


Fig. 2.14 Pictures with details of a SFT layout and of a microTCA crate taken during the construction and commissioning of the $3\times1\times1$ m^3 prototype [102].

The main role of the chimneys is to provide access to the cold FE without the necessity to open the cryostat. For this reason the FE cards are mounted on special 1.5 meters long blades which facilitates the extraction of the FE card trough the *warm flange* and to ensure that the card is correctly plugged into its connectors at the *cold flange*. In Figure 2.14 are shown pictures of one of the $3\times1\times1$ m^3 chimney and details of the SFT layout. In the current far detector design each ASIC chip supports 32 channels divided into groups of 16. The total number of channels per chimney is 640 for a total of 3 chimneys per CRP.

Digital electronics and DAQ architecture

The main requirement for the DUNE DAQ design is mostly related to the large scale data processing which calls for efficient and low-cost solutions to synchronize a large number of channels and to process and store large event samples at relatively high rate. The DAQ is designed to be highly scalable and it is divided into two level, called L1 and L2.
L1 DAQ is responsible for the digitalization of the individual signal recorded on each channel. The signal is continuously digitized in 12-bit accuracy at a frequency of 2.5 MHz, by Advanced Mezzanine Cards (AMCs) located in commercial microTCA crates [103]. Each microTCA is located at the top of each chimney and synchronized with the others thanks to a high-accuracy clock signal following the White Rabbit standard [104]. The length of the sequence covers the entire drift window for the ionizing charge to arrive from the cathode to the anode. If no zero suppression is implemented, the total number of time samples depends on the total drift length of the detector. For ProtoDUNE and DUNE respectively 10k and 20k samples are expected for each channel.
The backplane of each microTCA crate is connected through a 10GbE uplink cable to a dedicated back-end FPGA processing boards, which are part of the L2 DAQ. The 10GbE

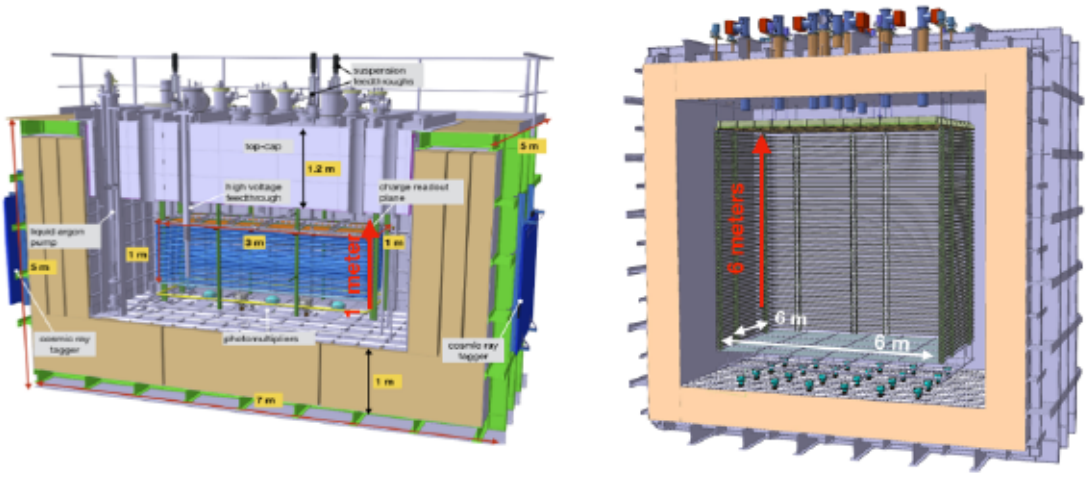


Fig. 2.15 The two DUNE dual-phase prototypes: the $3\times1\times1$ m^3 prototype (left) [102] and the $6\times6\times6$ m^3 or ProtoDUNE-DP (right) [66]. The two sketches are not in scale one with respect to the other (note the two different drift lengths highlighted with the red arrow).

uplik determine the maximal event rate sustained. Since each crate is connected to the 640 channels readout from a single chimney, and each chimney is read in parallel, the total event rate is found to be $\approx$ 100 Hz and 50 Hz for respectively ProtoDUNE and DUNE in case of a lossless transmission scheme. A factor 10 compression can be achieved in case zero suppression or similar algorithms are implemented. The L2 DAQ has the main role of interconnecting all the crates and implementing the event building logic. The card used follows basic commercial standards and it was chosen for its high-performance with massive data processing and low latency time.
The previous description aimed at giving an overview on the DAQ requirements and proposed architecture for the dual-phase design of DUNE. Presenting the full proposed architecture or specific implementations of it is far from the scope of this thesis. A broader description can be found in the related DUNE documentation [53, 56, 101, 102].

2.7 The DUNE dual-phase prototyping strategy

The design for the DUNE 10 kt module previously described is based on the performances of a small dual-phase LAr TPC with only three liters of active volume and one LAS with a 10×10 cm^2 LAS exposed to cosmic rays [99, 100, 105]. However, the design featuring a 50×50 cm^2 LASs assembly on a larger CRP had never been tested before. In order to prove assess the performances of the proposed setup at the size required by the DUNE far detector, the DUNE collaboration has planned a prototyping strategy articulated over more than 5 years and culminating with the publication of a Technical Design Report (TDR) explaining the feasibility of the technical solution as designed for the DUNE 10-kt module.

The scalability of the detector is tested in two stages with prototypes featuring increased size of the active volume, drift length and readout surface.

2.7.1 The $3\times1\times1$ m^3 prototype

The $3\times1\times1$ m^3 prototype is a 4 t dual-phase LArTPC with 1 m drift. A sketch of the detector is shown in Figure 2.15 (left). It operated at CERN during 2017 and collected more than 500k cosmic triggers. Its main goal was to test for the first time the performances of the charge readout over a 3×1 m^2 surface and the amplification provided by the 50×50 LEMs. A large portions of the work carried out for this thesis hinges on the characterization of the performances of this prototype and a broader description is presented in Chapter 4.

2.7.2 ProtoDUNE-DP ($6\times6\times6$ m^3 prototype)

The $6\times6\times6$ m^3 prototype also known as ProtoDUNE-DP [101] is being ultimated at the time this thesis is written and it is scheduled to take data during Autumn 2019. A schematic representation of the ProtoDUNE-DP layout it is shown in Figure 2.15 (right). Its active volume is 300 t (about 100 times larger than the previous $3\times1\times1$ m^3 prototype) and has a drift length of 6 meters. Its "twin" detector, ProtoDUNE-SP has collected data using both cosmic rays and a charged particle beam in Fall 2018 [64]. ProtoDUNE-DP use instead only cosmic rays to test and characterize the proposed charge readout in scale 1:1, with two instrumented planes of 3×3 m^2 each. Particular attention will be given to the study of the detector stability over a prolonged period of time and to the homogeneities of the effective gain over the entire readout surface. ProtoDUNE has also the role to support calibration efforts for the DUNE detector.
Upon the success of the operation with cosmic rays, there is also the possibility to gather data with a calibrated particle beam in 2021, after the shutdown of the facilities of CERN. Studying the detector response using charged particles in a calibrated beam at the same energies covered by the DUNE beam is of particular importance to understand the interaction between charged particles and Argon and understand the energy scale and resolution of the detector.

2.7.3 Developing software tools

In parallel to the test of the full scale detector components, the prototypes offers the possibility to develop dedicated software tools for the dual-phase detectors. This is a necessary ingredient to ensure the correct understanding of the detector and test novel analysis strategies in preparation to DUNE. The efforts done for this thesis to integrate a complete dual-phase simulation and automated reconstruction are described in Chapter 3 and applied for the first time ever on the the data collected with the $3\times1\times1$ m^3 prototype, as shown in Chapter 4.

Chapter 3

Dual-phase simulation and reconstruction in LArSoft

The chapter describes the status and the work done for this thesis on the software tools necessary for the main physics goals of DUNE. The simulation and the reconstruction here presented have been used for the analysis of the performance of the $3\times1\times1$ m^3 prototype described in Chapter 4 and future tests and improvements are planned for the reconstruction of cosmic tracks and showers in the $6\times6\times6$ m^3 prototype. The performance of the reconstruction of simulated beam neutrino at the DUNE dual-phase far detector will be presented in Chapter 5.
All the software tools here introduced are developed with LArSoft, the common software framework shared by many LAr experiments. Its characteristics are briefly introduced in Section 3.1. Only a simple definition of the dual-phase far detector geometry, illustrated in Section 3.2, existed for dual-phase within LArSoft before this thesis. The work described in this chapter is the summary of many collaborative efforts. For this thesis, we have focused mainly on:

- Interfacing the output of generators such as GENIE and CORSIKA with LArSoft within the dual-phase specific requirements.
- Including in the dual-phase simulation the parameterization of the gas processes, and the noise simulation.
- Adapting and improving for the dual-phase use the higher-level reconstruction algorithm for bidimensional clustering and tridimensional trajectory fitting.
- Developing a Deep-Learning model, based on a Convolutional Neural Network, capable to distinguish between track-like and shower-like energy deposits.
- Organizing the reconstruction and simulation reconstruction in a set of automatized processes executable with LArSoft and optimized for the LBL analysis of DUNE. More details about this aspect are given in Chapter 5.

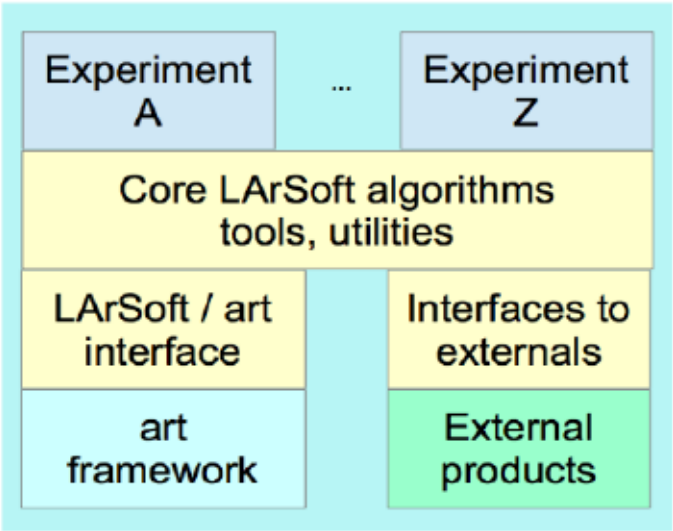


Fig. 3.1 Description of the LArSoft layered architecture. It is particularly highlighted how the specific LArSoft core algorithms and utilities bridges together the *art* framework, third part products, and experiment specific configuration.

The dual-phase simulation is described in Section 3.3 and Section 3.4. The standard dual-phase automatic reconstruction chain is described in Section 3.5, while the Deep-learning model, the preparation of the dataset, the training strategy, and the robustness of the results are described in Section 3.6.

3.1 LArSoft framework

The Liquid Argon Software (LArSoft) [106] was firstly design as offline software processing for the ArgoNeuT experiment [62] and then expanded to support all the LAr experiment operating at Fermilab. The framework is written in C++ and built on *art* [107], the event-processing system established at FNAL and used by most experiments hosted at the lab. Since the basic data format is common from most LArTPC experiments, the code has been designed to maximize interoperability and provides a common interface, infrastructure and algorithms for simulation, reconstruction and analysis. The structure of the framework is shown in Figure 3.1. LArSoft provides defined packages such as GENIE [108] or GEANT4 [109], which are easily accessed by particular experiment by a run-time configurable descriptions of their unique features, such as geometry and detector response.
User-written *modules* extend the *art* framework and allow to access and process the information stored event by event. There are two types of modules: the *Analyzers*, which allows to access data from the events stream, and *Producers*, which have also the ability to store new data-products into the events to be used for future processing. A reconstruction algorithm is an example of *Producers*, because it reads and aggregate the lower level information to form a higher-level data product to store in the event stream for further processing. Along with

Analyzers and *Producers* modules, the framework allows to introduce customizable *Services*, which live outside the event stream and may be used to obtain general information such as the LAr properties or information about the detector geometry whenever needed.
The Fermilab Hierarchical Configuration Language (*fhicl*) is used to define the configuration file of an *art*/LArSoft job by specifying the modules and services to use (and their order) and the configuration of their run-time parameters. The general end-to-end configuration for an experiment involves four standard stages: generation, propagation (usually executed by GEANT4), detector response simulation, and reconstruction. The producer module relatives to each stage is configured in a dedicated *fhicl* file. Custom *Producers* modules and *Services* are integrated into the standard simulation and reconstruction at different stages, implementing the dual-phase modification when needed.

3.2 Detector geometries

Figure 5.1 shows the first basic geometry of one of the 10 kt cryostats of the DUNE far detector, which existed in LArSoft at the time of the first CDR publication. The drift direction is along the horizontal x coordinate; this is related to a limitation embedded within the framework which was originally developed for single-phase TPCs and assumes a drift perpendicular to the gravity direction [1].
The readout surface is located immediately after a 1.3 cm thick gas Argon layer at the end of the drift region. The CRPs are simulated as two planes made up of arrays of wires, orthogonal with respect to each other, representing the copper-gold strips. Each direction defines one of the two projection of the detector: View 0 is defined as the projection along y, while View 1 is the projection along z. The geometry implements also the 1 cm thick borders between CRPs given by the G10 frame, but not the borders between the various LAS inside a CRP. The effect is introduced when considered relevant by declaring the readout wires at the border of the LASs as dead using a specific configurable parameter.
Due to the large computational resources required by processing the large amount of channels at the DUNE far detector, for most of the full-simulation studies, a reduced geometry is used. This geometry is called *workspace* geometry and it features only 8 CRPs, arranged in two groups of four along the beam direction. This ensure an active volume which is large enough to have most of the events contained and therefore allows to develop with much ease reconstruction and calibration tools that will be tested later with the full detector geometry.
The geometries of the $3\times1\times1$ m^3 and the ProtoDUNE prototypes have been introduced in LArSoft similarly to the far detector geometry. For the geometry of ProtoDUNE, the far detector geometry have been reduced to include only 4 CRPs and 6 meters drift, while for the geometry of the $3\times1\times1$ m^3 prototypes, only one CRP has been defined with a surface of 1×3 m^2 and only 1 meter drift.

[1] A correction for this is not simple and it was necessary to modify the framework structure and to coordinate strictly with the *art*/LArSoft team. At the time of this thesis, efforts are still ongoing to ensure the full compatibility of the new rotated framework with the existing geometries and therefore the anomalous drift direction is still in use.

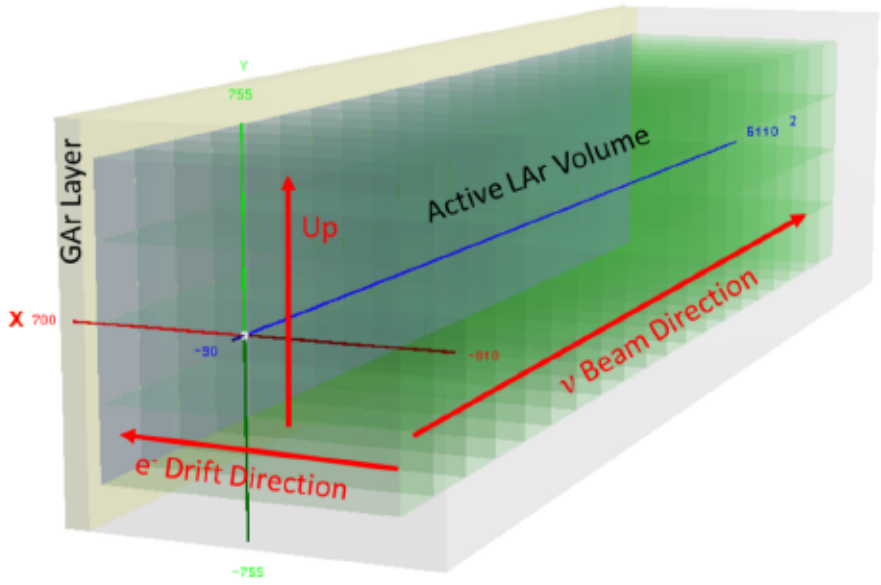


Fig. 3.2 The dual-phase 10 kt DUNE far detector geometry description in LArSoft.

Other elements of the detector, such as the cryostat vessel, the cathode and the PMTs have been also included to the geometry of both DUNE and related prototypes. However they are not described in detail here.

3.3 Event generation and propagation

The relevant event generators used in the scope of this thesis are GENIE and CORSIKA. Particle propagation is performed with GEANT4. For this part the existing LArSoft interfaces allowed a smooth integration with the dual-phase geometry, in particular for the particle propagation. Some modifications have been done to the event generators to account for the anomalous orientation of the drift direction, and on the simulation of the electron extraction from liquid to gas.

3.3.1 The neutrino event generator

The LArSoft generator *module* for beam neutrinos accepts histograms modeling the expected unoscillated on-axis fluxes at the DUNE far detector and models the neutrino interactions using the GENIE v2 event generator [108] to return a data product containing all the primary particles produced in the final state of the neutrino interaction which will be subsequently propagated in the LAr volume. The tuning of the GENIE setting used follows the DUNE TDR prescriptions and can be found in [50].
Although the LArSoft *module* works correctly along with the dual-phase geometry, it has been necessary to rotate the input beam direction contained in the fluxes in order to imitate the correct detector orientation.

3.3.2 The cosmic ray event generator

The simulation of cosmic rays is relevant for both the $3\times1\times1$ m^3 and $6\times6\times6$ m^3 prototypes. In particular for the $3\times1\times1$ m^3 prototype efforts were made to include in the simulation a realistic cosmic ray simulation featuring the correct particle multiplicity in order to simulate correctly the detector trigger, which is based on a threshold applied on the quantity of light seen by the photomulitpliers. Also ProtoDUNE is expected to observe cosmics rays and it is important to rely on accurate simulation in order to understand the effective gain.
Primary particles are selected from a pre-generated CORSIKA [110] cosmic ray sample containing more than 100k air showers generated at location and altitude of CERN. Air showers are drawn randomly from the samples, and their start positions and start times are randomized in order to be accepted within a buffer surface over the detector and inside the acquisition time. Time and space correlation of the particles inside the same shower are preserved. The buffer surface and the time duration of the acquisition window are user-configurable.
Also for this LArSoft generator module it was necessary to perform a small rotation of the input in order to ensure the compatibility between the LArSoft geometry and the realistic dual-phase setup which requires the drift to be parallel to the direction of gravity.

3.3.3 Particle propagation in LAr and ionisation electron drift

Particle propagation is handled by GEANT4. The GEANT4 interface with the LArSoft framework allows to configure details of the simulation such as the physics list, the enabled processes, and the relevant energy cuts. The QGSP_BERT physics list is a standard physics list used for high energy physics experiments[111] which is well suited for energy ranges above hundreds of MeV, and thus used for all the simulation described in this thesis. The optimal maximal step length and and the lower energy threshold cut implemented for the DUNE single-phase simulation have been applied identical also for dual-phase.
The energy deposit ΔE produced for each step is converted in ionization charge $\Delta Q = e\Delta E/W_{ion}$ with e being the elementary electron charge and $W_{ion} = 23.6$ eV the ionisation work, as given in Table 2.1.
The quenching on the charge due to recombination effects described in Section 2.2.4 is accounted by multiplying ΔQ by the recombination factor calculated according to the modified Box model [76]. The drift field configured is $\mathscr{E}_{drift} = 500$ V/cm.
The ionization charge produced at the position $\{x,y,z\}$ is projected on the readout plane. In case of a homogeneous drift field the final coordinates are found to be: $\{t_{drift} = (X_{readout} - X)/v_d, y_{readout} = y, z_{readout} = z\}$, where v_d is the drifting velocity at the field $\mathscr{E}_{drift}$. If the drift field is not homogeneous, the final charge position is extracted from specific tables reporting the adequate corrections which are externally configured into a specific LArSoft *service*. This is the approach commonly used to introduce Space Charge effects into the simulation.
Once the final projected coordinates $\{t_{drift}, y_{readout}, z_{readout}\}$ are found, some correction are applied in order to account for the effect of diffusion and the electron emission time from liquid to gas:

- **Diffusion:** is included in the simulation by smearing the coordinate $\{t_{drift}, y_{readout}, z_{readout}\}$ with a normal distribution having $\sigma = \sqrt{D_{L,T}\, t_{drift}}$. $D_{L,T}$ is the longitudinal (transverse) diffusion coefficient at the field $\mathscr{E}_{drift}$ as shown in Figure 2.4 (left).

- **Emission time:** It becomes significantly long when extraction field in liquid is lower than 2.0 kV cm^{-1}. In this cases, the electrons are not extracted immediately from liquid and the process may take up to several ms. For this reason a delay t_{ext} is added to t_{drift}. t_{ext} is randomly generated following an exponential law with time constant τ_{ext} depending on $\mathscr{E}_{ext}$. Although a parameterization of τ_{ext} as function of $\mathscr{E}_{ext}$ as shown in Figure 2.5 was already obtained in a previous work [83], the relation configured in LArSoft has been extrapolated from the $3\times1\times1$ m^3 prototypes data with the methodology presented in Section 4.3.1.

The final charge collected at readout $\Delta Q_{readout}$ may be calculated as follows:

$$\Delta Q_{readout} = \Delta Q \cdot \exp(-t_{drift}/\tau_e) \cdot G_{\text{eff}}. \tag{3.1}$$

The exponential term in brackets is the charge attenuation due to Argon purity, with τ_e depending on the concentration of impurities in LAr from Equation 2.9. The second term, G_{eff} represent the effective gain of the detector. G_{eff} depends on a non-trivial interplay between the extraction, amplification and induction electric fields, and requires dedicate simulations to compute the trajectories of the electrons in gaseous Argon and simulate the Townsend multiplication. More details about the current status of the simulation of G_{eff} are given in Section 4.3.2. After all the necessary corrections, $\Delta Q_{readout}$ is are assigned to the closest readout channel to form the signal current.

3.4 Detector response

The simulation of the detector response mimics the effect of the Front End amplifier and the DAQ on the collected charge signal to form the output waveform on the related readout channel. A simulation of noise is also implemented at this level and can be added directly on the simulated signal.

3.4.1 Signal convolution and digitalization

After the simulation of the drift effects described in Section 3.3.3, the projected charge $\Delta Q_{readout}$ and the associated arrival time t_{drift} is assigned to a readout channel to form the signal current $I(t)$. The final signal recorded by the detector $V_{out}(t)$ is obtained by the convolution between $I(t)$ and the response of the amplifier $h(t)$:

$$V_{out}(t) = \int_{t_0}^{t} I(t')h(t-t')dt'. \tag{3.2}$$

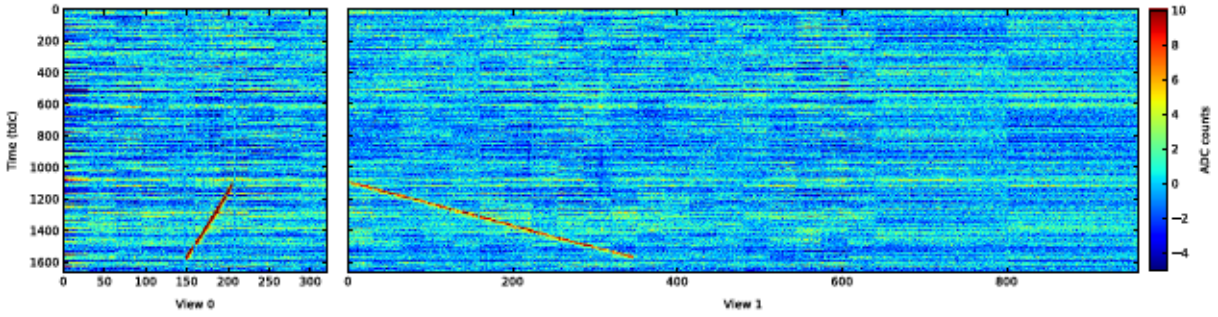
$h(t)$ depends on the electronics design. For the one described in Section 2.6.2 the analytical form of $h(t)$ is:

$$h(t) = \frac{\tau_1}{\tau_1 - \tau_2} \cdot \left(e^{-t/\tau_1} - e^{-t/\tau_1} \right), \tag{3.3}$$

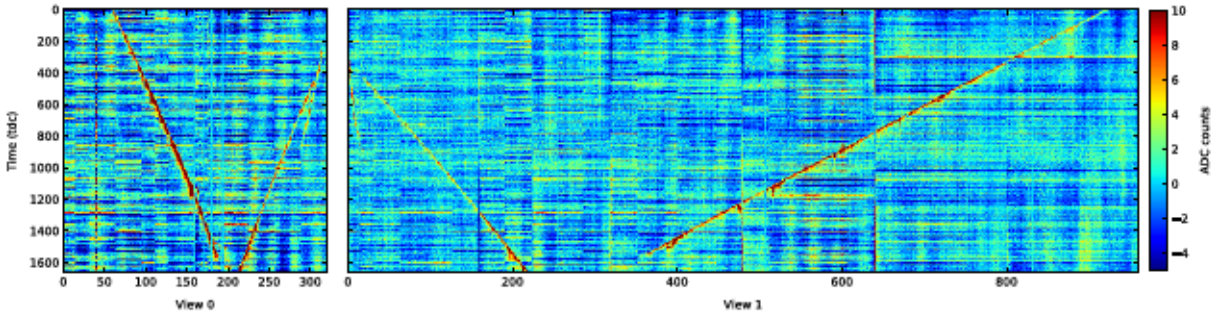
where τ_1 and τ_2 are the rise and fall time of the waveform. The convolution procedure is realized thanks to a LArSoft *Service* already implemented for the single-phase detectors and no modifications to the existing algorithm were found necessary for our case.

3.4.2 Data-driven noise simulation

In realistic operation conditions of a LArTPC, noise is formed on the readout strips. There are several possible sources which may induce noise, and often it is not possible to investigate and provide an analytical description of the noise level inside the detector.
The use of a data-driven noise simulation is mostly justified by the observation of coherent noise patterns on the events collected by the $3\times1\times1$ m^3 prototypes. This, coupled with the relatively low signal-to-noise ratio achieved during operations of the detector, has an impact on data reconstruction in the $3\times1\times1$ m^3 prototype and thus realistic noise simulation are needed in order to gain a formal understanding of the response of the charge readout and the waveform formation. An accurate description of noise in the $3\times1\times1$ m^3 detector is given in Chapter 4, while the main scope of this section is just to describe the simulation strategy adopted to have a general data-driven simulation of noise in dual-phase LArTPCs.
In general, coherent noise patterns are characterized by a signal with same frequencies and phases across several readout channels. It is possible to study the correlations between channels in order to understand which groups of channels are sharing similar signals in phase and frequency content, and replicate this behavior in the randomization of the simulated waveform. The noise simulation is done using a LArSoft service which accepts as input an average frequency distribution (spectrogram) for each channel, and the expected groups of correlated channels.
The modules uses then an Inverse Fast Fourier Transform (InvFFT) to extrapolate the expected time domain average signal on each channel from the spectrogram. The amplitude of each frequency is randomly smeared using a normal distribution with a user-configurable standard deviation. Once the signal is converted back to the time domain, each waveform gets a phase value randomly extracted from a uniform distribution in the range $[-\pi; \pi]$. The same random value of phase is applied to each group of correlated channels provided as input.
The LArSoft *Service* for the noise simulation has been designed to be agnostic of the detector geometry. Information on the number of channels and on the detector sampling window and frequency are extracted from the correlation patterns and the input frequency provided as input. However, if the simulated waveform in time domain after the InvFFT has not the same length of the expected sampling window of the detector, the module is programmed to mirror or trim the expected waveform in order to equalize the window. It is important to notice that by mirroring the simulated waveform to cover a longer drift window is not expected to produce a fully realistic representation of the expected noise waveform. This happens as consequence of the Nyquist-Shannon theorem, which sets relationship between the sampling



(a) Data-driven noise simulation injected on a simulated event in the $3\times1\times1$ m^3 detector.



(b) A cosmic ray event collected with the $3\times1\times1$ m^3 detector.

Fig. 3.3 Qualitative comparison between the data-driven noise simulation (a) and a real event (b) collected with the $3\times1\times1$ m^3 prototype. The correlation patterns found in data (see text and Chapter 4) is well reproduced by the simulation, along with the noise differences between the 3-meter strips long (left) and the 1-meter strips long (right) readout planes.

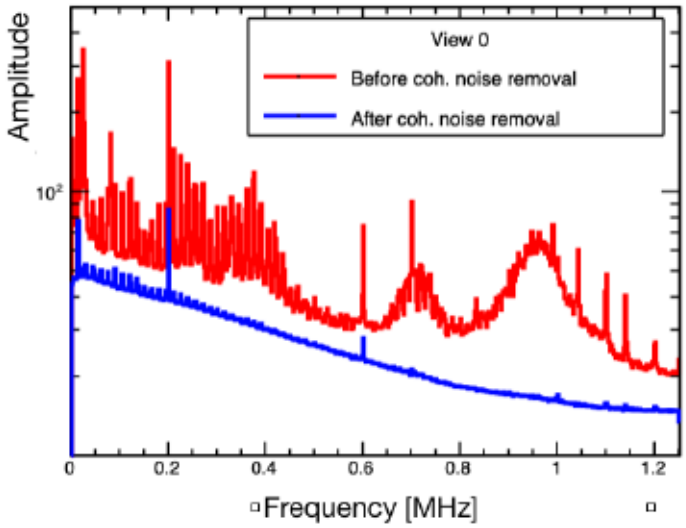


Fig. 3.4 The average frequency spectra before and after the use of a coherent noise filter (see Section 3.5.1) for one strip taken from the run used as noise reference for the $3\times1\times1$ m^3 detector. The coherent noise spectra in blue is used for the noise simulation in the $3\times1\times1$ m^3 prototype, while incoherent spectra the red is used for the DUNE far detector.

frequency f_{samp}, the number of time samples N and the lowest frequency f_{min} which can be resolved:

$$f_{min} = \frac{f_{samp}}{N}$$

The mirrored waveform is thus not representative of any frequency below f_{min}.
In spite of the caveat previously described, a mirrored waveform extracted from the $3\times1\times1$ m^3 incoherent noise pattern has been injected in the simulated sample of beam neutrinos at the dual-phase DUNE far detector. This was done in order to provide an approximated indication of the incoherent noise level coming from the 3-meters long strips.

3.5 Event reconstruction

LArTPC detectors allows to produce a fine tracking of the trajectories of the various particles produced in the final state of the interaction. Although this is one of the great advantages offered by the technology, quite sophisticated pattern recognition tools are necessary to extract the large amount of information available with great precision. Since the data are high resolution bidimensional projections of the interaction, it is necessary to have versatile reconstruction algorithms able to separate different types of topologies already with very low level information. Furthermore, matching the two-dimensional projections to form a tridimensional event requires often advanced algorithms, able to deal with topological ambiguities. This task is not easy and in certain conditions view matching is also complicated for a human brain. As a general overview, the approach to reconstruction commonly followed for LArTPC events can be divided into 4 main steps, schematically shown also in Figure 3.5:

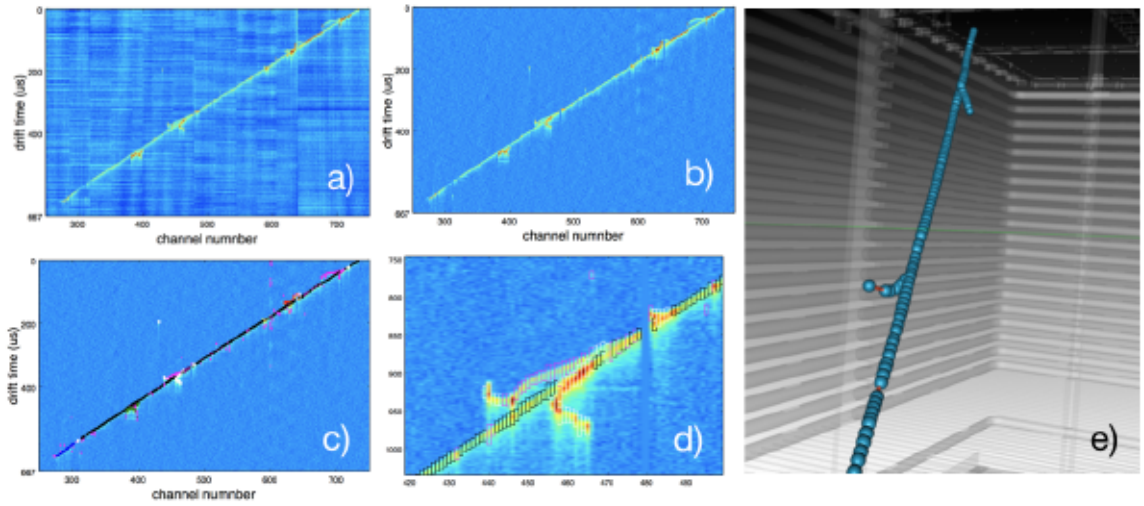


Fig. 3.5 Schematic example of the different reconstruction stage applied on a cosmic muon track collected with the $3\times1\times1$ m^3 detector. The raw data (a) are first preprocessed in order to remove the noise (b) and identify the signal hits (c). Hits are grouped into 2D clusters (d), and clusters are finally associated to form a 3D image of the event (e) [112].

a) **Signal processing**: during this stage the raw data collected on each strip are processed in order to filter the noise and identify the signal Regions of Interests (ROIs) as shown in Figure 3.5 (b).

b) **Hit finding**: during this stage, the signal on each strip is analyzed to identify and separate the signal from different particles and reconstruct the individual energy deposit on the strip.

c) **2D Clustering**: during this step hits on different strips are associated based on similarity criteria in topology, distance, and charge.

d) **3D Matching**: the individual bidimensional clusters of hits found on each view are matched together to build a 3D image of the event. High level information usually needed in support for physics analysis studies is often extracted here.

Each of the four steps is described in detail in the following sections.
In recent times, techniques of machine learning used for computer vision called Convolutional Neural Networks (CNNs) have started to be deployed to process the LArTPC images, both as support to more traditional reconstruction algorithms or as stand-alone models. In Section 3.6, we present a CNN created to perform a local classification of the electromagnetic energy deposits in support to 2D clustering algorithms.

3.5.1 Signal pre-processing

The ionization electrons induces a current on the anode plates which is transformed into a digital signal by the FE electronics and the DAQ systems. The signal is then recorded

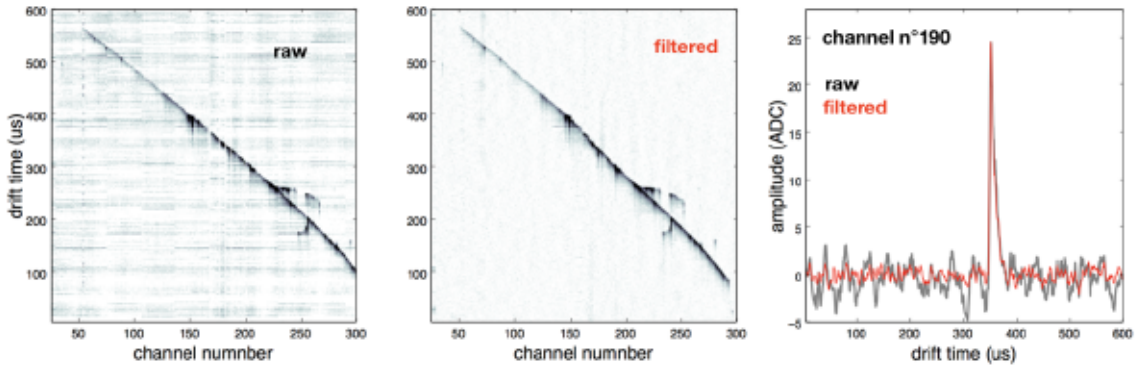


Fig. 3.6 A cosmic muon event collected with the $3\times1\times1$ m^3 detector before (left) and after (center) the pre-processing strategy described in Section 3.5.1. Pre-processing impacts minimally the expected signal defined within the ROI (right). [112].

as a waveform with a discrete time sampling. Significant electronic noise or *cross-talk* between neighboring strips may distort the expected signal and introduce external noise contributions. During pre-processing, this type of noise and detector artifacts are identified and filtered in order to maximize the signal-to-noise ratio and extract the physical signal more efficiently. In order to achieve that, pre-processing is composed of three main operations, which are implemented in a configurable LArSoft *Service* and can be executed in sequence or independently activated:

- **ROI identification.** The regions of interest on the signal are identified using a simple threshold logic to define the starting and ending time of the signal. The values of the start-stop thresholds are externally configurable, but they are in general rather loose at this step in order to maximize the signal region selected.

- **Low pass filter.** Most of the noise comes from sources which usually have relatively higher frequencies than the signal. For instance in the $3\times1\times1$ m^3 prototype operation, one of the main sources of noise was originated from the high-voltage power supply with frequencies around 1 MHz. More detailed about noise during the $3\times1\times1$ m^3 operations is described in Chapter 4. Using a Fast Fourier Transform algorithm (FFT) it is easy to convert the recorded signal from the time to the frequency domain. The FFT implementation used is the standard configured within the LArSoft framework. Once the signal is described in its frequency domain, it is convoluted with a Low pass filter function which removes all the frequencies above the cutoff frequency of the filter. The filtered signal is then reconverted to the time domain using an inverse FFT algorithm.

- **Coherent noise removal.** The main task of the coherent noise removal is to reduce periodic noise patterns that are seen identical on large sets of readout channels. In this case, the noise filtering is done directly on the waveform expressed in its time domain and relies on the signal ROIs previously identified. The idea behind this method is that

the noise structure in terms of amplitude, frequencies, and phase is repeated identical on groups of correlated channels, as visible in the left picture of Figure 3.6. The noise removal procedure acts as a baseline subtraction across groups of channel for any given time sample. The signal ROIs are excluded from the calculation of the baseline mean. The coherent noise filter is implemented in a LArSoft service, which requires to be configured with an array of correlated groups of channels.

A more detailed discussion about the nature and the performances of the noise removal is given in Chapter 4. Even if the signal processing strategies previously described were mostly developed for the data reconstruction for the $3\times1\times1$ m^3 prototype, the LArSoft modules implemented for this goal are able to operate with any given dual-phase geometry, and thus easy to be reconverted and applied on the ProtoDUNE and DUNE detectors.

3.5.2 Hit finder

The second step in reconstruction is the identification of individual energy deposit within the waveform. This procedure is called hit finding and it is probably the most important step of the reconstruction chain. Physically, a hit represents the signal produced by the electrons drifted from a specific location inside the active volume to the anode strips. We are interested in reconstructing with great precision the *integral* and the *peak time* of a hit. The first provides an information of the charge contained in the hit, and thus of energy deposited by the particle inside the LAr volume, while the second gives the drift, x, coordinate of the energy deposit. The y and z locations are straightforwardly reconstructed knowing the strip number on which a given hit is found.
The first step of the hit finding strategy consists of identifying peaks within the ROIs defined in the waveform. The *peak finding* is performed using a simple threshold logic to define the position and size of peaks in the waveform. According to criteria of distance and relative height, peaks may also be grouped together.
Each peak is then fitted with a function f_h which for isolate peaks has the form:

$$f_h(t) = \frac{e^{\frac{t-t_0}{\tau_1}}}{1+e^{\frac{t-t_0}{\tau_2}}}. \tag{3.4}$$

Groups of nearby peaks are instead fitted together with:

$$f(t) = B + \sum_{h}^{peaks} f_h(t), \tag{3.5}$$

where the parameters of the fit are t_0 (the hit arrival time), A and B (a normalization and shift factor respectively), and τ_1 and τ_2, the fall and rise time constants. Hits belonging to the same group are forced to have exactly the same parameters to facilitate the convergence of the fit. Important quantities which are extracted from the fit are:

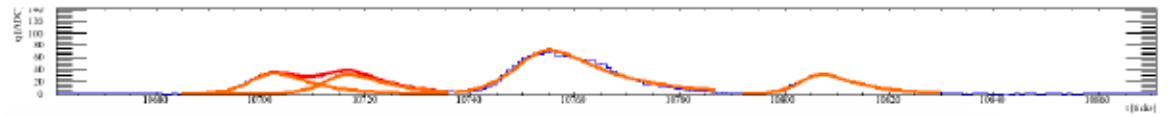
- the hit integral, which gives information of the deposited charge;

- the hit peak time (t_0), which gives the position of the charge deposit along the drift;
- the hit time resolution (or width), calculated as the FWHM of the hit, which represents the error on the drift position of the hit;
- the multiplicity, which counts the number of hits found in the same group.

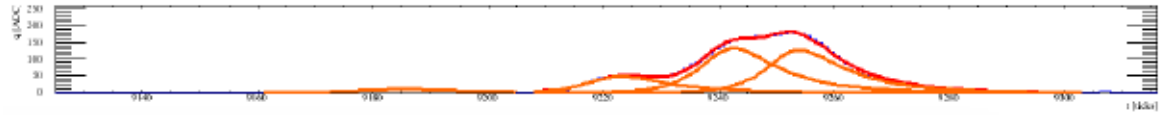
After the fit, the χ^2 is computed and if it is higher than a certain value the waveform is fitted again assuming the presence of one more peak which was not identified during the *peak finding stage*. This operation is repeated up to three times and the original fit is kept unless an improvement of the χ^2 value is found.
There are several advantages to justify the use of a fit-based hit finder. The most important is that the fit helps in assigning the correct integral and time value to each energy deposit, even when those have occurred very close in time. This is particularly useful to resolve better the vertex regions and large electromagnetic showers. An example for this is shown in Figure 3.7a, where the complex waveform identified on the given strip, is well fitted by a sum of several individual hits, whose contribution can be isolated and used to increase the resolution of the TPC in busy regions with many overlapping deposits such as around a neutrino interaction vertex. Refitting also helps to identify a secondary contribution which does not produce evident peaks on the raw waveform (Figure 3.7b).
Additional criteria test the number of peaks found in each group and the fluctuations on the waveform in order to isolate potential troublesome regions for the hit finder, such as along vertical tracks. In this case, the fitting stage is bypassed and a sequence of evenly spaced hits is saved instead. Figure 3.7c shows an example of this on one of the waveforms collected with the $3\times1\times1$ m^3 prototype.
The hit finder algorithm based on the technique here described is implemented in a LArSoft module and used as standard choice for the hit reconstruction in dual-phase LArTPCs. It is important to stress on the fact that both for signal pre-processing and hit finding are used specific algorithms studied specifically for dual-phase LArTPCs. The shape of the signal formed on the dual-phase collection planes is very different from the signals waveform registered on the induction planes of a single-phase LArTPC. Hence, the necessity of dedicated tools different from what is currently used on single-phase detectors.

3.5.3 2D clustering

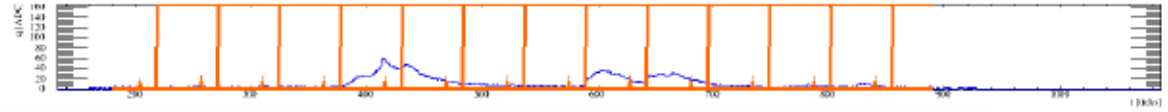
This steps of the reconstruction consists of associating hits from neighboring strips into two-dimensional groups, or *clusters*. The clustering associates the hits from the same particle using similarity criteria such as proximity in space and time, and charge similarity. It allows therefore to reduce the information redundancy from each two-dimensional projection and to mitigate the impact of fake hits reconstructed from noise.
There are several clustering strategies which have been developed for LArTPC event over recent years. The one used in the reconstruction chain for both the $3\times1\times1$ m^3 data and the DUNE far detector here proposed is called TrajCluster [113] and it was introduced for the first time in support to the data analysis of MicroBooNE [114], LArIAT [115], and ArgoNeuT [62].



(a) Hit finder



(b) Refit



(c) Fit failure

Fig. 3.7 Results of the hit finder in the dual-phase DUNE LArTPC. The ADC signal is in blue, while the function obtained from the fit is in orange. Three examples are shown: the fit of groups of peaks (a), the result of the refit (b), the hit fitter failure (c).

TrajCluster aims at reconstruction the projected trajectory of a particle. For each trajectory, the algorithm starts with identifying a small group or "seed" of nearby hits. Additional hits are attached to the leading edge of the seed cluster if they are similar to the hits already included. The similarity requirements are several, but the two most important requires a small impact parameter between the hit width and the projected trajectory, and a relative charge lower than a custom threshold.
The process is iterated until stopping conditions are met:

- lack of hits or too large number of wires without any signal hits;
- abnormally high or low charge fluctuation;
- presence of a large track kink or a 2D vertex

Once all the hits have been considered for clustering, the algorithm allows to perform a second or even a third pass, with different parameters settings. This allows to use looser or tighter parameter criteria to set the hit similarity constraints and thus to target the characteristics of certain topologies at each pass. Further algorithms are available within TrajCluster which may be run at the end of the entire process to merge or split the clusters found.
Since no constraints on the number of planes or wire spacing are imposed, the same implementation of TrajCluster within LArSoft made for single-phase TPCs has been used. Efforts have been put in identifying the correct threshold settings and cuts in order to exploit all the potentiality of the algorithm given the finer pitch between strips, the different shape of the hits, and the larger variation of charge due to the longer drift.

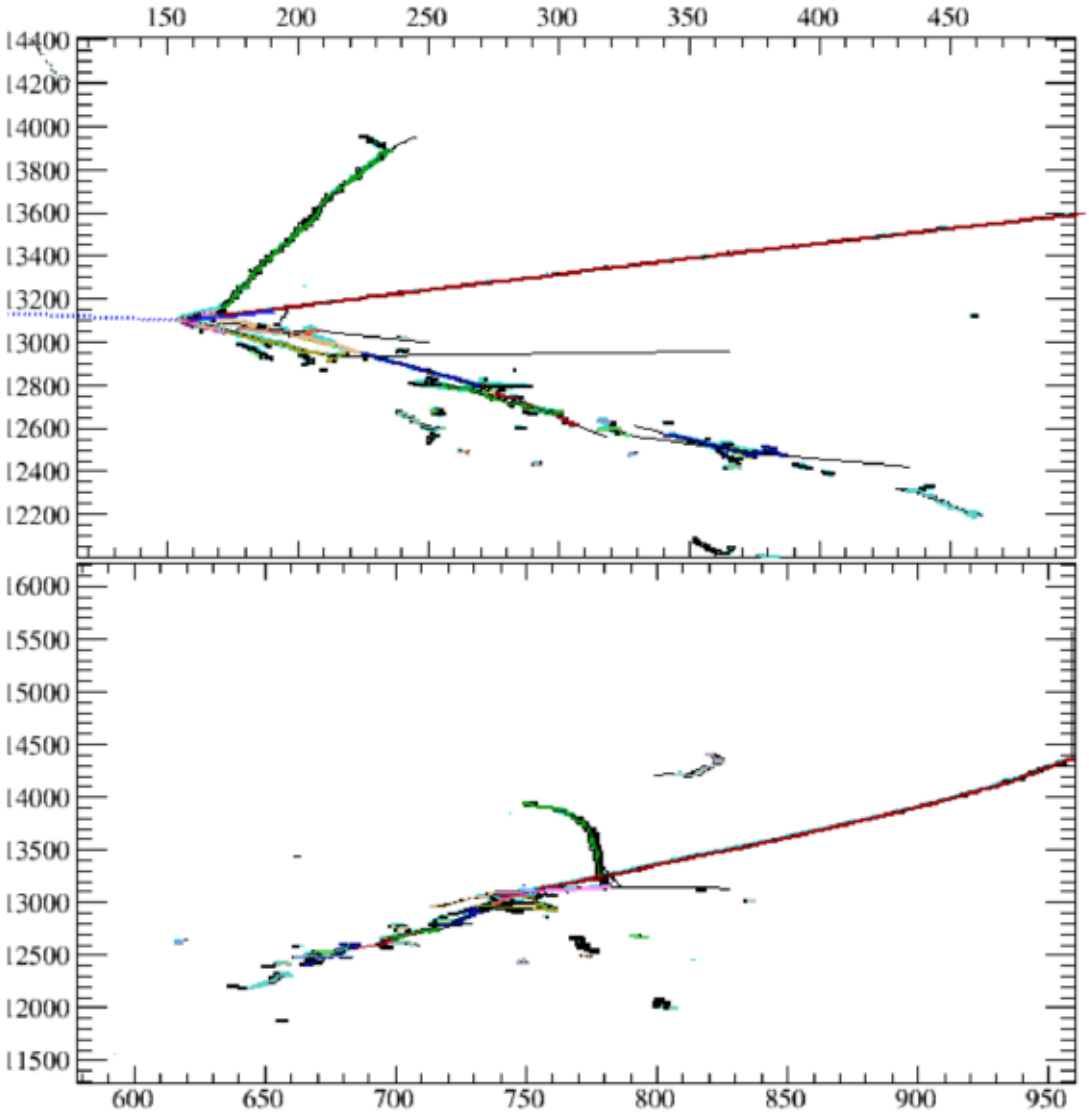


Fig. 3.8 The two 2D projections of a ν_μCC event at the dual-phase DUNE far detector after the use of PMA for 3D view matching. The hits assigned to the same tridimensional trajectory are colored in the same way on both views. When no match is found, hits are colored in black

3.5.4 3D matching

Once the two-dimensional clusters are formed, the final step of reconstruction consists of matching the energy deposit between the two views to build a tridimensional representation of the event. This step is important because it allows to extrapolate physical quantities, such as the energy deposited per unit of length by the particle, and its momentum, which are both fundamental for the event classification and energy reconstruction.
The method used for cluster matching and tridimensional trajectory reconstruction is the *Projection Matching Algorithm* (PMA) [116]. This algorithm was first deployed for precise tridimensional trajectory fitting in ICARUS [117], and successively implemented in LArSoft and extensively studied in many single-phase LArTPC, including LArIAT [115], and MicroBooNE [114]. The general concept behind the algorithm is to find optimal matches between bidimensional clusters found on different planes and provide a best tridimensional trajectory representing the path of a charged particle inside the LAr volume. This is done in various steps:

- First PMA tries to identify potential cluster matches between views using start-end time similarities.
- For each potential match, a tridimensional trajectory formed by a polygonal line with an increased number of segments. The placement and direction of the segment is optimized by minimizing a specific cost function based on the 2D distance between the hits and the bidimensional projection of the 3D trajectory formed.
- At the end of the iterative process, a score is assigned to the final trajectory, and used to determine the best cluster match. In LArTPC with three views, only two views are used for the matching and the score is calculated from the cost function of the projection of the trajectory on the third plane. If only two views are available the score is simply the total cost function across the two projection.
- A final score based on the mean squared errors of the 2D distances between the hits and the found trajectory is assigned to each track, and used to determine the best match
- The formed trajectory is used to assign an unique tridimensional position to the hits on the matched clusters. In this sense PMA is particularly advantageous compared to other reconstruction techniques, since it does not impose any requirement of drift time similarity between hits belonging to different projections. This allow to produce a better assignment between the hit and the instantaneous direction along the trajectory and have a more accurate estimation for the dE/dx along the track.

In the PMA implementation for dual-phase in LArSoft, the procedure previously outlined is done independently for the signal collected by each CRP, and then the results from different CRPs are merged together. The reason for that is to reduce the number of computation required, since the algorithm must loop several times over all the clusters found in order to identify the best matches. The obvious drawback of this method is that a local failure of the reconstruction for a single CRP may affect the whole event. Common known failures of the

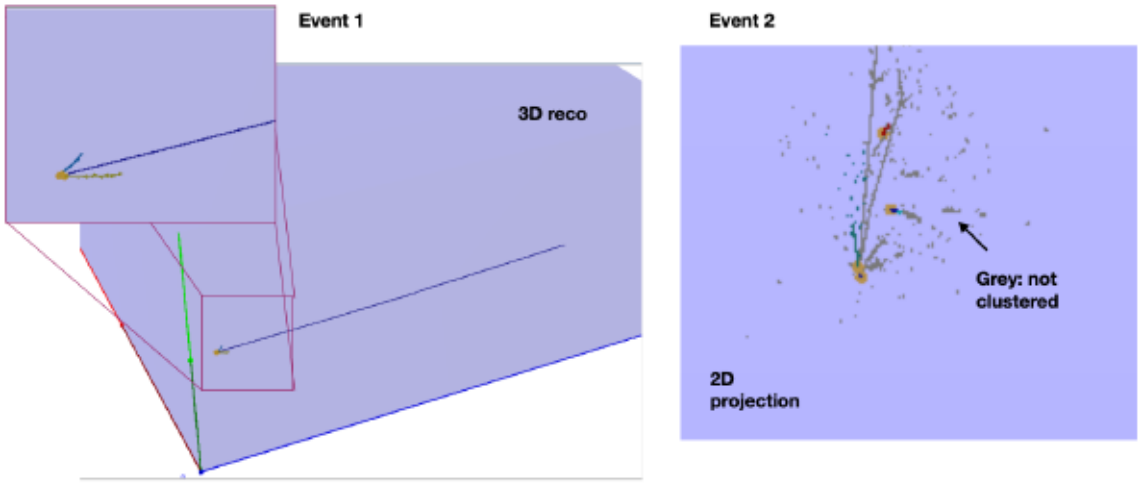


Fig. 3.9 The performance of Pandora on a ν_μCC (left) and a ν_eCC (right) event at the dual-phase DUNE far detector. The event on the left was fully reconstructed in 3D, while for the event on the right it was not possible to find any good match between views.

algorithm are described in [117]. For some orientations of the trajectory the 2D optimization of the fit is very complicated. This happens, for example, with trajectories parallel to one of the readout direction or to the drift, since there is not a sufficient number of hits on the projection to optimize the fit. These reconstruction inefficiencies are extensively illustrated in Chapter 5 where PMA is used for the reconstruction of neutrino events.
An example of a ν_μCC event in the dual-phase DUNE far detector successfully reconstructed with PMA is shown in Figure 3.8. The hits matched to the same 3D trajectory have the same color on both views. Black hits were not associated to any trajectory. From Figure 3.8, it is possible to observe that PMA is able to associate correctly track-like trajectories, but it is not able to reconstruct more complex topologies such as electromagnetic cascades.

3.5.5 Pandora

The Pandora software development kit (SDK) [118] was created to address pattern-recognition problems in fine-granularity detectors. Pandora differs from TrajCluster and PMA because it is capable to use several algorithm to cluster and match between views simultaneously. Pandora addresses first to straightforward topologies in order to reduce the complexity of the event, while it tackles with more complicated configurations in a second time. Choices previously made can also be reverted, if new information provide a more reasonable reconstruction of the event. Pandora has been already applied to many LArTPC data such as neutrino and cosmic ray events in the MicroBooNE detector [119] and it is currently in use for the study of the beam interaction in ProtoDUNE-SP.
Until recent times, it was not possible to use Pandora also for dual-phase LArTPC, due to an intrinsic assumption of a specific detector geometry expect in the LArSoft *module* responsible to interface with the Pandora SDK. At the time of writing, some tests of the

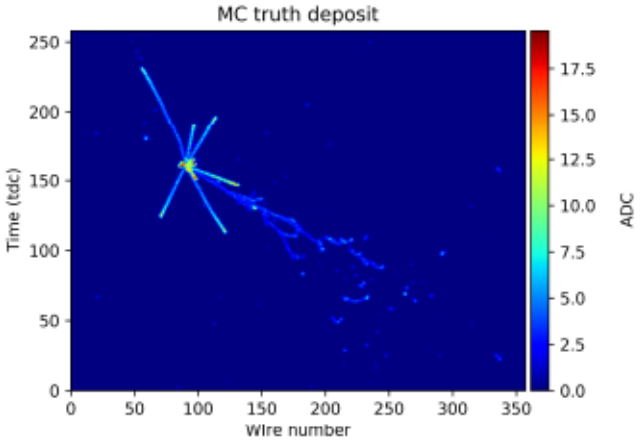


Fig. 3.10 Example of the ADC map of a simulated ν_μCC events. For a human eye, *Track*-like energy deposits are well distinguishable from *shower*-like deposits.

algorithm on dual-phase neutrino events reveals encouraging performances for neutrino events with simple topologies, such as in Figure 3.9 (left). A more systematic understanding of the algorithm is however necessary, in particular to understand the reason of its failure in complex events featuring large electromagnetic cascade, like in the example shown in Figure 3.9 (right). Efforts in this sense are ongoing, with ultimate goal to optimize Pandora for the full-reconstruction of neutrino events at the dual-phase DUNE far detector and for the analysis of the data from ProtoDUNE-DP.

3.6 Convolutional Neural Network for track/shower separation

One of the main advantages of LArTPCs is the possibility to form high resolution images of the charged particles contributing to the event. According to their interaction with matter, particles appear differently inside the TPC. Figure 3.10 shows an example of a ν_eCC interaction at the DUNE dual-phase far detector. The straight lines belongs to minimal ionizing particles (*mips*) such as muons, protons, and pions, while the confused dispersed energy deposits are signatures for electron or photon induced electromagnetic (EM) cascades. We refer to the former as *tracks* and to the latter as *showers*.
The possibility to distinguish between tracks and EM activity based on their topology brings a significant contribution to the DUNE physics goals. For example the presence of an electron induced cascade at neutrino vertex is a powerful discriminant between ν_μCC and ν_eCC interactions, where an electron is always expected in the final state. Also a precise identification and reconstruction of the Michel electron spectrum is important for the detector calibration [120]. Moreover, the capability to separate electron and hadronic contributions in ν_eCC events may allow to perform independent corrections to the reconstructed energy and increase the energy resolution.

Even if the distinction between tracks and showers is straightforward for a human eye, this task remains very challenging for the automated pattern recognition algorithms described in Section 3.5. Machine Learning (ML) algorithms and in particular Convolutional Neural Networks (CNNs) [121] are promising tools to solve this problem. CNNs are Deep learning algorithms which are widely used for computer vision tasks [122] and which are becoming a common tool for pattern recognition also in high energy physics [123]. This technology has been applied with success for neutrino classification in NOvA [124] and, due to the possibility to operate directly on image inputs, it also an excellent tool for image classification in LAr [125].
This section describes the work done during this thesis to develop a CNN for pixel classification to discriminate tracks-showers-Michel electrons and background at pixel level in LArTPC images. The algorithm is designed be included into the dual-phase automated reconstruction chain and to tag each hit in order to improve the performances of higher level algorithms. In particular Section 3.6.3, Section 3.6.2, and Section 3.6.4 describe respectively the network architecture, the preparation of the input and the training strategy. The performances of the algorithm, its robustness and generalization capabilities are instead described in Section 3.6.5, Section 3.6.6, and Section 3.6.7. The source code with the model architecture definition and training is available on GitHub[2].

3.6.1 CNN basic principles

Standard ML algorithms consist of a network of interconnected layers of nodes. The output of each node is a weighted sum of the output of the nodes in the previous layer, plus a bias term operated typically by a non-linear *activation function* [126]. The algorithm is trained to learn what is relevant for the classification on a prepared of input whose output, or *true* label, is known. This process is called *supervised learning*. During the learning process, the algorithm iteratively tunes weights and biases of each node in order to minimize an error function representing the distance between the output of the ML algorithm and the true label. Figure 3.11 outlines the basic architecture of a CNN algorithm. The main difference between CNNs and other ML algorithm are layers made up of of two-dimensional kernels introduced in order to use images as input instead of pre-selected arrays of features. The kernel layers learn how to highlight the relevant features for the classification directly on the image and feed their output to the classification layers which aggregates those information and provide the final output.

3.6.2 Data preparation

In ML algorithms, the data used to train the model can be considered as a part of the model itself. For this reason, it is very important to carefully define the problem and prepare the data in an appropriate fashion. LArTPC data are well suited for Deep-Learning applications, since they feature a high-resolution image of the event, whose pixels are given by the wire spacing and the size of the time samples. Our CNN algorithm has to be able to separate

[2]https://github.com/ascarpel/FeatureLabelling

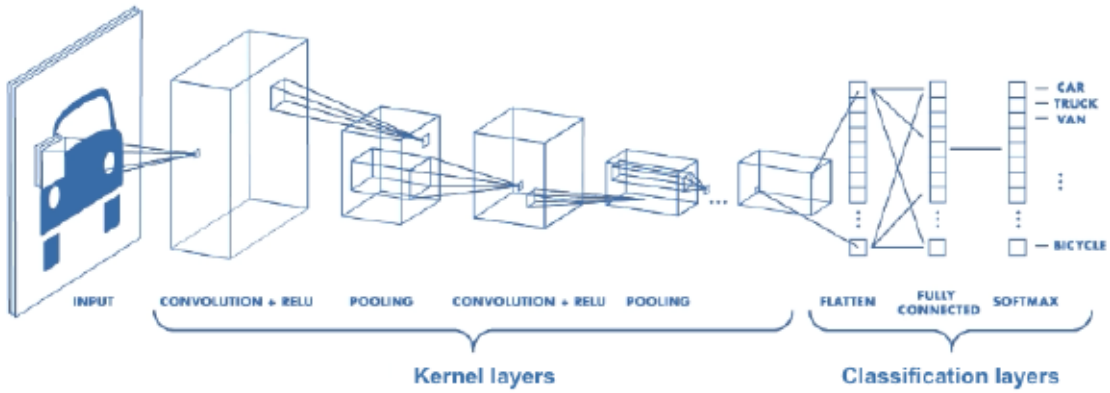


Fig. 3.11 A sketch of a typical CNN architecture for the classification of types of vehicles. The 2D input image is first processed by kernel layers which extrapolate the features necessary for the classification which are successively aggregate by classification layers. [127]

tracks and showers at pixel level and must be sufficiently fast to be included into the full reconstruction chain without significantly increasing the computational resources demanded. This is achieved by classifying each pixel within the context given by a small "region of interest" rather than the full picture. The approach adopted in data preparation is schematized in Figure 3.12 left. A context *patch* of 20x20 cm^2 (68x66 pixels) is built around each pixel of the image. The energy deposit of central pixel of each patch is associated to the corresponding GEANT4 particle to determine the true label of the input. The inputs are divided into three main classes and one subclass:

- Tracks;
- Showers;
  - Michel electrons as a subclass of the shower class;
- Background, which classifies pixels with small energy deposits but which are not associated to any specific particle, as for example in noise pixels.

An example of one patch for each class with its true label is shown in Figure 3.12 (right). The actions performed to downscale the larger image such as the one in Figure 3.12 left into a set of relevant patches and true labels are described in the two following paragraphs. The first is the *event dumping*, where is created a mask for each LAr image associating each pixel with the corresponding ADC value, GEANT4 pdg code, and true energy deposit. The second is the *patch production*, where the relevant regions of interest inside each image are identified, selected, labeled, and included into the dataset used for training the algorithm.

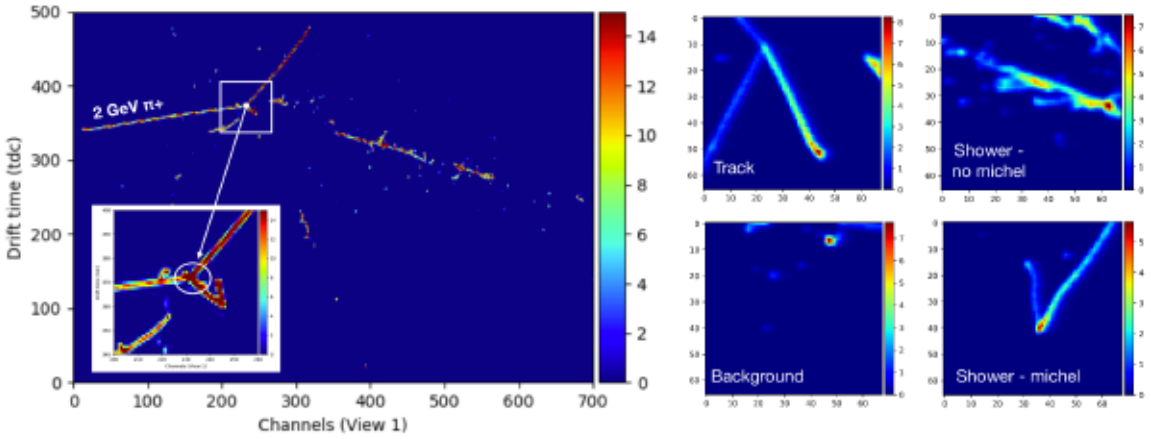


Fig. 3.12 Right: example of the data preparation on a simulated 2 GeV pion interaction. A patch (identified by the white box) is build around a pixel of the event. The label for the patch is given by the true information associated to the central pixel within the patch. Left: an example of patches of the four possible labels.

Event dumping

This first stage of the data preparation is performed by a dedicated LArSoft module, in order to access all the information stored after the particle propagation in GEANT4. The main actions performed by the module are the following:

1. Creating equally sized pixels in the time and wire number. For dual-phase the wire pitch is 3 mm and the sampling time difference is 0.4 µs. At nominal drift velocity $v_{drift} = 1.6$ mm/µs the number of tdc units necessary to travel a distance equal to the wire pitch determines the downscaling factor of 5 applied along the time dimension. The pixel intensity is the mean of the five corresponding ADC counts in the downsampled 5 tdc bins.

2. A calorimetric factor is applied to the ADC count to correct for the effect of the electron lifetime and the amplification in Argon vapor.

3. Since the *activation functions* of the CNN tend easily to saturate to 0 or to 1 making the neuron ineffective for the learning process, an appropriate scaling is perform in order to constrain the ADC values in the [0, 15] range. The pixel map with the new ADC values are saved in 2D histograms.

4. Along with the ADC pixel maps, two more histograms are created. The first contains an event mask with encoded the GEANT4 PDG codes associated to every pixel and the physical process contributing to it. The second simply stores for each pixels the true energy deposit of the associated GEANT4 particles. Both histograms are used in the following step of the data preparation to select the patches for training in interesting parts of the image (for example in proximity of a decay vertex).

Table 3.1 The number of events and the corresponding number of patches used to composed the training, validation, and testing samples. The fraction of classes present in each sample is also reported. Patches with Michel electrons are also counted as showers.

	Training		Validation		Test	
	Pions	Protons	Pions	Protons	Pions	Protons
Events	20000	19900	2450	2550	250	250
Patches	2M	1.9M	245k	250k	25k	25k
Fr. tracks	53%	66%	54%	66%	52%	65%
Fr. showers	41%	27%	40%	27%	41%	29%
Fr. background	6%	7%	6%	7%	6%	6%
Fr. Michel	2%	1%	2%	2%	1%	2%

Production of patches

During this step, the information from the 3 histograms previously made are aggregated to downscale the image to a set of patches and the corresponding labels. First a 66x68 patch is formed around each pixel of the ADC map. Then, the GEANT4 code, the physical process and the true energy inside the patch are looked up and used to form the true label. This information is also used to balance between the different type of patches and to reduce the redundancy by selecting only patches in proximity of interesting areas of the event.
The total number of patches produced is separated into three samples that will be used to train the network and test the performances. The training and validation samples are used for the learning process and to monitor its evolution. The testing sample is completely independent of the learning process and it is used to study the final performance of the model. This is done to minimize the possibility to introduce biases from learning and assess the generalization capability of the classifier. The composition of the three samples is reported in Table 3.1. Input patches are formed by downscaling pions and protons events with momentum between 2 to 5 GeV/c^2 and isotropical angle distribution. Patches from both views are equally considered.

3.6.3 Network architecture

The difference between CNNs and other ML algorithms is in their architecture. CNNs use one or more layers made up of layers with 2D kernels able to identify specific features inside the input image. Inner layers of various dimensions are able to aggregate information between kernels.
The complexity (or depth) of Deep learning models is determined by the number of kernels in the first layers and by the number of inner layers. Very complex models are in general able to better identify patterns in data and produce more accurate classifications. However, they are also more difficult to train and less versatile for different classes of problems. The model used for this work is not very complex. This choice was made in order to include it into the

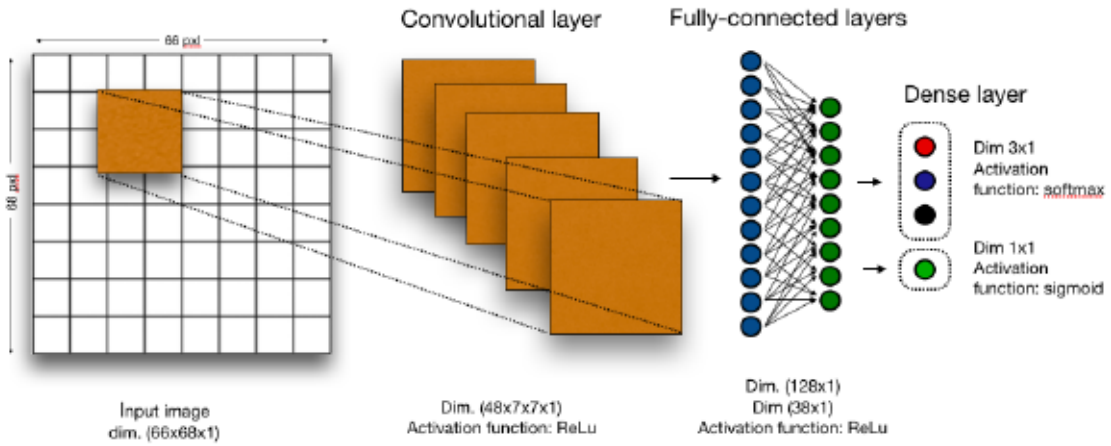


Fig. 3.13 The network architecture implemented for the track/shower separation in a dual-phase LArTPC.

automated reconstruction of Section 3.5 without a significant increase of the computation time required. The input to the CNN is a 66x68x1 tensor. The first two dimensions of the tensor are the height and width of the input image in pixel, while the third dimension holds the intensity of each pixel. The inputs may have one among the three true labels: track, shower, background, and if they are classified as showers, they might be also Michel electrons.
Figure 3.13 sketches the architecture of the CNN which is made up with the following layers:

- **Convolutional layer** . The network has only one convolutional layer with 48 7x7 kernels. The activation function for the nodes in each kernel is the rectified linear unit (ReLu) function [128].
- 1st **Fully connected layer**. The first of the classification layer is a monodimensional layer with 128 fully connected nodes. Fully connected means that each nodes receive a linear combination of inputs from all the nodes of the previous layer. The activation function is ReLU.
- 2nd **Fully connected layer**. The second classification layer counts only 38 nodes. It is fully connected and the output activation function is ReLU.
- **Dense layer**. The last layer is called dense layer because it receives the outputs from the previous layer and form the probability for the for classes. This layer has dimensions 3x1, where three nodes are for the three main classes. The activation function for those nodes is the *softmax* function [129], a popular function in categorical classification algorithms because it returns an probability distribution consisting of one probability for each class. The remaining node is for Michel classification. The activation function is a *sigmoid* [126].

The architecture previously described has been worked out from previous studies carried out for the single-phase detector [130]. The size of the inputs, the number and size of the layers and some of the activation functions have been modified in order to improve the network performances, although no systematic optimization was performed.

3.6.4 Training

It is during the training process that the CNN learns from data which features are relevant for the classification. During the training the output of the model is compared with the true label, the error of the model is computed and the weights at each node are updated in order to ensures the convergence of the error to a minimum. The function describing the error of the model is called *loss*. The *loss* chosen for the 3 class output l_{3class} is the *cathegorical crossentropy* function [131]:

$$l_{3class} = -\sum_{i}^{C} t_i log(s_i), \tag{3.6}$$

where s_i and t_i are respectively the model prediction and the expected true label for the i-th class of C. The name *categorical crossentropy* found its origin from information theory, but for classification problem, this function can be naively seen as the log-likelihood to get the i-th class having probability t_i with a frequency s_i.
For the Michel electrons output, the *loss* chosen, l_{michel}, is the mean squared error. The total model *loss* is obtained with a weighted sum of l_{3class} and l_{michel}, with an higher weight given to l_{michel} because of the relative class imbalance.
The network is optimized during training using the Stochastic gradient descent (SGD) technique [132]. With SGD, the weights of the network are iteratively adjusted of a quantity proportional to the gradient for the loss function averaged over batches of 286 images. This proportional constant is called *learning rate* and control the speed at which the model is learning. We trained the model with a fixed learning rate of 0.01. The quality of training is monitored by using the classification accuracy, defined as the ratio between all the images and the correctly classified images.
A *dropout* [133] regularizator is used to prevent overfitting. The *dropout* algorithm randomly set to 0 a fraction (0.2 for our model) of the nodes in each inner layer. In this way, the importance that the model give to random fluctuation in the dataset is reduced. Since the *dropout* is not applied on the validation sample, the computed accuracy and *loss* for this sample may be higher than in the training sample
The evolution of the *loss*, and the classification accuracy of the model are shown in Figure 3.14 (left) and right as function of the 50 epochs of the training. One epoch defines the time needed to consume all the images in the training sample. Training was stopped before the appearance of overfitting.
The network scheme in Section 3.6.3 and the aforementioned training strategy have been implemented using Keras [134] and trained on an NVIDIA-GPU GeForce GTX 1080 with 8 GB of memory. With the training and validation samples of Table 3.1, training of one epoch takes about 15 minutes.

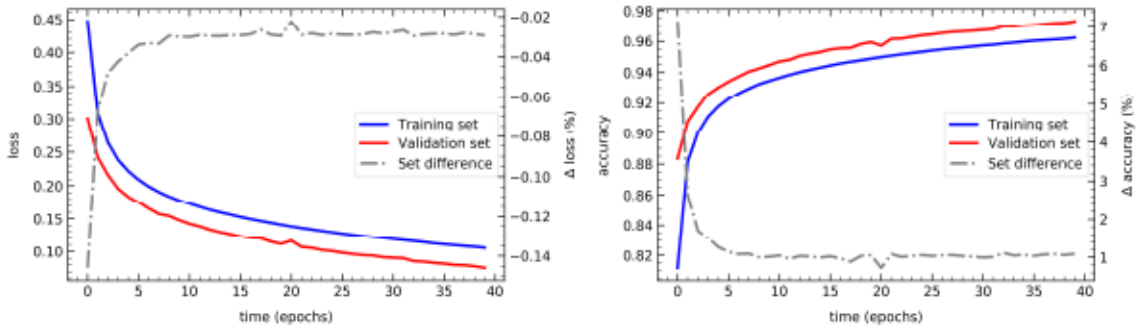


Fig. 3.14 The evolution of the total *loss* (left) and accuracy (right) as function of the training epochs. The dotted gray line is the difference (Δ) between the score on the training and the validation sets.

3.6.5 Performances

Figure 3.14 shows the evolution of the model *loss* and accuracy on the training and validation sample as a function of the training epochs. The maximal accuracy achieved by the model over 50 training epochs is 0.98. Accuracy is not a metric able to describe alone the performances of the model made and the and quality of the classification. A more useful tool to asses the performances of a classifier is the *Confusion matrix*. In a *Confusion matrix*, each row represents the instances in a predicted class, while the rows represents the instances in the true class. This representation allows to visualize and quantify if and how the model is confusing between classes.
Since the *softmax* and *sigmoid* response functions of the last layer in the architecture of Figure 3.13 return normalized probability values for each class, the calculation of the correctly classified labels and the correspondent contamination is straightforward:

- For the track/shower/background classification the predicted label is the label which has received the highest probability by the network.
- For the Michel class, a threshold of 0.5 has been put on the output probability from the network to assign the Michel predicted label to the input.

Figure 3.15 shows the confusion matrix for our model when applied to the testing sample of Table 3.1. The number quoted on each bin is the fraction in % of true events labeled as the corresponding row which has been classified by the model as the label in the column.
From Figure 3.15, it is clear that our network is able to well distinguish between track, showers, and background at a level of few %. A confusion lower that 10 % occurs between showers and tracks. This happens because the topological discriminant is not sufficient to ensure a precise separation between the case of δ rays emission (considered as EM activity and thus labeled as shower) and simple fluctuations in the ionization energy deposit.
For Michel electron, classification is less efficient than for the other classes. Michel events are often confused with showers and less frequently with tracks. An accurate investigation of this is outside the scope of this thesis, however a 74% true positive rate for Michel electron

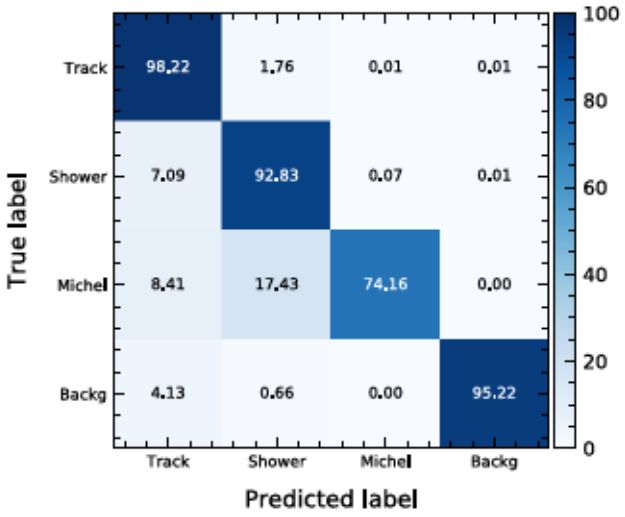


Fig. 3.15 The *confusion matrix* calculated on the validation sample of Table 3.1.

pixels classification is an encouraging results that might be improved simply by increasing the number of Michel events in the training sample.

3.6.6 Robustness of the model

CNN models learn how to classify the input data by selecting relevant features in the training sample. Although sometimes a human eye can understand the reasoning behind some choices, often the model learns at such an abstract level that is almost impossible to find a straightforward correspondence with the physical information of the input. For this reason it is instructive to study how well the model adapts to variations in the input that may occur once the model is applied to unknown LArTPC data.
For this this purpose, we binarize the output of the model into just track/shower and ignore the classification into Michel electron, or Background. The output probabilities for track and shower are consequently renormalized. We define the track hypothesis as signal and the shower hypothesis as background (to be distinguished form the background label previously mentioned). A selection threshold between the two hypothesis is set to 0.5. The model performances are described in terms of selection *efficiency* and *Purity*, which represents respectively how many true tracks are selected in the sample and the degree of contamination from the background. In more quantitative term, efficiency (E) is computed as:

$$E = \frac{TT}{TT+FS}, \tag{3.7}$$

with TT and FS being respectively the number of true-tracks and the number of false showers selected. A true-track is an event having track label, which has been correctly identified by the model, and consequently a false-shower is an event having track label which has been classified as shower. Their sum represents the total number of tracks labels in the sample

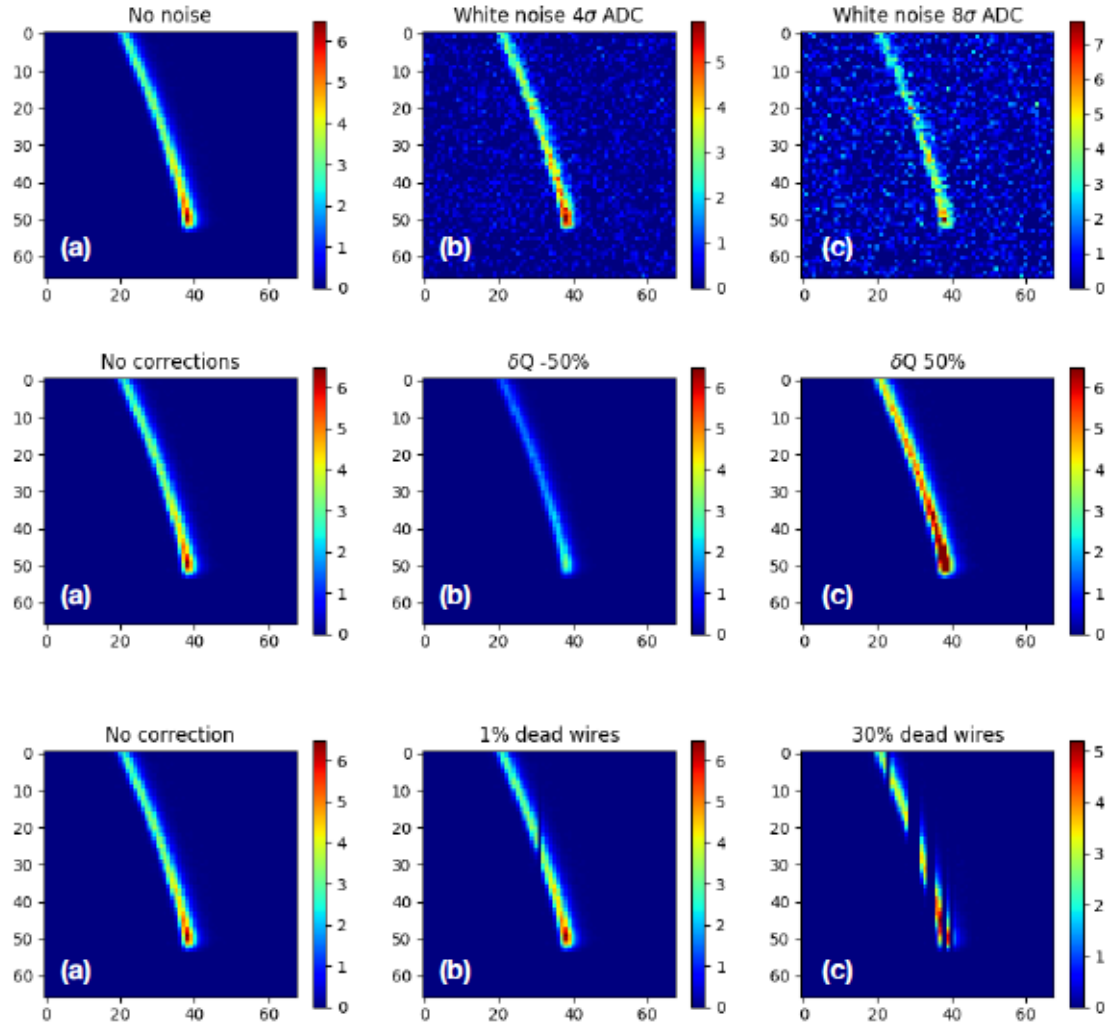


Fig. 3.16 Example of how random noise (top), charge uncertainties δQ (center), and dead wires (bottom) transform one of the patches from the validation sample.

Purity, P, is defined as:

$$P = \frac{TT}{TT + FT}, \tag{3.8}$$

where FT is the number of false tracks in the sample. The sum $TT + FT$ represents the number of events of the sample which are declared of being tracks. In LArTPCs the input image might be altered in several ways. We consider three common effects to prove the model robustness: random noise variations, uncertainties on the charge calibration, and local density of dead wires.

Random noise variation

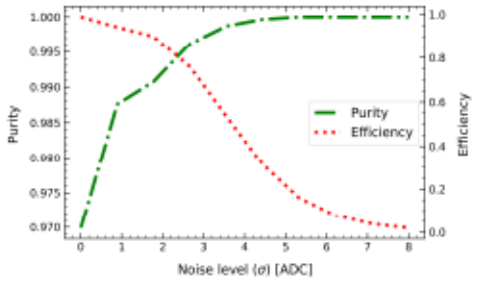
Random noise is a common peril in image processing. For this test, a random noise following a Normal distribution with mean equal to zero and standard deviation σ has been injected in the validation samples before the classification. Figure 3.16 top shows how an example patch is transformed by injecting noise with σ equal to 0 (a), to 4 ADC (b), and to 8 ADC (c). Purity and efficiency are calculated on the validation sample as function of the noise σ. Results are shown in Figure 3.17a. As expected the efficiency of track selection drops drastically with the increase of the noise level, because the model uses the confusion of the image to determine whether the input is a track or not. P increases with the increase of noise, since it is sensitive to the number of tracks erroneously classified as showers. In presence of a high level of noise, the model rarely confuses a shower for a track, while the opposite case is more likely.

Uncertainties on the charge calibration

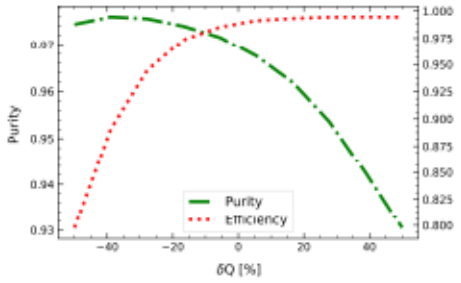
According to Equation 3.1, the pixel intensity is corrected with a factor C_{calo} before it is provided to the algorithm. The correction is done in order to homogenize charge variations across the drift, however it relies on a prior knowledge of the LAr purity and the detector gain. The uncertainties on the knowledge of C_{calo} are introduced as a systematic relative error (δQ) on the charge value assigned to each pixel. The central row of Figure 3.16 center shows how the example patch (a) is modified by adding a negative (b) and positive (c) correction δQ of 50 %. *Efficiency* and P have been calculated for the validation sample corrected for different values of δQ varying in a symmetrical range between $\pm$ 50 % (as shown in Figure 3.17b)
In Figure 3.17b, purity and efficiency show opposite behaviors. The model relies strongly on charge information in order to separate tracks from showers. Let us consider for example the case of a patch containing only one straight portion of a muon. δ rays are commonly emitted by those types of tracks and in some projections they are not distinguishable from the main trajectory. Therefore the CNN algorithm can solely rely on charge information to distinguish between this case and the case of a clean track. Applying a relative correction on the charge value therefore has a huge impact on the algorithm when it tries to classify this type of events. If the correction is negative, the prediction of the track case is encouraged, increasing therefore the number of false-showers predicted and increasing the number of false-track. In the opposite case, when the correction is positive, the preferred outcome is reversed: false-tracks tends to decrease and the true showers to consequently increase.
It must be stated however, that the range of $\pm\delta Q$ considered for the plot of Figure 3.17b has been chosen to exaggerate its impact on the classification. In DUNE δQ is expected to be constrained at the 2% level, thus the effect here described is expected to scarcely impact on the final performances.

Density of dead wires

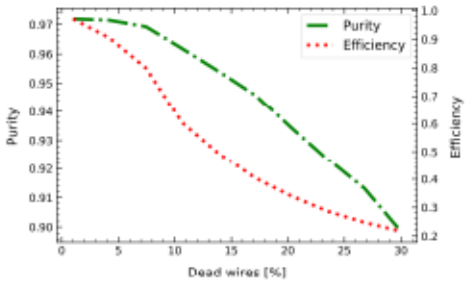
During detector operation, loss of information from some channels may occur. The dead wires look like empty gaps in the images and they may have an impact on the performances



(a) Random noise level.



(b) Calibration uncertainties.



(c) Density of dead wires.

Fig. 3.17 Impact of random noise (a), charge calibration uncertainties (b) and local density of dead wires (c) on the classification performances of the neural network model expressed in terms of *Purity* and *efficiency*.

of the neural network classifier. Figure 3.16 bottom shows how a local fraction of dead wires transforms the example patch. In analogy to the previous case, the input have been altered shading an increasing fraction of columns of pixels from 1 % to 30 %, and the performances evaluated. Results are shown in Figure 3.17c.
Increasing the number of dead wires increases the sparsity of the energy deposits in the image, and therefore reduces the amount of information available. This is the reason why in Figure 3.17c both purity and decreases: the algorithm has to base its classification on less information, and therefore the resulting classification does not favor any case in particular but commits a random error between the two.

3.6.7 CNN in the reconstruction chain

One of the main objectives for the shower/track classification algorithm is to be included into the main standard reconstruction at very low level in order to target specific algorithms on specific types of hits. This is done via a dedicated LArSoft module, which is able to build a region of interest in similar fashion as presented in Section 3.6.2 around the peak position of every hit found in the event. Every hit receives a CNN score for each class which is saved along with other hits information. Higher lever reconstruction algorithms, such as 2D clustering, may either use the hit score to target the reconstruction on a certain type of

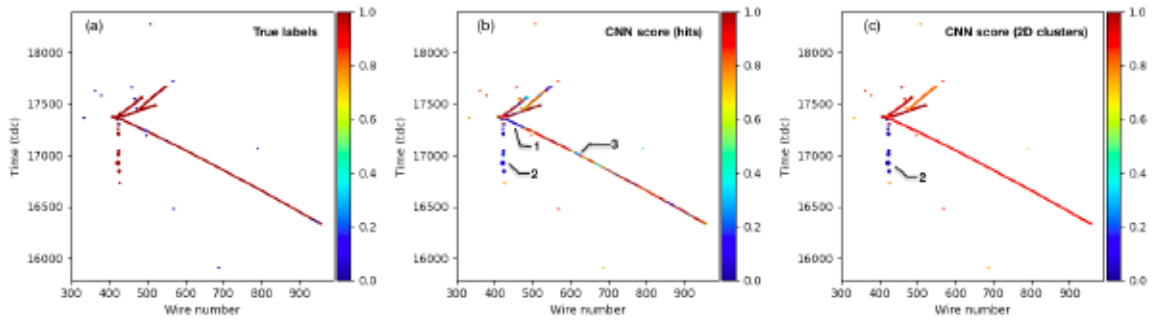


Fig. 3.18 Performances of the CNN model on a ν_μCC event in the DUNE-DP detector. Color scale represents the classification probability: 0 is a perfect shower, 1 is a perfect track. The truth case (left) is compared with the CNN output for the single hits (center) and for clusters of hits (right).

hits or may aggregate the scores of many hits to produce a cumulative classification of the reconstructed object. Figure 3.18 shows a ν_μCC event with hits (Figure 3.18 (center) and clusters (Figure 3.18 (right) classified according to their binary CNN output and compared with the true label (Figure 3.18 (left). From the hit probability map of Figure 3.18 (center) it is possible to identify three problematic areas where classification fails:

1. Vertex area: some hits in proximity of a vertex region might be confused with showers because of the increasing particle density
2. Vertical tracks: the larger amount of energy deposited by vertical tracks on a single wire might trick the classification of this type of hits from tracks into showers
3. Long clean tracks: delta rays are easily confused with energy deposit fluctuations along the track.

Local variations and inefficiencies are often mitigated when the single probability are aggregated together to form a cumulative score for a single cluster object, as in Figure 3.18 (center), where the issues highlighted in 1. and 3. are absorbed in the total cluster score.

3.7 Summary and outlook for reconstruction

The current status of the simulation and reconstruction of events in a dual-phase LArTPC has been presented in this Chapter. The work has been implemented within LArSoft, which is the common software framework shared by most liquid Argon experiments, and in particular by DUNE. Thanks to LArSoft, both simulation and reconstruction are customizable and can be adapted to different detector geometries with little efforts.
A completely functional dual-phase simulation is now in place and its applications are further described in Chapter 4 and Chapter 5.
An automated reconstruction is in place and it is capable to a tridimensional reconstruction of straight track-topologies. A Convolutional Neutral Network is used to separate track-like

and shower-like deposits with a classification accuracy larger than 95% for all the classes. Its robustness against variations of noise, charge, and density of dead wires were also tested, showing overall good performances.
Simulation efforts of the dual-phase are overall good and it will further tuned thanks to data from the two dual-phase $3\times1\times1$ m^3 and $6\times6\times6$ m^3 prototypes. It this one of the goals achieved with this thesis to have an almost entirely automatic reconstruction for dual-phase. However more work is necessary in order to complete the reconstruction chain, in particular for tridimensional shower reconstruction. Efforts in this sense have already started using the Pandora framework on dual-phase events.

Chapter 4

Operation and data analysis of the $3\times1\times1$ m^3 dual-phase prototype

This chapter describes the operation and data analysis of the $3\times1\times1$ m^3 prototype of the dual-phase DUNE LArTPC. The detector was built and commissioned between 2016 and 2017 at CERN and collected more than 500k cosmic ray triggers until late fall 2017. It was the first time a 4 ton dual-phase LArTPC was operated, representing a step of a factor 1000 in volume if compared to previous prototypes, and the first test for the 50×50 cm^2 LEM assembled on an extended readout. [105]. The detector setup and operation are briefly described in Section 4.1.
The primary goals of the $3\times1\times1$ m^3 prototype are:

- The characterization of the detector performances with the extended three squared meters CRP, which requires to obtain an effective gain of more than 15 and 10% resolution to the collected charge from minimum ionizing particles, with inter-LEM fluctuations constrained at the 5%.
- The study of the electron amplification on the 50×50 cm^2 LEMs.
- The validation of the component, their inter-operability, and stability in cryogenic conditions.

Regretfully, some issues at the level of the extraction grid have made it impossible to reach the expected nominal conditions of amplification and therefore accomplish to the main detector purposes in an entirely satisfactory manner. The data taking conditions and the type and quality of data collected are described in Section 4.2.
In spite of the detector issues, the prototype offers the opportunity to test the LArSoft simulation and reconstruction of Chapter 3 on real data from a large dual-phase LArTPC. More broadly, in this thesis we attempt to provide a general characterization of the performances of the $3\times1\times1$ m^3 prototype, bearing however in mind that the results obtained are not representative of the performances of the future DUNE far detector. Particular attention is given to the analysis of the noise, whose characteristics were studied to be introduced in the simulation presented in Section 3.4.2. The work on the dual-phase LArSoft simulation and

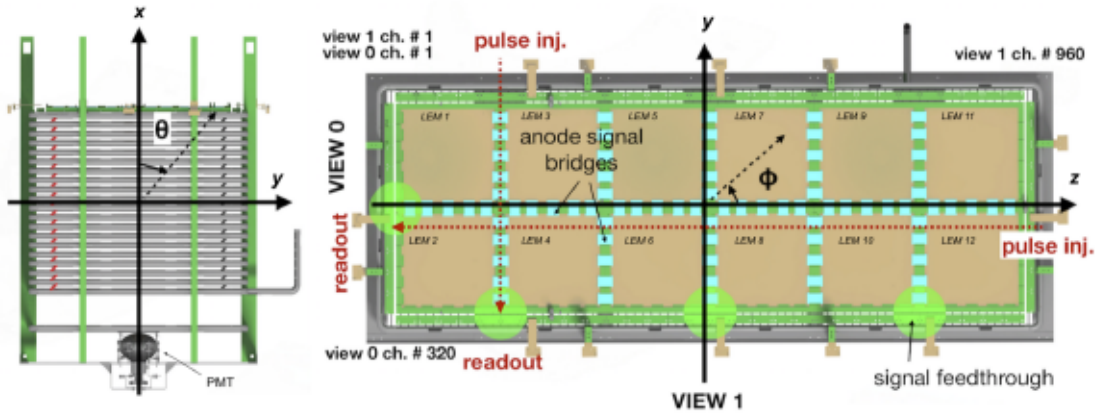


Fig. 4.1 Transverse (left) and top (right) sections of the 3×1×1 m^3 TPC. The drift direction is along x. The CRP is composed by 12 50×50 cm^2 LAS, electrically bridged to form an uniform readout surface of 3×1 m^2. The three meters long strips are parallel to the z direction and constitute View 0; the one meters long strips are parallel to y and constitutes View 1. The polar angle ϕ is defined on the $y-z$ plane. [102].

reconstruction using data-driven inputs is presented in Section 4.3 and Section 4.4. Finally, Section 4.5 illustrates the analysis of the 3×1×1 m^3 data done for this thesis, where we estimate the maximal gain achieved within the 50×50 cm^2 LEM, and its homogeneity across the CRP surface. Electron lifetime and resolution are also measured. The data collected are also used to attempt a parameterization of extraction and amplification efficiencies as function of the respective fields. The results are discussed Section 4.6.

4.1 Experimental setup

The experimental apparatus of the prototype is shown in Figure 4.1. It is composed of a 3×1×1 m^3 active volume (4.2 t active mass) TPC inserted inside a passively insulated cryostat with a total internal volume of ∼23 m^2. The TPC field cage and readout plane are both suspended under a 1.2 m thick insulator lead, called *top cap*. The use of a *top cap* to close the detector is necessary due to the small size of the cryostat. For larger prototypes, such as ProtoDUNE, and for DUNE it is possible to build the cryostat as a single structure with a uniform membrane coverage.

Five photo-multiplier tubes are located beneath the drift cage. Their role is to record the primary scintillation light (S_1) from the Argon excimer states formed by the charge particle crossing. Secondary scintillation (S_2) coming from the scintillation of the electrons extracted in vapor is also collected. The S_1 signal from the PMTs is used to provide the initial time of the event and provides the trigger timing to the data acquisition system. An alternative trigger signal is provided by the timed coincidence of the signal provided by the two cosmic ray taggers (CRT) located on the short side of the cryostat.

Table 4.1 List of resolution and operative range of the relevant devices used during the regular operation of the $3\times1\times1$ m^3 detector

Name	Resolution	Range
Temperature	< 0.5 K	50-350 K
Pressure	1% full range	near atmospheric
Coaxial level meters	1 mm	470-1200 mm
Plate level meters	100 µm	0-25 mm
HV LEM	10 mV	0-8 kV
HV Grid	1 volt	0-12 kV
HV Cathode	0.001% (volt)	0-300 kV

4.1.1 CRP arrangement

The charge readout plane is suspended from the top cap. Its design follows what is already described in Section 2.6.1 and has the role to provide a high-efficient extraction, amplification and collection of the drifted electrons. It is made up by twelve 50×50 LEM-anode *sandwiches*, electrically bridged together to instrument the full 1×3 m^2 surface of the TPC. The main difference between this arrangement and the design of Section 2.6.1 are the different lengths of the readout strips in the two directions.
The strips are connected to the front-end ASIC cards in two groups of 16 channels, for a total of 32 channels per strip. The cards are located in cold, just above the CRP, inside 4 signal feed-thought (SFT) chimneys. The signal is brought along the chimney by 1.5 m long twisted ribbon pairs cables from the electronics cards to external micro-TCA crates, which handles the digitalization and data acquisition operations. This arrangement follows what outlined for DUNE and described in Section 2.6.2

4.1.2 CRP instrumentation

Both the extraction and amplification of electrons are delicate processes and even small variations of temperature, pressure, high voltage, or the distance between the LEM and the liquid surface have a strong impact on the process. For the success of DUNE it is therefore necessary to have in place a precise and continuous monitoring system inside the cryostat. One of the main goal of the $3\times1\times1$ m^3 was also to test the temperature, pressure, and level probes and their interface with a common *slow control* database [102]. The characteristics, resolution and operative range of the main CRP instruments are reported in Table 4.1
Pressure gauges are located inside the cryostat main volume and in the insulation space. They consist of a set of transducers providing either relative and absolute reading in a range close to 1 bar. Temperature inside the cryostat is monitored with platinum resistance thermometers with a precision better than 0.5 K over the 50-350 K range. Those sensors are located in various positions on the CRP surface and at different heights, in order to provide a sampling of the temperature variations inside the cryogenic volume and monitor the temperature gradient in the amplification region.

The system to measure the level of liquid Argon consists of a set of coaxial and plate capacitors which change their capacitance as function of the depth in liquid to which they are submerged. Coaxial level meters provide in general a gross level reading and they are used during the phase of the detector filling. Parallel Plate capacitor are installed on the CRP and provide a real-time precise reading of the Argon level once it reaches to the top of the field cage.
The high voltage for the extraction grid and the LEMs is provided by a SY4527 CAEN mainframe and transported inside the detector volume with coaxial cables. The LEMs are biased independently on each electrode through the high voltage pins that traverse the dielectric surface. Specific connectors provides the appropriate bias to the grid.

4.1.3 Detector operation and collected data

The commissioning phase of the detector is articulated in three steps whose execution take about 1 month. First, the cryostat is purged, cooled down and filled with liquid Argon until the field cage is completely immersed. Camera, temperature, and coaxial level meters are used to monitor this phase. Once filling is completed the Argon recirculation system is switched and maintained during the detector operation. Finally the CRP level is adjusted parallel to the LAr level using the capacitor level meters.
The full operation period of the $3\times1\times1$ m^3 prototypes is comprised between the end of June and mid November 2017, including two shut-down periods of 2/3 weeks in August and October, where pulsing measurements and tests of the high-voltage settings were performed. A total of 5×10^5 events are collected, with simultaneous charge and light readout. Two different trigger configurations are used:

- **CRT triggers:** about 15% of the event are collected using the timed coincidence between the two scintillator CRT plate located on the short edges of the detector. The event rate in this trigger condition was approximately of 0.3 Hz.
- **PMT triggers:** the trigger condition is sent when all the PMTs sees a simultaneous signal greater than 125 mV within a 80 ns time window. With this conditions the event rate is of 3 Hz. A 200 ms veto window is implemented after each trigger signal in order to reduce losses of data packets on the network for events too close in time. The 85% of the $3\times1\times1$ m^3 events is collected with this configuration.

Each run is defined by the high-voltage settings and may have different time duration. A run is composed by subrun files with 335 events each. For the $3\times1\times1$ m^3 we define an event the signal formed on each strip in a 1667 time samples window (which is equivalent to the electron drift time with a drift field of 0.5 kV cm^{-1} after the trigger signal has been sent to digital acquisition system.
Both CRT and PMT triggered events were taken in various configuration of high-voltage settings for all the various regions: drift, extraction, LEM amplification, and induction fields. The regions are defined from the charge readout scheme of Figure 2.10. For *extraction* we define the region between the grid and the bottom electrode of the LEM, the *amplification* region is comprised between the two LEM electrodes where the electron avalanche occurs.

Table 4.2 The electric field settings explored during the operations of $3\times1\times1$ m^3 LArTPC. Nominal values are taken from Figure 2.10 and are the settings necessary to reach an absolute detector gain of 20 in a small dual-phase LArTPC with a 10x10 cm^2 readout area [99].

Region	Scanned range kV cm^{-1}	Best $3\times1\times1$ m^3 settings kV cm^{-1}	Nominal settings kV cm^{-1}
Induction	1 - 5	1.5	5
Amplification	23 - 31	28	33
Extraction (in liquid)	0.6 - 2.5	1.7	2
Drift	0.18 - 0.7	0.5	0.5

The *induction* region is comprised between the LEM top electrode and the anode. Each region is scanned individually by reducing the other field values in order to be compatible with the overall limitation due to the grid HV breakdown, which are commented in the following section. The general procedure for the scans required to increase step by step the electric field in a specific region up to a limit imposed by the high voltage breakdown. The configuration reached at each step is kept for half an hour during which charge and light data are collected. The fields from the other regions are modified during the scan in order to maintain the grid below its maximal operation condition below -5 kV. The ranges of values scanned for each region are reported in Table 4.2 and compared with the nominal design values of Figure 2.10.
In parallel to the scan of the different field regions, a specific tuning of the high-voltage settings was found out in order to identify the best configuration able to maximize the gain in stable condition. A more than 8 hours-long run, *Run 840*, was taken using the optimal settings of Table 4.2.

4.2 Detector operation stability

This section focuses on the stability of the $3\times1\times1$ m^3 detector during the operation period and the issues encountered. One of the goals of the $3\times1\times1$ m^3 detector is to ensure stable operation of the detector in cryogenic conditions. The fluctuations of temperature, pressure and liquid argon level are described, as well as the highest high-voltage settings reached. A detailed description of the noise characteristic and origins is also given, with the main goal to capture the most relevant noise patterns for the noise simulation described in Section 4.3. Finally, we shortly comment on the detector calibration.

4.2.1 LAr level, pressure and temperature stability, and electron lifetime

Temperature, pressure, and level of the liquid Argon have been proven stable during the entire operation periods. Fluctuations of temperature and pressure are checked routinely in order to monitor the heath load from the insulator. Figure 4.2 (top) shows the variation of pressure

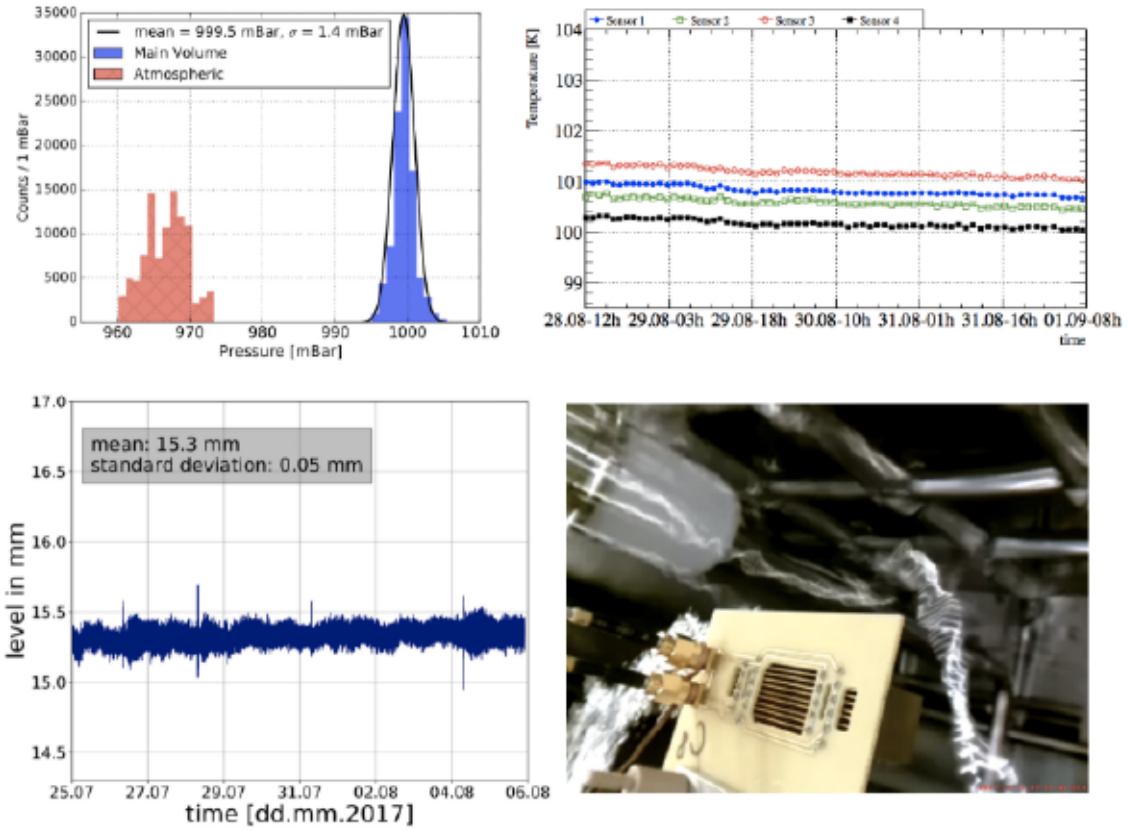


Fig. 4.2 Top: Pressure (left) and temperature (right) variations over one week of detector operation. The fluctuations of temperature are related to 4 different sensors located at different points on the CRP, 2.6 cm above the LEMs. Bottom: fluctuations of the liquid Argon level for a 12 days period (left) and visual inspection of the flatness of the level using a camera located inside the cryostat (right) [102].

(left) and temperature (right) over a one week period of detector operation. Pressure inside the cryostat is found stable around $P = 999.5 \pm 1.4$ mbar and temperature gauges located in different places of the CRP record temperature differences between them of ± 1 K. The level meters are automatically checked to maintain a constant liquid inside the TPC. Figure 4.2 (bottom) shows the excellent stability of the liquid Argon level, with only small fluctuation of 0.3% (left).
The electron lifetime has been measured and found around 4 ms, value larger than the minimal requirement of 3 ms . The methodology used is described in [102].

4.2.2 High-voltage stability and maximal field achievable

Meeting the nominal high-voltage settings of Figure 2.10 requires to provide simultaneously an extraction, amplification, and extraction fields of respectively 2 kV cm^{-1}, 33 kV cm^{-1} and 5 kV cm^{-1}. Drift field is set to be 0.5 kV cm^{-1}. This configuration requires to bias

the extraction grid at a nominal voltage of -6.8 kV. However, due to technical issues, the grid could only be operated with stable high voltage down to -5.0 kV. Above this value discharged occurred.
The results of investigations pointed to a combination of a faulty high voltage contact on the extraction grid and an untensioned group of wires. The latter hypothesis is supported by the presence of shortcuts observed between the grid and two of the twelve LEMs electrodes when the Argon liquid is increased. The visual inspection inside the cryostat after the end of the data-taking confirmed an issue on the tensioning system holding a group of 32 one meter long wires located under those specific LEMs. We also ensured that the extraction grid was fully immersed in liquid by testing the maximal voltage sustainable in air of -2 kV.
The grid issues forced the detector to be operated at different high-voltage settings and the maximally achievable electric field sustainable was explored independently for each region (drift, extraction, amplification, induction). The scanned high-voltages are reported in Table 4.2. The results for the scans of each region are commented below:

- **Extraction:** despite the fault on the grid system previously described, it has been possible to scan the extraction field up to 2.5 $\mathrm{kV\,cm^{-1}}$
- **Induction:** a few tests with low induction between 1 and 5 $\mathrm{kV\,cm^{-1}}$ were performed.
- **Amplification:** a maximal stable field of 31 $\mathrm{kV\,cm^{-1}}$ is achieved with -0.2 kV and -3.3 kV on the top and bottom electrode of the LEM respectively and grid biased at -5 kV. The four corner LEMs must however operate at 27 $\mathrm{kV\,cm^{-1}}$. Appropriate trip current thresholds on all the high-voltage channels are set ensure stable operation of the whole CRP and occasional discharge on single LEMs do not prevent the performances of the others. By biasing individually each LEM, with everything else set at 0 kV, a field of 32 $\mathrm{kV\,cm^{-1}}$ is sustained for about one hour. This is in agreement with previous studies on the breakdown voltage of 50×50 $\mathrm{cm^2}$ LEMs [87], however it is not yet entirely clear why it is not possible to achieve this configuration with all the LEMs at the same time.
- **Drift:** a stable configuration 0.5 $\mathrm{kV\,cm^{-1}}$ with cathode at -56.5 kV is sustained during the entire period of the detector operation. Dedicated test with lower and higher fields were performed. Operation with 0.7 $\mathrm{kV\,cm^{-1}}$ (-75 kV at the cathode) where possible for about half an hour.

Table 4.2 shows also the optimal high-voltage settings found maximizing the detector performances in terms of quality of the TPC charge image. It was possible to operate the detector with this configuration for about 8 hours.

4.2.3 Noise level

The noise origin, characteristics, and evolution were monitored during the $3\times1\times1$ $\mathrm{m^3}$ operation, with a dedicated set of measurements. Fluctuations on each channel are quantified using the Root Mean Square (RMS) of the signal in ADC counts. The ADC counts are

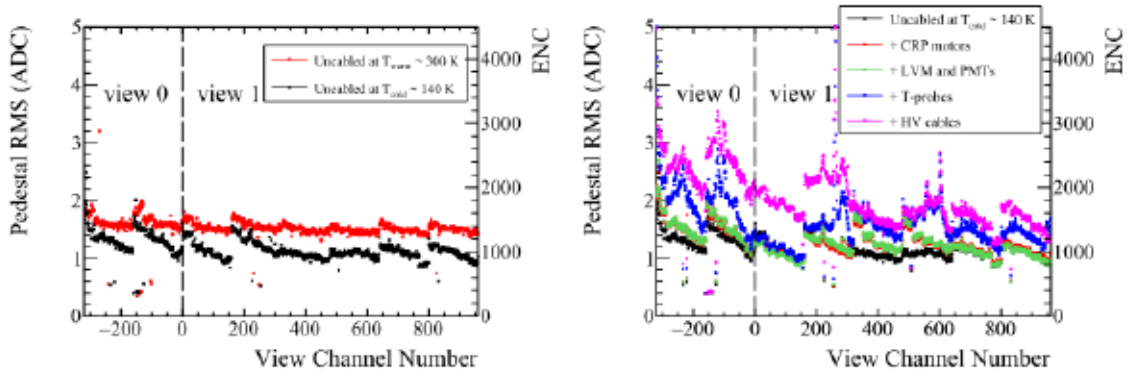
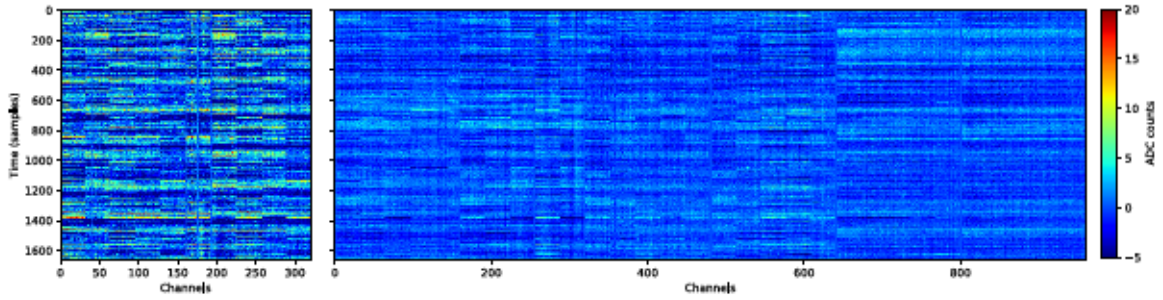


Fig. 4.3 Left: comparison of the noise RMS collected in both *cold* conditions (at 140 K) and at *room temperature* (near 300 K). Right: Noise RMS at cold measured for different slow control configurations [102].

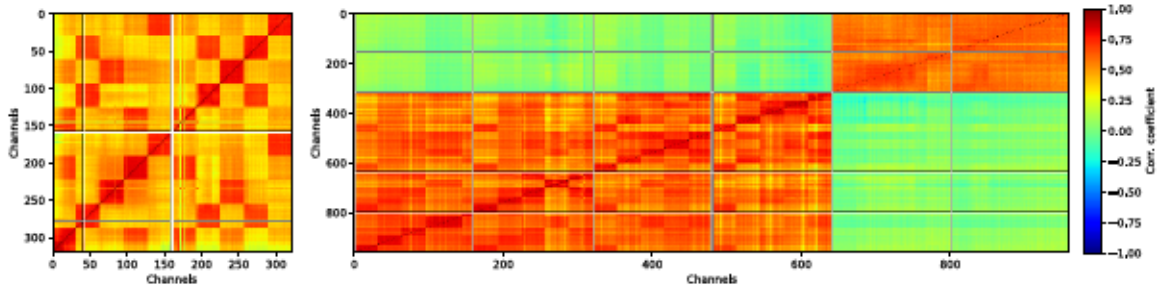
immediately related to a measure of the Equivalent Noise Charge (ENC) using the resolution of the digitizer for which holds: 1 ADC = 900 ENC. The results of the noise investigation are used as parameters to inject into the simulation of 3×1×1 m^3 detector.

Source of noise

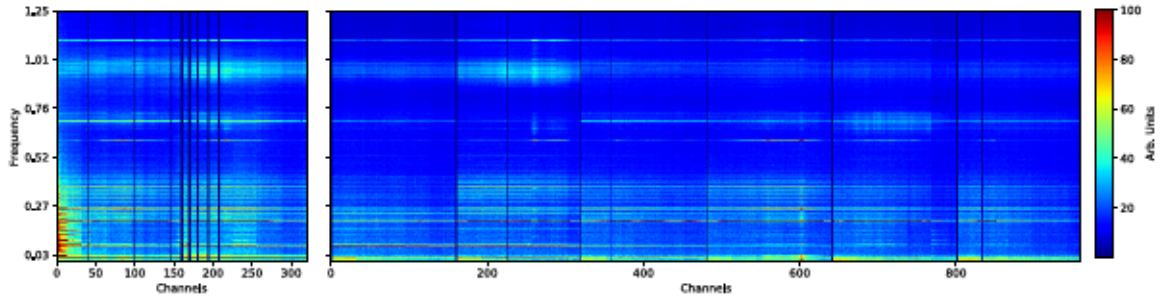
Noise measurements were taken both in *cold* conditions (at 140 K) and at *room temperature* (near 300 K). A comparison between the noise RMS on each channel at cold and room temperature is shown Figure 4.3 (left). The measure at room temperature are taken with all the slow control connector disconnected; it was not possible to repeat the same measure at cold since the Argon re-circulation pump and the high-voltage connectors for the cable could not be removed. Despite this, an improvement of about 25% is observable. It is interesting to notice that at room temperature the FE electronics outperform the predicted noise of Figure 2.13 (right): 1350 ENC are measured while 1530 ENC were expected for an input capacitance of 1.6 pF. It is possible that the intrinsic detector capacitance has been overestimated from design; this will be investigated further during the commissioning of the 6×6×6 m^3 prototype.
The impact of the different slow control systems on the noise level was also studied. The results are illustrated in Figure 4.3 (right): starting from cold conditions with all the cables disconnected except for the cathode high-voltage and the re-circulation pump, each slow control sub-system is progressively re-connected. Most of the noise comes from the pickup of the temperature probes and the high-voltage connectors. The magenta points represent the usual data-taking condition, with a noise level of about 2 ADC counts (1800 ENC). An interesting noticeable effect on both Figure 4.3 (right) and Figure 4.3 (left) is the repeated patter for groups of 160 neighboring channels, which reproduces the LAS arrangement. This effect might be due to noise reflections on the electrical bridges between the various 50x50 cm^2 anode units. Small differences between the noise RMS of the 1 meters-long strip (View 1) and the 3 meters-long strip (View 0) views is also observed in the conditions of Figure 4.3 (left).



(a) The channel map for the raw detector noise in time domain.



(b) Correlation matrix between channels in time domain.



(c) The channel map for the raw detector noise in frequency domain.

Fig. 4.4 The channel maps of the $3\times1\times1$ m^3 detector for both View 0 (left) and View 1 (right).

Noise patterns

The characterization of the most relevant noise patters is done using the noise reference run from July 2017 and assumed to be representative of the detector operation conditions. Figure 4.4a shows an example of the raw noise patterns after the baseline subtraction (not described here) across the channels of the detector for both views. Known dead wires have been excluded from the analysis and their signal set to 0. From Figure 4.4a, the noise appears correlated in time across common groups of channels. These coherent patterns are investigated with the correlation matrix in Figure 4.4b. Each entry of Figure 4.4b is the correlation coefficient r_{ij} calculated between the time signal on the i-th and the j-th channels. Strong correlation pattern are found every 32-groups of channels on View 0. This behavior is caused by the reflection of the waveform on the connectors to the FE card which can hold

maximum 32 pins arranged in two groups of 16. Also the electrical bridges between the 50×50 cm^2 anodes holds the same arrangement. On View 1 the correlations is between pairs of LAS, thus in groups of 380 channels. This depends on arrangement of the signal feed-through Chimneys, which connects the LAS in groups of two.
The relevant frequencies that make up the observed noise are studied by applying a Fourier transform on the recorded time signal. Figure 4.4c shows the average frequency spectrum for each channel of reference run for noise. The two main contribution to the spectrogram are identified at 0.1 and 0.2 MHz representing respectively the frequency of the power supplier and the AC/DC converters. The origin of the large burst around 0.9 MHz is still unclear. A similar pattern has been observed also by MicroBooNE [135], and its origin identified in power switching circuits of the high-voltage main board. Some observation of the frequency spectrum performed with the high-voltage boards on and off agrees with this hypothesis, however no conclusive explanation has been put forward so far. Since the frequencies composing the burst are narrowly localized around 0.9 MHz, a low-pass filter is used to mitigate this effects.

Stability of noise

Unfortunately pedestal measurements were not routinely taken during the detector operation, thus the evolution of the noise has to be inferred from data. However, it is possible to provide an estimation of the noise RMS using the TPC charge data by excluding the signal regions from the noise calculation. Figure 4.5 shows the evolution of noise as function of time during the detector operations. The dashed bars identified the two shutdown periods when operation of maintenance were performed to the main CAEN high-voltage boards. The noise RMS increased after each of the maintenance operation, however good stability is observed before, after and in between the two shutdowns. Also no strong night-day variations of the noise were observed The noise RMS level is significantly reduced by the coherent noise filter described in Section 3.5.1. Most of the noise introduced by the slow control grounding has a coherent origin, and the resulting incoherent contribution is stable and similar to the nominal design value of 1.1 ADC (1000 ENC) on both planes. The coherent noise filter seems also to mitigate the noise increase that occurred after the two shut-down periods.

4.2.4 Detector response with pulsing measurement

Specific pulsing runs were taken at different stages of the detector operation. Such measurements were performed in order to quantify realistically the detector response to known charge pulses and the cross-talk between neighboring strips. Due to the characteristics of the detector, it is possible to pulse only a minimum of groups of 32-channels at the same time. With pulsing measurements it has been possible to work out the calibration for the Front-end electronics and the level of cross-talk between wires. The latter has been found caused by capacitative couplings between parallel and perpendicular stripes quantified to be of few percents for every channel on both views. For the results presented in this thesis, the effect of the cross-talk is considered to be negligible. The setup of the pulsing system and the discussion of the results obtained is described in [136].

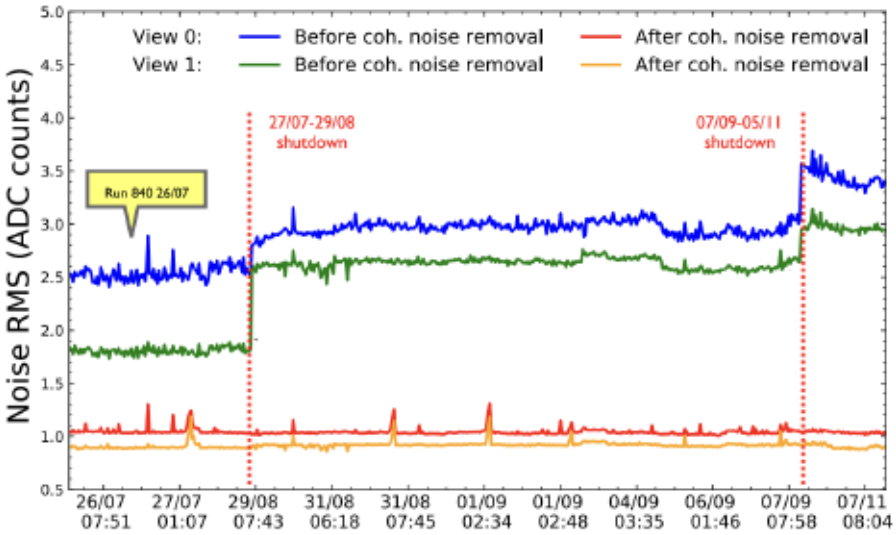


Fig. 4.5 The evolution of the noise RMS in ADC counts over the entire operations of the $3\times1\times1$ m^3 prototype for both the coherent and incoherent noise. The dotted red lines represent the shut-down periods, while the yellow box identify when the reference run for noise was taken.

4.3 Data-driven simulation

The LArSoft dual-phase simulation has been used to produce Monte Carlo samples in support to the study of the $3\times1\times1$ m^3 prototype. The steps of the simulation were outlined in Chapter 3. First, particles are generated from a database of pre-generated cosmic rays using CORSIKA to reproduce the correct correlation between angle and energy in the cosmic rays. Then the interactions between the particle and the Argon medium of the detector is performed using GEANT4, and finally the electrons produced by the ionization are projected on the readout planes and convoluted with the electronic response signal. The noise patterns investigated in Section 4.2.3 are configured in the noise simulation algorithm described in Chapter 3 and included as well in the simulation. A data-driven parameterization for the emission time of the slow component of the extraction and realistic probabilities of the electron transmission through the LEMs holes and collection on the anode have been introduced in the simulation, as detailed below.

4.3.1 Simulation of the electrons from the liquid to the gas phase

A careful study of the electron emission from liquid to gas is important in order to understand the detector signal recorded for the low-extraction runs of the $3\times1\times1$ m^3 detector. As previously described in Section 2.3.1, a fraction of electrons is emitted as *fast*, in a time lower than 0.1 µs, while a second *slow* fraction of electrons is emitted with time that can be larger than 0.1 ms. The relative amount of *fast* and *slow* electrons depends on the applied extraction field and determine a different time distribution of the charge collected of the

signal. For an extraction of about 1.5 kV cm^{-1}, the estimated emission time, τ_{slow}, is of ~ 8 µs [83]. This time is comparable with the shaping time of the electronics, and thus it impacts on the expected signal shape.
The simulation procedure for the *slow* electrons consists of summing the arrival time of the ionization electrons on the strips with a delay distributed according to an exponential with time constant τ_{slow}. The parameterization of τ_{slow} as function of the electric field from [83] and illustrated in Figure 2.5 (left) has been obtained for a configuration fairly different from the 3×1×1 m^3 detector and for electric field lower than 1 kV cm^{-1}, when the effects of the field become dominant over the thermal effects in the emission. For this reason, we propose a new data-driven parameterization of τ_{slow} for the 3×1×1 m^3 detector. The value of τ_{slow} is modified in the simulation until the average width of the hits matches with the value found in data. The comparison of the width is done by selecting tracks which have the same directions of the tracks generated with the Monte Carlo. This procedure is repeated for different values of the extraction fields.

4.3.2 Simulation of transmission and collection in gas

The peculiar operation conditions of the 3×1×1 m^3 detector requires to produce a dedicated simulation of the paths of the electrons along the field lines which are produced by the combination of the three electric fields. An example of this is shown in Figure 4.6 (top). The simulation has been done using the GARFIELD++ and ANSYS softwares [137, 138]. A detailed description of the simulation is beyond the scope of this thesis, however this is crucial to understand the interplay between the different electric fields and provide a prediction of the electrons losses on the bottom and top electrodes of the LEMs. The electron transmission efficiency from the liquid into one hole of the LEM is shown in Figure 4.6 (bottom left) as function of the extraction and the amplification fields; Figure 4.6 (bottom right) shows instead the collection probability at the anode as function of the amplification and the induction fields. The transmission and collection efficiency from the maps of Figure 4.6 (bottom) are used in the simulation to account for the average electron losses at different configurations of the high-voltage settings. A dedicated parameterization of the LEM amplification in the 3×1×1 m^3 is not available at the time this thesis is being written and, as consequence, its value needs to be inferred from data.

4.4 Data reconstruction and selection

The LArSoft reconstruction introduced in Chapter 3 has been used for the first time on real data of 3×1×1 m^3 detector for the reconstruction of the tri-dimensional particle tracks. The reconstructed trajectories are then filtered using some selection criteria to identify good-quality minimum ionizing particles (muons or pions), which are used for the data analysis. Two *mip* candidates are shown in figure Figure 4.7 (top).

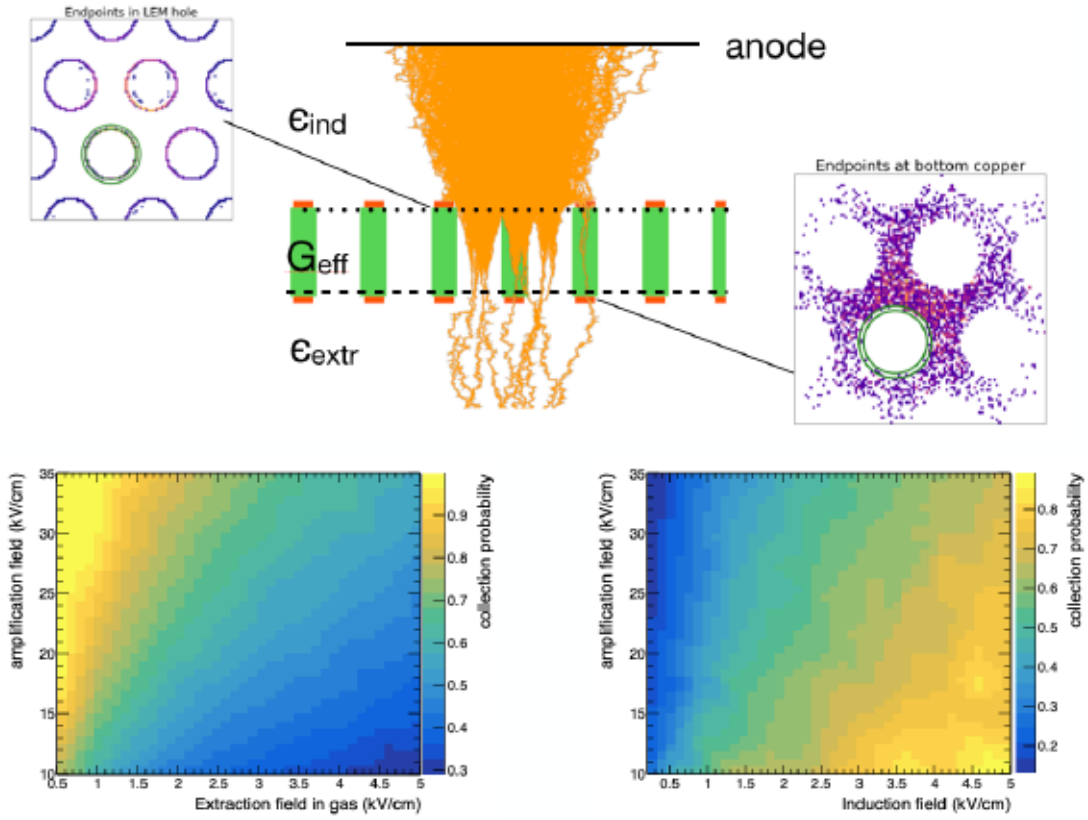


Fig. 4.6 Top: Simulated trajectories for electrons in gaseous argon. The scatter plots show the areas where part of the charge is collected (for this specific simulation settings) and thus does not contribute to the signal. Bottom: The electron transmission probability map from the liquid surface to the LEM holes (left) and the collection probability from the hole to the anode (right). [112]

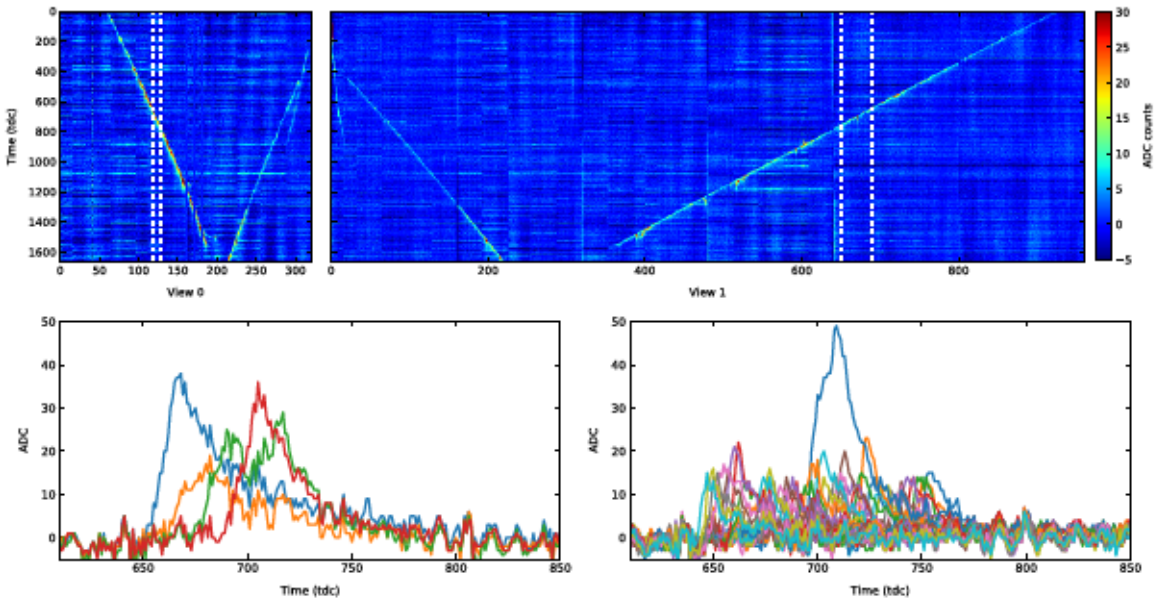


Fig. 4.7 Top: Raw event from *Run 840* showing two cosmic muons candidate. Bottom: Superposition of the waveforms found on the strips between the dotted white lines of the event display above for both View 0 (left) and View 1 (right). Each color denotes the signal collected on a different strip before noise removal.

4.4.1 Event reconstruction

Data reconstruction is performed offline within the LArSoft framework and uses the reconstruction chain described in Chapter 3. First the raw data collected are processed to remove the baseline, identify ROIs, and remove the coherent noise. The hit-finder algorithm is applied to the waveforms in order to identify the location of the peaks and their associated charge. Two dimensional clustering and projection matching algorithms are applied to form the tridimensional trajectory of the interacting particles.

4.4.2 MIPs selection

A trigger signal is sent when the PMTs collect in coincidence enough light above the threshold. This condition produces events which are agnostics to the type of particles that contains. For example *mip* candidates can be found in the same events with electrons, showering pion candidates, or stopping protons. For this reason it is important to have a good selection capable to discriminate *mips* from the shower candidates. The selection is done using two variables: the Charge Box Ratio (CBR) and the reconstructed track length.
An example of the CBR method is shown in Figure 4.8 (left): two boxes are built around a section *mip* candidate trajectory, one small in orange and larger one in green. CBR is defined as:

$$\mathrm{CBR} = \frac{Q_{\mathrm{big}} - Q_{\mathrm{small}}}{Q_{\mathrm{small}}}.$$

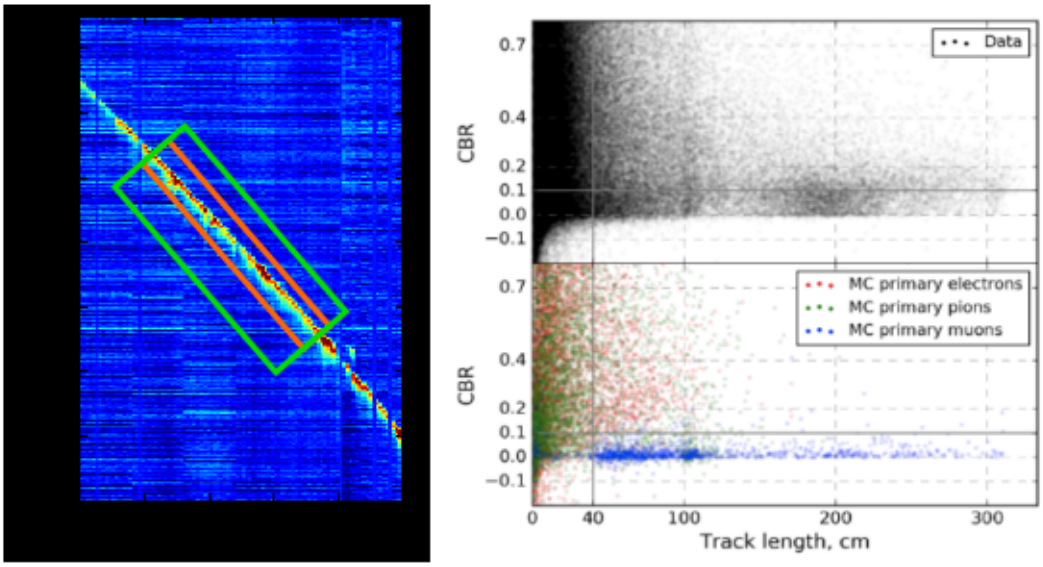


Fig. 4.8 Left: example of the CBR method applied to one *mip* candidate. The relative charge between the green and the red box is computed along the track trajectory. Right: distribution of the track length and cbr ratio for the reconstructed track in data (top) and in Monte Carlo (bottom) for pions, electrons, and muons. [112]

where Q_{small} and Q_{big} are the total charge found respectively in the small and in the big box. The philosophy behind this observable is mostly related to the different topology of the energy deposit which distinguish tracks from showers: trajectories reconstructed in a shower branch are often surrounded by charge deposits from the other showers branches, and thus they would result in a relatively larger CBR compared to a *mip* candidate trajectory, which is usually isolated from other energy deposits. The second cut used is the reconstructed track length. This is one other effective criteria to separate *mips* candidates from electron or photon induced cascades which is based on the *bremsstrahlung* radiation length of showering particles which is ~ 15 cm in LAr.

The tuning of the selection thresholds for the observables previously described has been done with both an eye scan of the selected data and using a Monte Carlo produced with one single type of interacting particle. The distribution of CBR and reconstructed length for every selected track is shown in Figure 4.8 (right). A cut at CBR< 0.1 and track length > 50 cm is found effective to select a very pure ($> 95\%$) sample of muons. The residual contamination comes mainly from pions, which at low momenta are also minimal ionizing particles. Selection efficiency with those cuts on a pure Monte Carlo sample of muon is about 80%. The following cuts are not sufficient to isolate stopping particles. Their contribution is however expected to be negligible.

Further quality cuts are imposed on the reconstructed hits in order to ensure a full discrimination of residual noise and a good approximation of the charge deposited using the fit. In order to be used, a hit must be associated to a tridimensional trajectory and have χ^2 larger than 0 and smaller than 5 and local direction of at least 2° off the vertical and horizontal direction. Due to the small size of the detector (and the relatively high LAr purity, as discussed later),

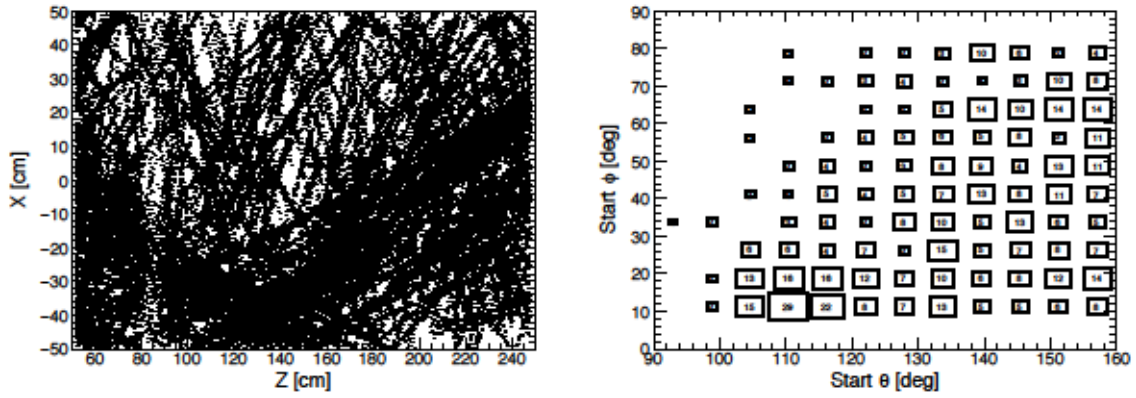


Fig. 4.9 Distribution of the coordinates of the hits in the selected tracks of run *Run 840* for the drift direction x and the three meters-long projection z (left). Distribution of the direction of the reconstructed tracks of run *Run 840* with respect to the azimuthal (θ) and polar (*phi*) angles, as defined in Figure 4.1 (right). The numerical value associated to each box is the number of reconstructed tracks for that particular bin.

lifetime corrections are not applied to the estimated charge. Also the first and the last 3 hits on both views which are associated to a track are excluded since the trajectory is often not well reconstructed at the border of the active volume. As last, according to the Run, hits located under the LAS modules which were not biased with the same tensions as the others are in general excluded from the analysis.

4.4.3 Distribution of the selected track and Monte Carlo sample

Figure 4.9 (left) illustrates the distribution of the selected *mip* candidates of *Run 840*. The coordinates of the hits in the selected tracks are shown in Figure 4.9 (left) for the drift direction x and the three meters-long projection z, while Figure 4.9 (right) displays how the direction of the reconstructed tracks are distributed with respect to the azimuthal (θ) and polar (ϕ) angles, as defined in Figure 4.1. Most of the reconstructed and selected cosmic rays *mips* are in general quite horizontal and parallel to the three meters long direction. This is caused mainly the light trigger condition used which tends to privilege long trajectories with a large photon yield across the entire TPC volume. The distribution is also slightly asymmetric towards *mips* coming from the right side of the detector. This is the consequence of the calibration of the PMTs, which is not uniform. This asymmetry has a negligible impact on the analysis since tracks coming from the opposite side of the detector have similar projections on both views.
The distribution of the direction of the reconstructed tracks is the result of the convolution between the primary cosmic rays distribution, detector acceptance, trigger conditions, and reconstruction efficiency. Although it is not straightforward to produce a full simulation of the light trigger, some studies are currently ongoing [139]. For the purpose of this thesis it is sufficient to produce a Monte Carlo sample of *mips* distributed as observed in data. The

sample is generated assuming that every *mip* candidate selected is a muon with energies drawn from the cosmic ray database produced with CORSIKA for the relevant angles. The minimal energy allowed is set to be above 2 GeV. The sample is prepared using the simulation of Section 4.3 at it is designed to reproduce the characteristics of *Run 840*.

4.5 Results from the $3\times1\times1$ m^3 prototype

A complete characterization of the detector performances requires a study of the effective gain achieved, its fluctuations in space and time, and its dependence on the different high-voltage settings. This can be done in the $3\times1\times1$ m^3 by comparing the energy deposit of *mips* with the expected distribution. For the data analysis we consider two subset of the data sample:

- *Run 840*, since taken at the best $3\times1\times1$ m^3 settings of Table 4.2, is used to measure and study the fluctuations of the effective gain.
- The runs used for scanning the high-voltage settings of Table 4.2 have been used to compare the measured valued for the extraction and amplification with the expected trends.

The detector resolution is discussed separately considering both the field scans and *Run 840*. The results obtained for the $3\times1\times1$ m^3 are compared with the performances of a previous dual-phase LArTPC, with an active volume of three liters and a single 10×10 cm^2 LAS for readout. We refer to this setup as *3L*, and its results constitutes most of the literature available for the charge readout capabilities of dual-phase LArTPC [87, 98–100, 140]. The 3L was exposed to cosmic rays, as the $3\times1\times1$ m^3, and stable at the nominal high voltage settings of Figure 2.10. Although the design of the 50×50 cm^2 LAS of the $3\times1\times1$ m^3 detector is derived from the 10×10 cm^2 LAS of the 3L, some cautions are necessary when comparing the two setups since the cryogenic conditions, high-voltage settings, and noise are different between the two cases. Furthermore, the 10×10 cm^2 LAS of the 3L is a single independent unit, while the 50×50 cm^2 anodes of the $3\times1\times1$ m^3 detector form an uniform readout. Also the noise, the signal processing and the event reconstruction are different between the two setups.

4.5.1 dQ/ds benchmark

The 3D reconstruction allows to retrieve the length of the track ds on each strip of view 0 and 1, along with the charge dQ reconstructed for each hit. The distribution of the ratio dQ/ds is proportional to the energy locally deposited by the track in liquid Argon per unit of length, which, for a minimum ionizing particles in a thin detector, is the Landau-Valivov distribution introduced in Section 2.2. Figure 4.10 shows the reconstructed dQ/ds distribution for the minimum ionizing particles recorder during *Run 840* on View 0 (left) and View 1 (right). The distributions on both views are relevant since dual-phase LArTPCs feature two collection views, and therefore ignoring one of the two projections would result in

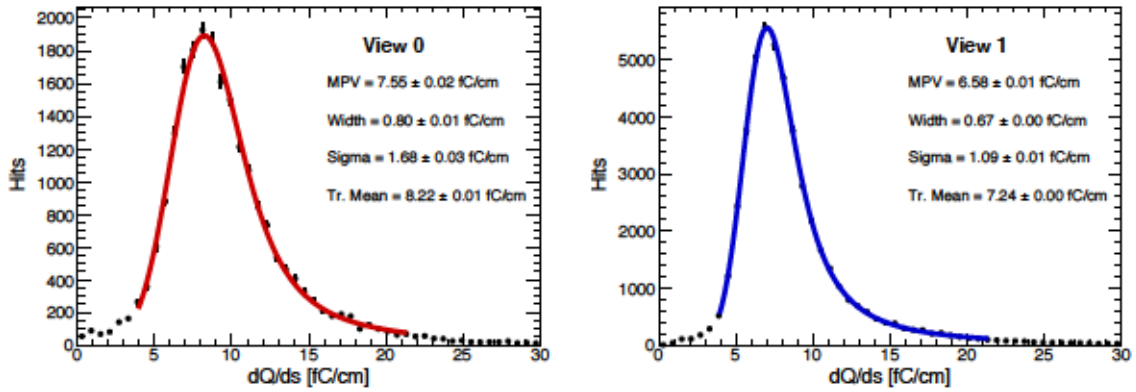


Fig. 4.10 Reconstructed dQ/ds distribution of *Run 840* for both View 0 (left and View 1 (right). The convolution of a Landau with a Gaussian fits the data: the parameters used for the fit are listed on the figure. The reported uncertainty is entirely statistical.

a loss of information. The data are described with the convolution of the Landau distribution with a Gaussian, representing the fluctuations in the charge collected introduced by detector effects (i.e fluctuations in the size of the electron avalanche produced inside the LEM).
As previously discussed in Section 2.2, the parameters of the LVD are sensitive to the events located on the tail of the distribution. δ-rays gives in general a significant contribution to the tail, however they are not easy to reconstruct and attach to their production point using the reconstruction proposed. This leads to a under-estimation of the mean of the distribution. For this reason the most probable value is in general used, as it is less dependent on the tails compared to the mean [1].

4.5.2 Charge sharing between views

One of the requirements of the dual-phase design for DUNE is ensuring a good-quality charge sharing between the two views. This is particularly important to ensure a good measure of the gain and to improve the overall resolution of the detector. The dual-phase anode has been conceived to provide perfect charge sharing independently of the particle orientation. This has been confirmed with the 3L setup, which has measured variations between the two views at the level of 1% [100]
To study the charge sharing, we observe the variations of the MPV of the Landau distribution for particles selected at different values of the polar angle ϕ in *Run 840*. For each value of ϕ considered, both the asymmetry coefficient and the resolution are calculated. The view

[1] A 30% *truncated* mean is often used as well. Truncated mean is quite practical because it does not require to be extrapolated from a fit. However, it is less straightforward to compare with its theoretical expectation than it is for the MPV.

asymmetry coefficient $\mathscr{A}$ is:

$$\mathscr{A} = \frac{\langle dQ_1/ds \rangle - \langle dQ_0/ds \rangle}{\langle dQ_1/ds \rangle + \langle dQ_0/ds \rangle}, \tag{4.1}$$

where $\langle dQ_i/ds \rangle$ is the MPV of the distribution on the i-th view. The variation of both the MPV values and $\mathscr{A}$ are shown as a function of ϕ in Figure 4.11 (left). The average asymmetry is around -7% with a sharp increase to -18% for $\phi \approx 15°$. An equal sharing is achieved for $\phi = 45°$. The large asymmetry found at extreme angles may be understood if systematics effects for charge reconstruction are taken into account. The particular trigger conditions and detector acceptance causes tracks at small values of ϕ to be almost perpendicular to View 1 and parallel to the three meters long strips on View 0. This means that the charge deposited by the particle is seen on View 0 by only few wires. This produces large waveforms, which are rather complicated to fit and whose integral in general overestimates the charge deposited. The reverse happens on View 1 where all the charge deposited is shared among many strips, and the resulting signal-to-noise ratio is thus smaller. This effect, in combination with large noise, has the consequence that the charge of some hits is systematically underestimated on View 1, because it is more difficult to isolate the hit from the noise. An example of the typical waveforms of a *mip* candidate traveling parallel to View 1 is illustrated in Figure 4.7 (bottom). This assumption is supported by observing the relative contribution of View 0 and View 1 to $\mathscr{A}$. View 0 is dominant over View 1 for angles below 45°, while the situation is reversed for angles larger than 45°. The different length of the strips and the larger coherent noise observed on View 0 further enhance this effect.
This behavior has qualitatively been observed also with the Monte Carlo simulation, however, due to some limits in the simulations that are illustrated later, it was not possible to reproduce it also quantitatively.
With the 3×1×1 m^3, it is not possible to disentangle entirely the biases introduced by reconstruction from possible evidences of charge asymmetries. The small asymmetry measured at 45° and the reasonable explanation put forward for the observed behavior are both encouraging evidences of a possible good charge sharing. More studies in this sense are expected to be carried on in ProtoDUNE-DP.

4.5.3 Effective gain

The *effective* gain of the detector G_{eff} is calculated by comparing the measured MPV from the reconstructed dQ/ds on View 0, $\langle dQ_0/ds \rangle$, and View 1, $\langle dQ_1/ds \rangle$, with the expected value of predicted for the LVD:

$$G_{eff} = \frac{\langle dQ_0/ds \rangle + \langle dQ_1/ds \rangle}{\langle dQ/ds \rangle_{exp}}. \tag{4.2}$$

$\langle dQ/ds \rangle_{exp}$ has been estimated from the Monte Carlo sample previously described after the GEANT4 propagation and it is found to be 7.75 fC.

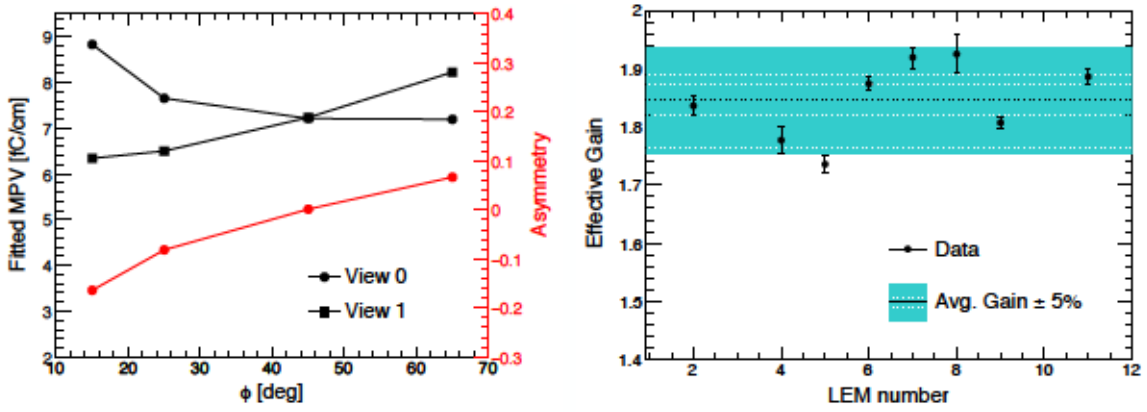


Fig. 4.11 Left: Fitted MPVs for both View 0 and View 1 and asymmetry coefficient (in red) as function of the polar angle ϕ. Right: Effective gain calculated for every LAS. The solid line and the blue bands indicate respectively the average effective gain and $\pm 5\%$ design tolerance (see text).

Because of the dependence of the fitted MPV on the polar angle ϕ observed in the previous section, we calculate the effective gain of the detector only considering tracks in a narrow angle of $45° \pm 5°$, where biases introduced by the reconstruction should be minimal and affecting equally both views. In these conditions, a $G_{eff} \approx 1.84$ is found. The low value of G_{eff} is related to the non optimal high-voltage settings of the run considered and thus, it is not representative of the actual amplification factor provided by the LEM, which is instead discussed in Section 4.5.6. Moreover, G_{eff} is measured before the *charging up* of the dielectric [99]. This phenomenon describes the degradation of the amplification performances of the LEM related to the accumulation of positive ions around the dielectric rim of the holes over several days and has the effect to produce an attenuation of the gain of about a factor three. Although *Run 840* is the longest stable run acquired during the 3×1×1 m^3 operation and lasted for more than 8 hours, this time is not sufficient to study the *charging up* effect with sufficient detail on the three 1 m^2 CRPs.

4.5.4 Gain uniformity

A big challenge for the design of the dual-phase CRP is to ensure a good uniformity of the gain over the entire surface of the CRP. Since amplification and collections are performed locally by each of the 50×50 cm^2 LAS composing the CRP, small fluctuations in the total gain achieved are expected from LAS to LAS due to the disaligment of the different elements or to small manufacturing errors (see Table 2.2). Gain fluctuations on the single LAS are expected to be lower than 5% by design.
Unlike the 3L, the 3×1×1 m^3 detector offers the possibility to study the gain uniformity over a 3 m^2 CRP. Figure 4.11 (right) shows the gain fluctuations for *Run 840* calculated from the fitted MPVs of the dQ/ds distribution recorded by each LAS. Data shows that the local gain fluctuations around the mean value are constrained within the 5% tolerance set by design. Regretfully, due to relatively short data taking period of *Run 840*, it has not been

possible to study the stability of the gain of each LAS as function of time. As discussed in the following section, such a measurement is important to perform an *in situ* calibration for each LAS and therefore constrain better the systematic effects and improve the resolution.

4.5.5 Impact of systematic uncertainties

In this section we try to sort the impact of systematic effects on measurement of the effective gain from the MPV measured using the fit of Figure 4.10. Four main contribution are identified and studied:

- **Fit systematics.** We have studied the impact of the chosen range on the fitted values of the MPV. Due to the characteristics of the Landau-Vavilov distribution, it is possible that a fit including the contribution of the tail may yield an overestimation of the measured MPV. A toy Monte Carlo has been prepared to study the variations of the fitted MPV for different ranges compared to the expected value injected. No significative contribution is found compared to the expected statistical fluctuations.

- **Calibration uncertainties.** Equation 2.6 is used to convert the deposited energy to ionization charge. The uncertainties on the ionization potential and the recombination have been both estimated at the order of 2% each [71, 74]. Both errors are treated as fully correlated and they are found contributing to the gain at the level of 2% on the expected most probable value of the LVD.

- **Instrumental uncertainties** The extraction, amplification, and collection processes are sensitive to the applied electric fields and gas density. The extraction field depends on the knowledge of the liquid Argon level with respect of the grid, the amplification field depends on the knowledge of the LEM thickness, pressure and temperature, while the collection depends on the measure of the distance between the LEM and the anode. For extraction and induction, the uncertainties on the respective fields are propagated to the respective efficiencies using the field maps of Figure 4.6 (bottom). For the amplification field, the errors on pressure, temperature, and lem thickness are instead propagated using the parameterization of the LEM amplification as function of the LEM field before *charging up* found for the 10×10 cm^2 anode [99]. The uncertainties on the extraction, amplification, and induction efficiencies are treated as fully correlated and they impact the effective gain at the level of 9%. The larger contribution is brought by amplification with a 7.4%. The instrumental uncertainty here proposed differs from the 5% systematic tolerance from the LAS manufacturing discussed in Section 4.5.4 because the variations of pressure and temperature are also considered.

- **Angle fluctuations** As illustrated in Section 4.5.2, there are systematic variations on the measured MPVs depending on the polar angle ϕ. This impacts on the measure for the effective gain with an additional 4%. The effect depends on the signal-to-noise ratio of the run, and it is not possible at the time of writing to provide a correction based on Monte Carlo.

The three main systematic contributions previously described are assumed uncorrelated. Their combined impact on the measure of G_{eff} is at the level of $\pm10\%$.
This large systematic contribution is related for the greatest part to the impossibility to operate the detector in stable conditions for very long time and it does not represent the level of uncertainties expected at DUNE. Both the instrumental and calibration uncertainties can be constrained with a local calibration of the gain performed for each LAS. The error derived from the fluctuations of the MPV as function of the angle depends on the poor amplification and has great impact because of the peculiar trigger conditions. A specific correction for this effect may be studied using a simulation, however its impact should be greatly reduced with the detector operating in its nominal configuration. Within limits of the detector, it is not possible to discuss further the detector calibration and the related treatment of uncertainties.

4.5.6 Field scans

With the field scans it is possible to study the variation of the detector response as function of different high-voltage settings. This is important in order to study the interplay between the different factors that compose the gain. The *effective* gain provided by each of the readout steps illustrated in Figure 2.10 can be factorized as follows:

$$G_{eff} = (f_0 + f_1) \cdot \varepsilon_{ext} \cdot G_{LEM} \cdot \varepsilon_{col}, \tag{4.3}$$

with f_0 and f_1 being respectively the charge sharing factors of View 0 and 1, ε_{ext} the extraction efficiency, G_{LEM} is the amplification provided by the LEMs, and ε_{col} is the collection efficiency on the anode. In general we have $(f_0 + f_1) = 1$ with $f_0 \approx f_1$, ε_{ext} and ε_{col} are ≤ 1, and $G_{LEM} \geq 1$.
We consider the scan of the extraction and the amplification fields. The former is used to qualitatively confirm the transmission probability computed in Figure 4.6 (bottom left), while the latter is used to quantify G_{LEM} as function of the amplification field. This is particularly useful to have an estimation of the amplification provided by the 50×50 cm^2 LEMs. For these studies, we use *mips* selected from the extraction and the amplification scans of Table 4.2. Due to the conditions of the field scans, it is not possible to study each field independently while keeping the others constant. For this reason *ad hoc* corrections are applied to the results of the scans. Moreover, the non-optimal detector settings and the reconstruction that has been specifically studied for *Run 840* limits the number of *mip* candidates found. Runs with less than 500 valid hits on each view are excluded from the analysis.

Extraction field scan

The extraction efficiency, ε_{ext}, is the fraction of electrons successfully extracted from liquid into gaseous Argon and subsequently collected inside the LEM holes. ε_{ext} can be further factorize in two contributions:

$$\varepsilon_{ext} = \varepsilon_{ems} \cdot \varepsilon_{trans}, \tag{4.4}$$

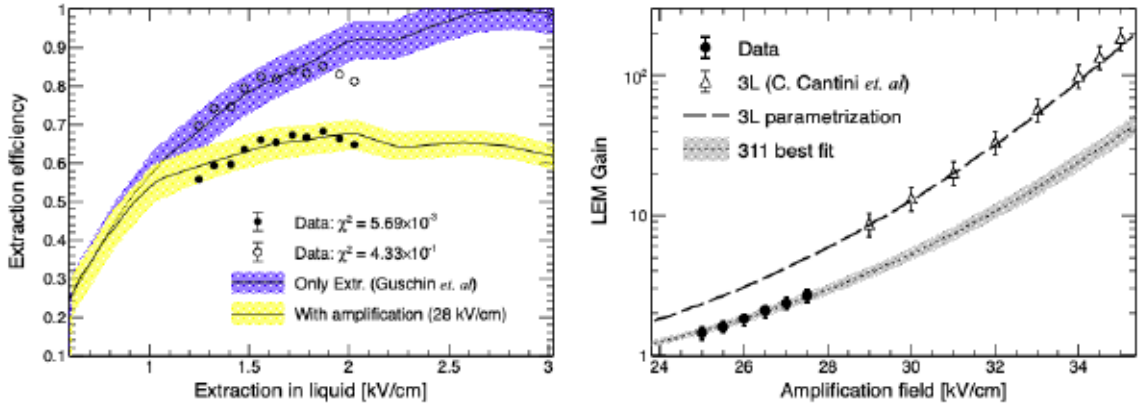


Fig. 4.12 **Left**: results of the extraction field scan. The extraction probability (blue) is convoluted with the probability of Figure 4.6 (bottom left) to obtain the transmission probability for electrons with an amplification of 28 kV cm^{-1} (yellow). Two different normalization are applied to the data (see text) to match blue (open circles) and yellow (full circles) curves. A χ^2 is calculated to show the relative agreement of data with the two expected probabilities. **Right**: results of the scan of the amplification field. The data collected are normalized to the high-voltage settings of the 3L setup [99] and fitted with an equivalent parameterization (see text). The dotted line shows the best fit of the data, while the gray area includes the contribution of the uncertainties on temperature and pressure.

where ε_{ems} is the efficiency of the field-enhanced thermionic emission of Section 2.3.1, and ε_{trans} is the transmission probability of the electrons as function of the extraction and amplification fields from Figure 4.6 (bottom left). Figure 4.12 shows both ε_{ems} (in blue) and ε_{ext} (in yellow), calculated for a fixed amplification field of 28 kV cm^{-1}. The difference between the two curves for increasing values of the extraction field proves the impact of both the extraction and amplification fields in driving correctly the emitted electrons into the LEM holes.

The G_{eff} is measured for each run of the extraction field, taken at fixed amplification of 28 kV cm^{-1} and induction at 1.0 kV cm^{-1}. The value of ε_{ext} is obtained by inverting Equation 4.3, where ε_{col} is taken from Figure 4.6 (bottom right). Since G_{LEM} remains unknown for the 3×1×1 m^3 detector, it is necessary to choose an arbitrary normalization for the measured ε_{ext}. Figure 4.12 shows the data points of the extraction scan normalized by two different factor chosen to match either the blue or the yellow prediction curves. A χ^2 is used to quantify the agreement of the data with both curves. The data points renormalized to the yellow curve show a trend that is in better agreement with the factorization of the extraction efficiency proposed in Equation 4.4 and calculated from Figure 4.6 (bottom left).

Amplification field

The scan of the amplification field allows to attempt a parameterization for the amplification provided by the LEMs. This is important in order to understand the amplification capabilities

Table 4.3 The parameterization of Equation 4.5 chosen for both the 3L and the 3×1×1 m^3 amplification fit. Values relative to the 3L are taken from [99] and references therein, with the exceptions of $A\rho$ and $B\rho$ which are taken from [87]. For the 3×1×1 m^3 the value of induction field $\mathscr{E}_{\text{ind}}$, extraction field $\mathscr{E}_{\text{ext}}$, temperature and pressure are relative to the running conditions during the scans, l and $\mathscr{T}$ are taken from [99]. k is found thanks to a simulation for the 3L and measured with a fit in the 3×1×1 m^3 (see text).

LEM	$\mathscr{E}_{\text{ind}}$ kV cm^{-1}	$\mathscr{E}_{\text{ext}}$ kV cm^{-1}	P bar	T K	$A\rho$ cm^{-1}	$B\rho$ (kV cm^{-1})$^{-1}$	l cm	k	$\mathscr{T}$
3L	5	2	980	87	3576	145.3	0.98	0.90	0.59
3×1×1 m^3	1	1.5	999	93	3749	152.4	0.98	0.80	0.59

of the 50 × 50 cm^2 LEMs in a realistic dual-phase LArTPC setup. The results are compared with the expectations from the 10 × 10 cm^2 LEMs of the 3L detector [99].
G_{LEM} is found from the measured G_{eff} in data by inverting Equation 4.3. Each measured point is further corrected for an appropriate factors ε_{ext} and ε_{col} to be consistent with the settings of the 3L setup. The corrected data points are shown in Figure 2.6 for values of amplification field from 25 to 27 kV cm^{-1}. The gain measured with the 10 × 10 LEM is also shown in Figure 4.12 along with the best fit of the Townsend avalanche as reported in the paper [99] :

$$G_{\text{LEM}} = \mathscr{T} \cdot e^{\alpha(k\mathscr{E}_{\text{amp}})l}, \tag{4.5}$$

where l is the LEM thickness, α is the first Townsend coefficient from Equation2.12 given as function of the reduced amplification field $k\mathscr{E}_{\text{amp}}$. k is a constant introduced by the authors of [99], which accounts for the effective amplification provided by the LEM given the distribution of the field lines. This is necessary because the field lines are not parallel and homogeneous everywhere inside the LEM holes. The value of k was estimated for the 10 × 10 cm^2 using COMSOL (see [99] and references therein). $\mathscr{T}$ in Equation 4.5 is the electronic transparency of the 3L setup. $\mathscr{T}$ depends on the position and size of the different components of the LAS, as well as on the high-voltage settings. Since the data points collected with the 3×1×1 m^3 detector are normalized to the conditions of the 3L setup, it is reasonable to believe that the electron transparency is similar. The run conditions and the fit parameters used in both the 3L setup and the 3×1×1 m^3 are reported in Table 4.3.
The 3×1×1 m^3 data are fitted according to the same parameterization of [99], scaled to the temperature and pressure conditions of the 3×1×1 m^3 detector, and leaving k as free parameter. A good agreement is found with a value of $k = 8.085 \times 10^{-1} \pm 4.9 \times 10^{-3}$, which is in tension with the value found with the simulation in the 3L. An explanation from this may involve the field lines of the collection and amplification fields which, in the anomalous condition of the data taking are different inside the holes of the LEMs with respect to what predicted for the 3L. Alternatively, it is possible that the 50 × 50 cm^2 LEMs underperforms compared to the analogous 10 × 10 cm^2, because of the larger surface. Unfortunately, the data collected in the 3×1×1 m^3 prototypes does not allow to further investigate this discrepancy. A scan of the amplification field performed in exactly the same conditions as the 3L setup

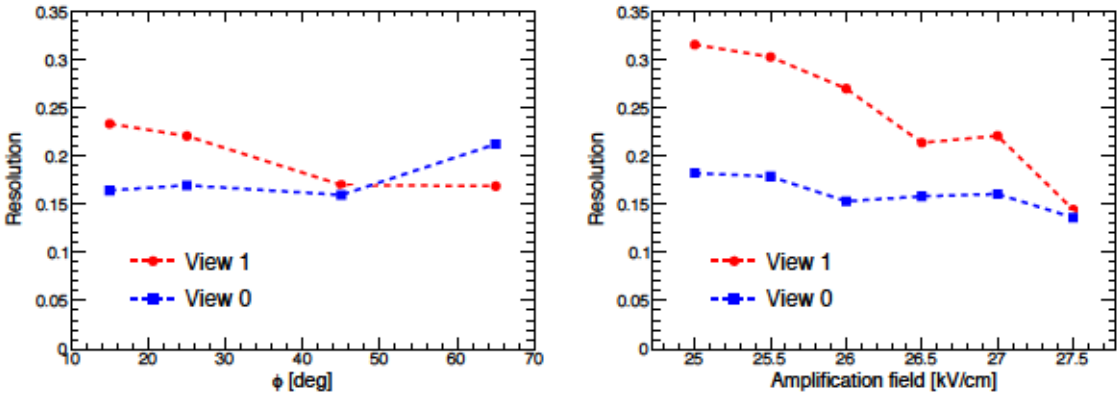


Fig. 4.13 Left: Resolution of *Run 840* as function of the polar angle ϕ. Right: Resolution as function of the amplification field for the runs of the amplification scan of Table 4.2.

would be the only way to confirm if the effective amplification performances for both LEMs are compatible despite their different surfaces.

4.5.7 Resolution of the detector

To correctly characterize the 3×1×1 m^3 detector response it is also important to study the detector resolution to minimum ionizing particles. This quantity is important for the discrimination power of the detector to different types of particles (for example muons or protons). We define the *detector resolution* with the ratio between the standard deviation of the Gaussian smear and the MPV of the LVD:

$$R_i = \frac{\sigma_i}{\langle dQ_i/ds \rangle}, \tag{4.6}$$

where, for the i-th view, σ_i is the observed standard deviation of the Gaussian convoluted with the Landau used for the fit of the dQ/ds distribution as previously described.
Based on the performances of the 3L, the expected resolution for *mips* in the DUNE dual-phase detector is set around 10%. The measured resolutions in *Run 840* from the fits of Figure 4.10 are 22% and 16% for respectively View 0 and 1. This poorer resolution observed with the 3×1×1 m^3 data might be explained considering the following contributions:

- **Stochastic fluctuations in the emission of electrons from liquid:** variations of the expected resolution as a function of the extraction fields have already been observed in the 3L [87], and are mainly due to the fact that the applied electric field alters the distribution of the kinetic energy of the electrons. Besides the qualitative observation of the 3L, no additional studies of this phenomena exists, at the best of our knowledge.

- **Stochastic fluctuations of the electron amplification:** in Section 2.3.2 it has been observed that the distribution of the final number of electrons produced in a Townsend avalanche depends on the initial number of electrons n_0 and the relative resolution

is independent on the detector gain and decreases with n_0 increasing. The map of Figure 4.6 (left) shows that the probability for an electron to be extracted and collected inside the holes of a LEM is smaller in 3×1×1 m^3 prototype than it is at the nominal settings. This means that fewer electrons are responsible for the amplification in the 3×1×1 m^3 LEMs and, as consequence, the relative fluctuations of the final avalanche size are larger.

- **Detector noise and signal processing:** the estimated charge from the reconstructed hits may be overestimated or underestimated according also to fluctuations of the noise, or reconstruction errors.
- **LEM gain uniformity:** since the gain observed fluctuates across the CRP, a local correction applied for each LAS may help in reducing the resolution observed on the total distribution of Figure 4.10.

It is not easy to unravel the relative impact on the detector resolution of the effects listed above. Biases from the reconstruction may be observed with the study of the resolution as function of the polar angle ϕ shown in Figure 4.13 (left). The resolution on both views is larger for extreme angles at $\phi < 20°$ and $\phi > 70°$ and almost equal at $\phi \approx 45°$, a trend similar to what was previously observed for the MPVs in Figure 4.11. A minimum seems reached by both view around 15%. This value can be interpreted as the detector resolution disentangled from the biases introduced by the track orientation. Figure 4.13 (right) shows instead the resolution as function of the amplification field calculated for different runs of the amplification scan previously described, taken at the same extraction and varying collection fields. On view 1 the measured resolution appears quite stable and fluctuates slightly above 15%. The same behavior is not observed on View 0, where much poorer resolutions are observed for low amplification fields. This is probably an effect introduced by the reconstruction, which was optimized for *Run 840*. Due to limitation in statistics it is not possible to observe the variations of the resolution as function of the extraction field or for the individual LAS. Data suggest that the limit resolution for the 3×1×1 m^3 detector is around 15%, however it is not possible to isolate the stochastic fluctuations related to gas processes from other type of random effects related to the signal formation and reconstruction.

4.5.8 Comparison with the expected distribution

The value of G_{eff} measured from *Run 840* is compared with the parametrization obtained with the field scans and found in good agreement. This value has been injected in the Monte Carlo simulation to reproduce the measured dQ/ds distribution of *Run 840*. The comparison between Monte Carlo and data is show in Figure 4.14 (left). The simulated and the reconstructed MPV differs by less than 2% on both views. Resolution is instead larger in data than in Monte Carlo. Our simulation has been done based on the previous studies which neglected possible fluctuations of the number of electrons introduced during extraction and amplification. Furthermore, the simulation assumes the absence of artifacts in the signal formation process (e.g wire cross talk) and consider an uniform gain for each LAS. As previously described, the experimental conditions of the 3×1×1 m^3 prototype do not allow

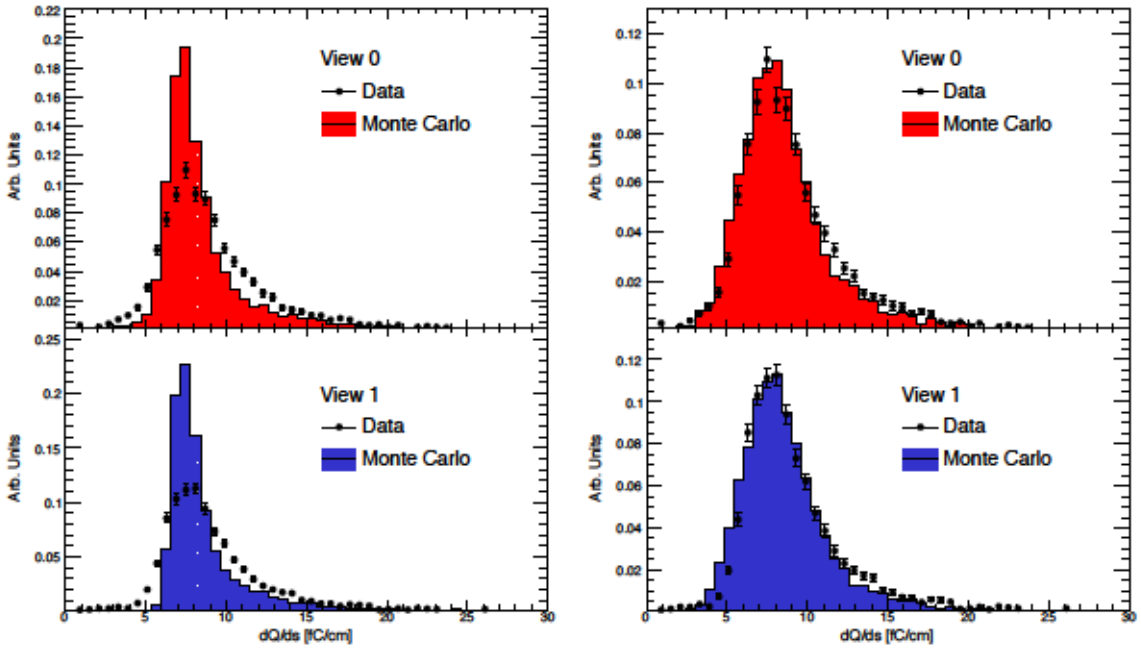


Fig. 4.14 Comparison of the distribution of the charge collected per unit of length in *Run 840* and the equivalent Monte Carlo sample without (left) and with (right) a Gaussian smearing of the charge (see text).

to unravel the relative impact of random contributions to the detector resolution, however the agreement with data is improved when a Gaussian smearing is applied *a posteriori* to the simulated distribution, as shown in Figure 4.14 (right). This suggests that the hypothesis of not accounted random fluctuations may provide a reasonable explanation to the poor resolution observed in data, however it is not possible to draw any further conclusion at the time this thesis is being written.

4.6 Summary of the performances of the 3×1×1 m^3 prototype

The 3×1×1 m^3 prototype represents the first large scale test of a 4 t dual-phase LArTPC with a three meter squared CRP. It has proven the overall good stability of the cryogenic conditions of the detector, pressure, temperature, and liquid level. Due to a faulty connector and loose wires on the extraction grid, it was not possible to test the nominal setting, nor to have long time high-voltage operations. The work described in this thesis is an attempt to use the data collected to assess the detector performances and understand its response. Results were compared with the dual-phase design requirements derived from performances of a smaller three liters dual-phase LArTPC with one single 10×10 cm^2 readout exposed to cosmic rays. The work focused in particular on the following aspects:

- The analysis of the characteristics, origin and stability of noise, with the goal to produce a realistic simulation
- The improvements of the reconstruction and selection of minimum ionizing particles: the $3\times1\times1$ m^3 has allowed to use the LArSoft reconstruction developed in Chapter 3 on real data for the first time.
- The analysis of the data collected.

Further comments are necessary regarding the last item of the above list: the analysis of the data was mainly oriented towards the characterization of the detector performances and the scan of the behavior for different high-voltage settings. The first analysis has been carried out using a relatively long run taken with the most stable detector condition possible. The effective gain before *charging up* is found to be 1.84, with a systematic contribution estimated at the order of 10%. It was not possible to isolate bias introduced by the reconstruction to prove the expected equal charge sharing between the two views of the detector. The gain recorded on different LASs of the CRP fluctuates within the design tolerance of 5%. The scan of the extraction field shows that data describes qualitatively the expected trend of the extraction efficiency convoluted with the transmission probability from the liquid interface to the LEM holes. The amplification scan has shown that the parameterization proposed for the 10×10 cm^2 LEMs holds quite well also for the 50×50 cm^2 LEMs of the $3\times1\times1$ m^3 prototype once the measured points are rescaled in order to have compatible run conditions and the differences in temperature and pressure are taken into account. The resolution of the detector is found to be around 15%, slightly larger than the expected 10%. The expected value for the resolution was however estimated for the nominal condition of the detector, therefore it is reasonable to assume that stochastic effects in gas may have different relative importance for different high-voltage settings. The Monte Carlo simulation is well in agreement with data for the MPV of the Landau distribution. A Gaussian smearing applied to the simulation improves also the agreement for the resolution.
With the $3\times1\times1$ m^3 detector it has not been entirely possible to disentangle the detector effects from reconstruction biases related to the small signal-to-noise ratio, nor to study the long term stability of the performance. The measured LEM amplification and extraction requires hand-made corrections to be compared with the expected relations, and thus they should be regarded more as qualitative trends. After the $3\times1\times1$ m^3 experience, the ProtoDUNE experiment is expected to finalize the characterization of the dual-phase design addressing the following topics:

- Measure the variations of the effective gain for an extended period of time on each individual LEM.
- Study the amplification with the 50×50 cm^2 LEM with a scan performed in the same condition of the 3L detector.
- Improve the detector simulation with a parametrization of both the amplification and the fluctuations of the number of electrons.

Chapter 5

Neutrino reconstruction and energy resolution with DUNE-DP

In this chapter the efforts towards a dual-phase oscillation analysis are described, with a particular focus on the energy reconstruction. The main goal to assess the performances of the full simulation and reconstruction developed in Chapter 3 for the first time on neutrino events and attempt to reconstruct the neutrino energy using the dual-phase setup. There is value in understanding the results using a full simulation because it is the only way to introduce the dual-phase advantages related to the signal processing on the *collection* views. The work here proposed constitutes a first step towards a study of the neutrino oscillation comprehensive of both single and dual-phase designs.
Section 5.1 describes the detector geometry used, the current simulation configuration accepted for the TDR studies and the neutrino sample produced. A remark on the dual-phase far detector acceptance to beam neutrino is also made, since this is relevant to understand the results of the reconstruction performance and the energy resolution. The reconstruction performance for primary particle tracks in ν_μCC events is detailed in Section 5.2, highlighting some potential issue in the detector design. The results are compared with a similar study carried out for the single-phase DUNE LArTPC. The reconstruction is then used to produce an estimate of the energy resolution and bias for ν_μCC and ν_e events. This is described in Section 5.4. The compatibility of the results obtained with the DUNE requirements, limits of the methodology, and possible improvements are also discussed in Section 5.5

5.1 DUNE-DP far detector neutrino samples

In this section, we describe the beam simulation and the Monte Carlo configuration used in preparation to the dual-phase oscillation analysis. The simulated sample has been produced bearing in mind the integration of the dual-phase detector into the neutrino oscillations analysis and therefore produced to be as conservative as possible and consistent with similar single-phase studies ongoing.

We present briefly the flux considered, the geometry and the simulation settings chosen. More information about the dual-phase simulation in LArSoft are given in Section 3.1. Eventually we will briefly comment about the detector acceptance.

5.1.1 Far detector geometry

The neutrino sample has been produced within the LArSoft framework using the configuration described in Section 3.2. The detector geometry used is a reduced copy (1:8) of the full detector geometry previously described. This is done in order to reduce the computing power required to simulate the full detector sample. This geometry is called *workspace* geometry and it features only 8 CRPs, arranged in two groups of four along z. This ensures an active volume which is large enough to have most of the events contained and therefore allows easier development of reconstruction and calibration tools that will be tested later with the complete detector geometry.

5.1.2 Monte Carlo settings

Neutrino fluxes and sample

The sample considered for the studies in this chapter is composed by 100k ν_μ and 100k ν_e, and an equivalent number of antineutrinos. The energy and angle of the incoming neutrinos at the far detector are taken from the unoscillated beam spectra at the DUNE far detector. The reference fluxes of beam neutrinos are consistent with the optimized beam design described in the CDR [46] and recommended for the TDR studies. A small rotation of the incoming flux has been performed in order to account for the unrealistic drift direction currently supported by the dual-phase LArTPC implementation in LArSoft (see Section 3.2).

GENIE configuration

The generation of the neutrino interaction at the DUNE far detector is handled by the GENIE event generator. The version used is v2 with all the models and simulated processes following the recommended settings of the collaboration. Some more details about the GENIE settings for DUNE are given in Section 3.3.

Particle propagation

The propagation of the particles that populate the final state is done within the GEANT4 v4 package. GEANT4 is fully integrated within the LArSoft framework and no different configuration settings are tested different from what previously described in Section 3.3.3. Light propagation is tuned off.
For what concerns the simulation of the gaseous Argon layer, effects related to extraction and induction have been turned off. This is done because, in the nominal configuration of the DUNE far detector, the extraction efficiency is expected to be close to 100%, and all electrons are extracted as *fast* electrons. A detector effective gain of 20 is set in the simulation which

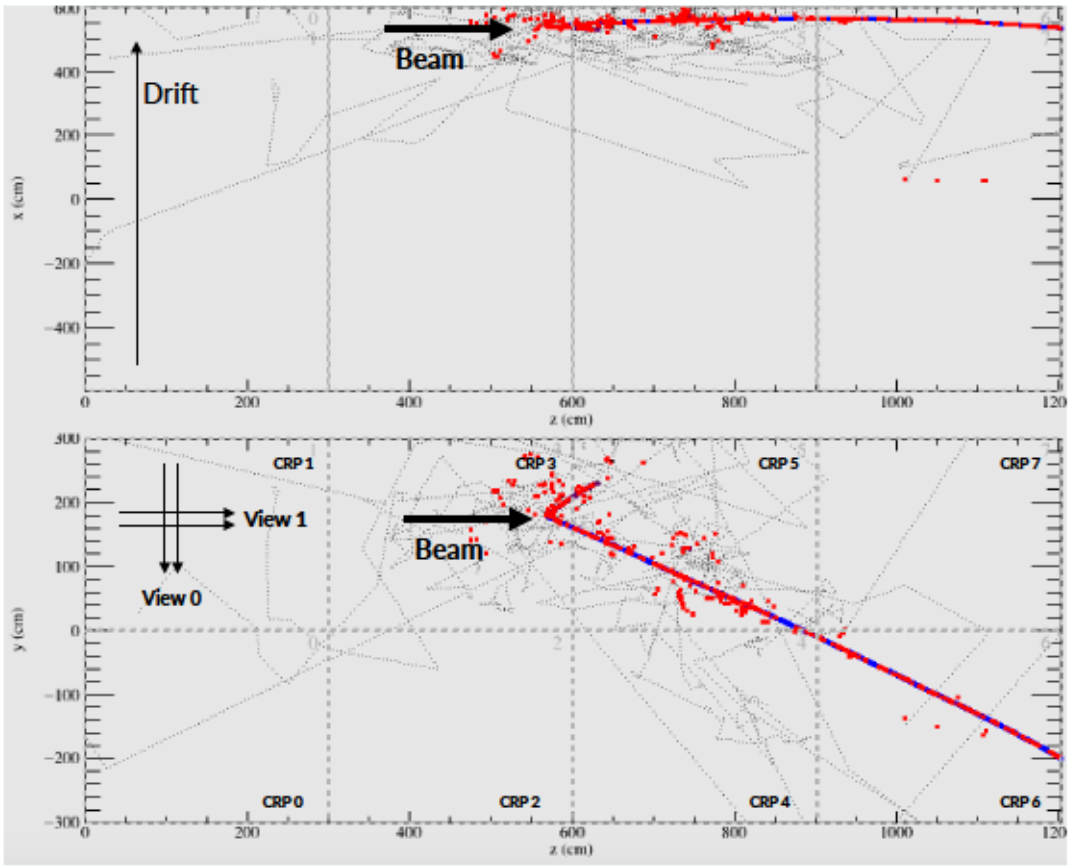


Fig. 5.1 The dual-phase DUNE far detector *workspace* geometry with a simulated ν_μCC event. The upper display shows the transverse drift x-z plane, while the lower display is the y-z readout surface. The y-z is divided into 8 3x3 cm^2 CRP regions separated by a 1 cm thick readout area with no readout, in order to simulate the frame between CRPs. The two readout strips samples respectively the directions along to y (View 0) and z (View 1). The neutrino beam direction is coming parallel to z. In the simulated ν_μCC event shown here, the red dots represent the energy deposits associated to electrons and photons, while the blue dots represent the deposits associated to muons.

the minimal value necessary to have a signal-to-noise ratio of 10:1 for the 12 meters drift given an electron lifetime of 3 ms. The value of the electron lifetime was chosen to be more conservative than the value measured in both the $3\times1\times1$ m^3 and ProtoDUNE single-phase detectors. Clearly, improvements on the lifetime are of direct impact on the minimal gain required to have Signal-to-noise ratio equal to 10:1.

Simulation of the detector response

The detector response is simulated as described in section Section 3.4. The signal current collected at the anode is convoluted with the shaping function of the electronics of equation 3.3 and digitized assuming a 13-bit ADC resolution. This is done in order to maintain the correct electronics correction factor, without saturating the ADC counts. This has the only consequence to increase the dynamic range while leaving unaltered the conversion between voltage and ADC counts. Also the shape of the waveform is expected to remain unchanged. A future implementation of the detector response should include a more realistic simulation of the correct electronics gain which uses a linear response below 400 fC and a logarithmic one above, as was described in Section 2.6.2.
An incoherent noise is added to the waveform based on the average noise pattern observed on 3 meters strips at $3\times1\times1$ m^3 detector after coherent noise removal. Incoherent noise fluctuations with root mean squared of 1.2 ADC are injected in the simulation using the simulation strategy described in Section 3.4.2.

5.1.3 Sample phase space and detector acceptance

In this subsection, the detector acceptance is briefly studied, since it has an impact on the reconstruction efficiency and energy resolution as discussed in the following sections of this chapter.
We first consider the angular distribution of the primary particles produced in beam events compared to the readout direction. In Figure 5.2 (left) it is shown that most of the particles produced by a neutrino interaction are emitted in a direction parallel to one of the readout views. The effect tends to be more prominent for primary muons than for other particles due to momentum conservation. About 88% of the primary muons produced in ν_μCC events are emitted in a $\pm10^\circ$ interval around the direction parallel to z. View matching for events produced in this topology is in general more complicated due to a more ambiguous projection of the event along the y direction. The impact of the readout orientation on the event reconstruction efficiency and reconstruction quality is described in Section 5.2.
Along with the readout direction, also the use of the smaller *workspace* geometry impacts on the expected detector acceptance to ν_μCC events with contained muons. It is possible to define the *inelasticity* of an event as the fraction of energy shared by the neutrino with the hadronic system. Due to the relatively small dimension of the *workspace* geometry, most of the muons contained inside the detector belong to events with large inelasticity. This unfortunately does not reproduce well the reality of the DUNE-FD, which is expected to have a larger fraction of contained muons regardless the inelasticity of the event Figure 5.2

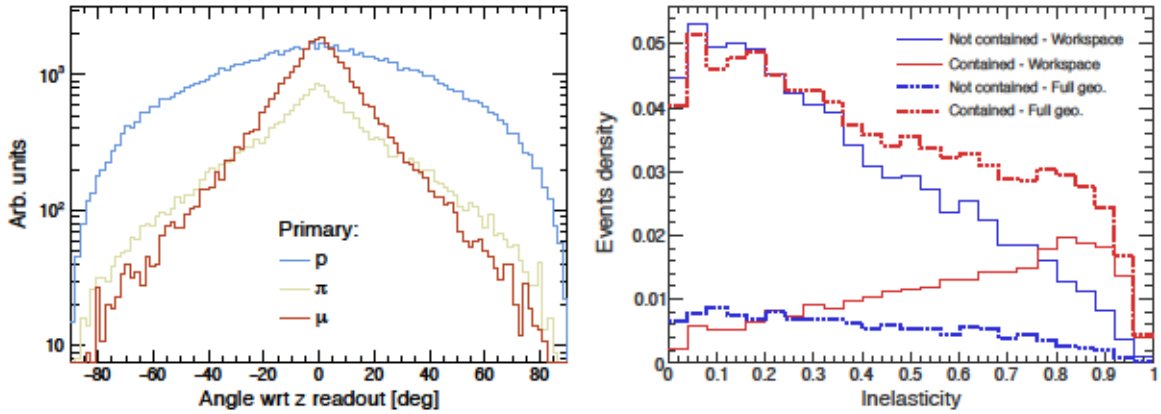


Fig. 5.2 Left: *inelasticity* distribution for ν_μCC events with respectively contained and not-contained primary muons in both the *workspace* and full detector geometry. The histograms are normalized to the total number of ν_μCC events. Right: distribution of the angle between the particle trajectory and the z axis for primary muons, pions, and protons emitted in ν_μCC events.

(right). A more complete discussion on the impact of the workspace geometry on the detector energy resolution is reported in Section 5.4.

5.2 Track reconstruction and study of the performances

The reconstruction algorithms described in Chapter 3 have been applied on the simulated neutrino events in the reduced *workspace* geometry of the DUNE dual-phase far detector. In this section, we focus on the reconstruction performances of track-like particles such as muons, protons, pions. Those particles appear in the TPC events with straight tracks topologies and therefore they are well suited to be reconstructed with the algorithms of Chapter 3.

Probing the possibility to reconstruct correctly straight track topologies is of primary importance to extrapolate important physics quantities, such as the stopping power and the particle momentum, which are used by algorithms of particle identification. An example of this is given in Section 5.4 where is introduced a technique to reconstruct the energy of ν_μCC based on the estimate of the primary lepton momentum. The total neutrino energy resolution would also benefit from the identification of individual track-like particle components inside hadronic cascades.

The reconstruction has been tested on the ν_μ sample produced with the setup of Section 5.1. The reconstruction performances obtained with the ν_μ sample are considered representative also for the $\bar{\nu}_\mu$ case, since the particles expected in the final state have similar topological features and angle distribution. Every event of the sample is processed using the same reconstruction chain of Chapter 3, with the fit-based hit finder, the TrajCluster algorithm for twodimesional clustering, and PMA for the tridimensional projection matching and trajectory fitting.

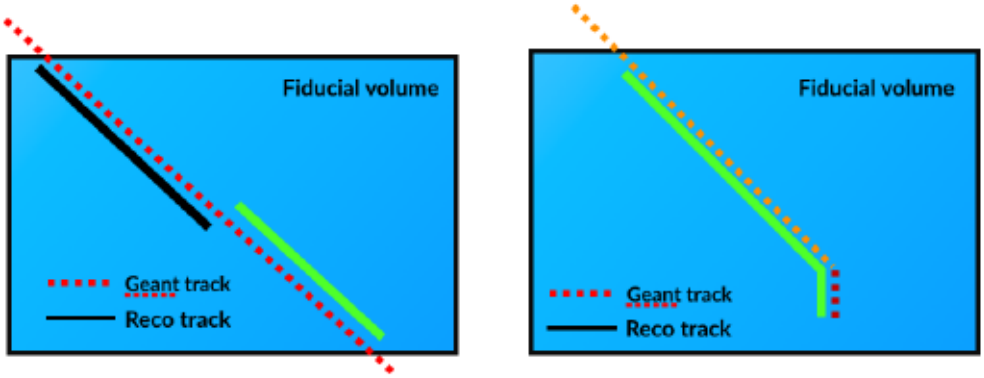


Fig. 5.3 A sketch of a track showing bad *Completeness* (left) and bad *Purity* (right). Dotted lines represent MCParticles and solid lines the reconstructed tracks. Different colors indicate different particles.

The estimation of the performances of the reconstruction is divided into three main steps. First a description of the metrics used will be given, then the performances will be evaluated on single primary particles, and finally the reconstruction quality is studied event by event. The results are eventually compared with those obtained with an alternative dual-phase detector geometry and those from a similar reconstruction applied to the single-phase LArTPC.

5.2.1 Definition of the quality metrics

Reconstruction performances are estimated by matching the reconstructed track trajectories with the simulated Monte Carlo particles (MCParticles). The matching is done at local level, by assigning each hit to the corresponding energy deposit inside the LAr volume. If more than one MCParticle has contributed to the same hit, only the largest energy deposit is considered.
For this study we use only the MCParticles which have been directly produced from the neutrino interaction, also called *primary* particles. We focus on those because they are expected to contribute directly to the energy estimation of ν_μCC events, in particular the primary muons. Also a primary MCParticle must have more than 5 hits assigned in each view to be considered "visible", and thus used for the matching.
Using the previously described matching criteria it is possible to calculate the following performance metrics for each reconstructed track:

- *Completeness:* the fraction of 2D hits of a MCParticle which are associated to the reconstructed tracks
- *Purity:* the fraction of 2D hits of a reconstructed tracks which are associated to a MCParticle.

Completeness and *Purity* are also simply described with the sketch of Figure 5.3. Both provide important insight on the track reconstruction quality: the first one defines whether

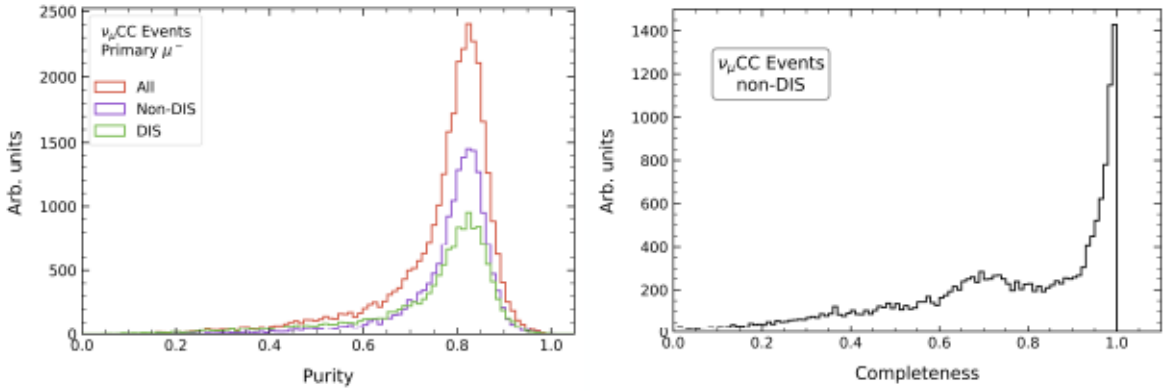


Fig. 5.4 The distribution *Completeness* (left) and *Purity* (right) for primary muons in ν_μCC interactions

the reconstructed trajectory is able to approximate entirely a MCParticle, while the second one defines the contamination on a trajectory coming from other MCParticles.
The reconstruction efficiency (ε) is calculated as the ratio between the number of *good* reconstructed particle tracks and the total number of MCParticles considered. A reconstructed track is defined *good* when it has *Completeness* larger than 10% and *Purity* larger than 50%.

5.2.2 Performances on individual reconstructed tracks in ν_μ CC events

The performance and the quality of the reconstruction are first evaluated on individual reconstructed tracks associated to a MCParticle. Because we are considering ν_μ CC events, we discuss the performances of reconstruction of primary muons separately from those of pions and protons. This is justified because in general muons tracks are longer and rather well-separated from the event, thus they are expected to be easier to reconstruct. Pion and proton tracks are usually rather short and localized in proximity of the interaction vertex.

Primary muons

The distribution of the quality metrics for the reconstructed tracks associated to the primary muons for the ν_μCC interactions is shown in Figure 5.4. The distribution of *Purity* from Figure 5.4 (left) has a most probable value around 0.8. This represents the intrinsic contamination along muons tracks which mainly comes from δ electrons and Bremsstrahlung photons emitted along the track.
The distribution of *Completeness* in Figure 5.4 (right) shows that muon tracks have in general good quality. There are, however, some situations where reconstruction can be troublesome and part of the trajectory may be lost, resulting in a low *Completeness* of the reconstructed tracks. This in general happens when there are not sufficient hits for the trajectory to be correctly fitted. It is possible to select two cases when this may occur:

- For particle travelling perpendicular to one of the readout projection, the energy deposited in the LAr volume is collected on a large number of different wires. Due

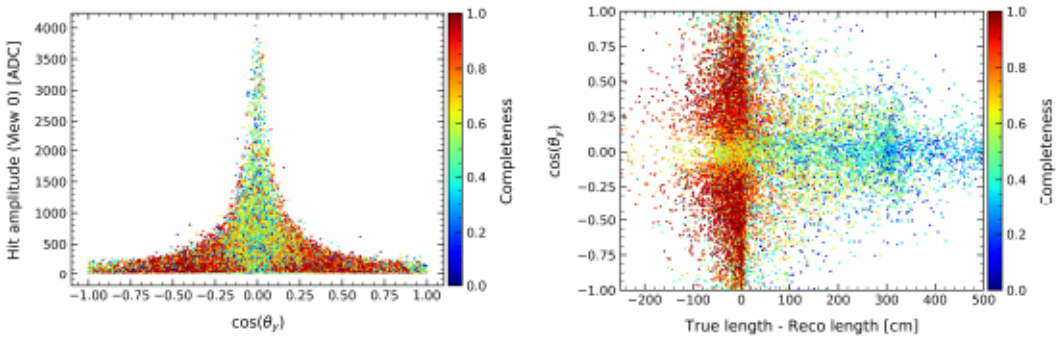


Fig. 5.5 Scatter plots with the *Completeness* calculated for each reconstructed track matched with a primary muons as function of $\cos(\theta_y)$ and the average signal intensity (left) or the residuals of the track length (right).

to intrinsic fluctuations of the energy deposit, it might happen that on some wires the signal-to-noise ratio of the resulting waveform may be not sufficiently large to be reconstructed as a hit. In a dual-phase large LArTPC this effect is enhanced if the track is located at the bottom of the detector and the effective gain is not large enough to compensate for the electron losses along the drift

- The number of hits is not large enough for the trajectory fitting, even if the signal-to-noise ratio of the signal is sufficiently large. This happens if the particle is travelling parallel to one of the readout projections or to the drift direction. In this cases, all the energy deposited is read only on few wires. Local features in the track topology (kinks, delta rays, ...) are in this case more complicated or impossible to resolve.

In the dual-phase LArTPC configuration proposed for DUNE, the two views are orthogonal (Figure 5.1). This means that the two cases described above happen simultaneously if a track is travelling parallel/perpendicular to one of the readout views. Unfortunately at the dual-phase DUNE LArTPC, tracks parallel to the z direction and perpendicular to y represent a large majority of the primary muons produced by the ν_μCC interactions, according to Figure 5.2 (right).
The scatter plot of Figure 5.5 left shows the completeness for reconstructed tracks matched with a primary muon as function of the average signal intensity and the direction cosine $\cos(\theta_y)$ with respect to the y direction (along View 0). At $\cos(\theta_y) = 0$ the track is parallel to z. The plot suggests that the reconstruction troubles are mostly related to the track topology along y, rather than a low signal-to-noise ratio along z. This hypothesis is also supported by the fact that no degradation of the reconstruction performance is observed as function of the drift distance.
Since the reconstruction of the neutrino energy described in Section 5.4 largely depends on the correct estimate of the leading muon momentum, it is important to understand the main reasons behind the reconstruction failures. In Figure 5.5 (right), the *Completeness* calculated for each track is shown as function of $\cos(\theta_y)$ and the residuals of the reconstructed

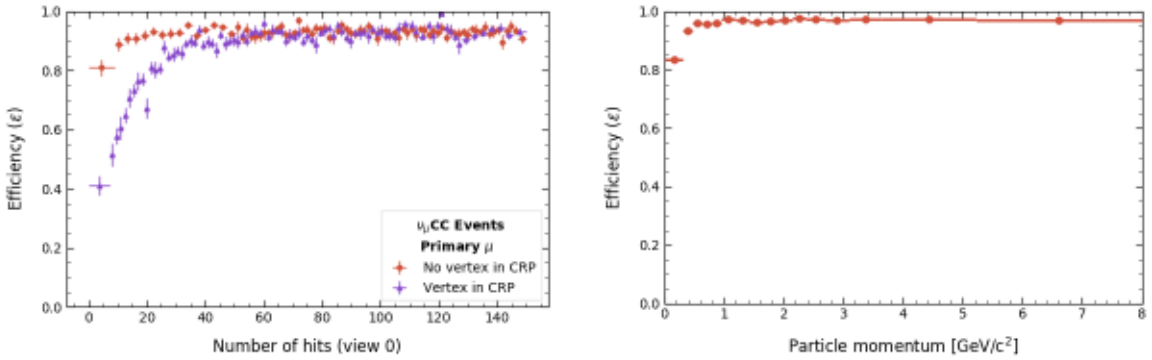


Fig. 5.6 The reconstruction efficiency of tracks associated to the primary muon as function of the number of hits found in CRP with or without the interaction vertex (left) and of the momentum (right) in ν_μ CC events.

track length. In the plot is possible to identify two regions, which provide insight on the reconstruction performances.
The first region is located at $\cos(\theta_y) \approx 0$ and residuals around 0, and contains mostly tracks whose fitted trajectory interpolates over part of the hits collection or for which a wrong plane match occurred. In both cases, even if a length similar to the true length was computed, some of the hits are not correctly reconstructed or do not belong to the track.
The second region is for residuals around 300 cm. Since this is equivalent to the size of one CRP, it is reasonable to suspect that this region contains tracks with portions that are poorly or not reconstructed around the vertex region. This is caused by the approach of PMA to projection matching, which stitches together the results from the reconstruction on each CRP individually. This clearly exposes the track reconstruction to failures on the single CRP region, which are more likely to occur in proximity of the vertex, due to a larger amount of information available. This represents a problem for the PMA algorithm, which aims mostly at finding optimal clusters matches between views.
Using the definition of efficiency previously given, it is possible to calculate the integrated track reconstruction efficiency for primary muons being 96%. In Figure 5.6 (left) the reconstruction efficiency for muon tracks is shown as function of the number of reconstructed hits associated to the track. The results are shown separated for each CRP in order to highlight the contribution given by the vertex region, following the considerations made previously for Figure 5.5 (right). The minimal number of hits required (on View 0) to have a reconstruction efficiency higher than 90% is 40 on View 0 for tracks in the vertex region, while it is just 15 for a track in CRP without the interaction vertex.
Figure 5.6 (right) shows the total reconstruction efficiency of the muon tracks as function of the lepton momentum. No significant reduction of the efficiency is found for momenta above 200 MeV.

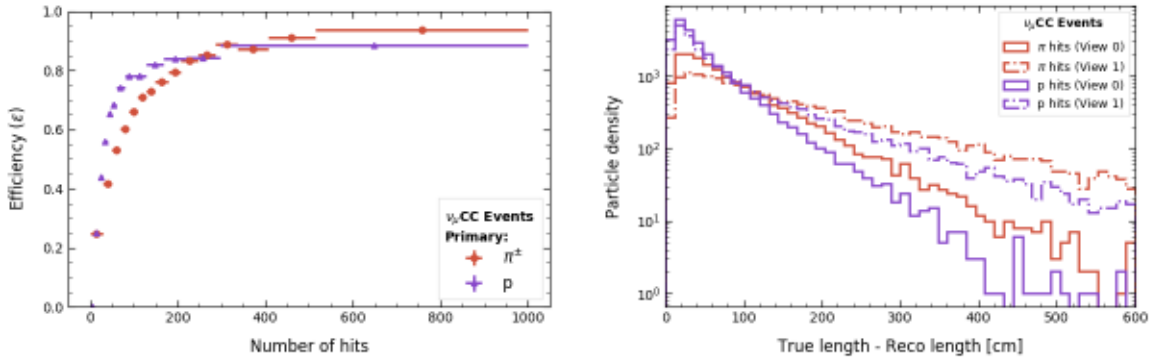


Fig. 5.7 The reconstruction efficiency of tracks associated with primary pions and protons as function of the number of reconstructed hits of the track (left). The distribution of the reconstructed hits associated to primary pions and protons on both views (right).

Primary pions and protons

The reconstruction of primary proton and pion tracks is in general more complicated than for muons. The reason for this are two. In low inelasticity events, protons and pions tend to appear as very short tracks close to the vertex and therefore more difficult to resolve. In high inelasticity events, pions and protons contribute to an intense hadronic activity, hence projection matching is more complicate without introducing ambiguities. Moreover, the previous observation made on the projection matching for the track trajectories holds also for protons and pions, although, in this case, the particle dispersion around the beam direction is wider than it was for muons.

The reconstruction efficiency as a function of the total number of hits is shown in Figure 5.7 (left). As for muons, increasing the total number of hits available gives sharp improvements to the view matching and trajectory fitting. From the distribution of the reconstructed hits associated to both protons and pions Figure 5.7 (right) it is possible that the main impact on the reconstruction is again given by the number of hits available along View 0. In case of low energy particles, those tracks have in general shorter length, which results in even shorter projections along this View.

The reconstruction efficiency as a function of particle momentum is shown in Figure 5.8 (left). It is clear that both for protons and pions there is a sharp increase of the efficiency as function of the momentum increase between 0 and 0.4 GeV, and 0 and 1 GeV for pions and protons respectively. This increase is clearly related to an increase of the track length and thus of visible hits available for the fit. After that, the efficiency reaches a plateau or even slightly decreases because of the increase of the inelasticity of the interaction, which results in a more complicated confusion around the vertex and an increase of complexity for the view matching algorithm.

The length residual for reconstructed trajectories associated to protons and pions is shown in Figure 5.8 (right). The distribution is rather broad, with a clear bias toward the negative tail. This is a sign for either a poor reconstruction in terms of *Purity* or view matching, since

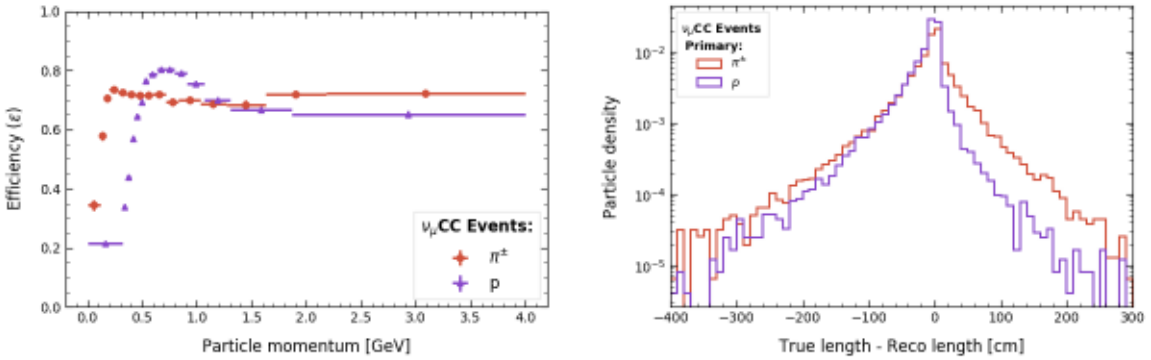


Fig. 5.8 The reconstruction efficiency of tracks associated with primary pions and protons as function of the particle momentum (left). The distribution of the residuals of the reconstructed track associated to protons and pions (right).

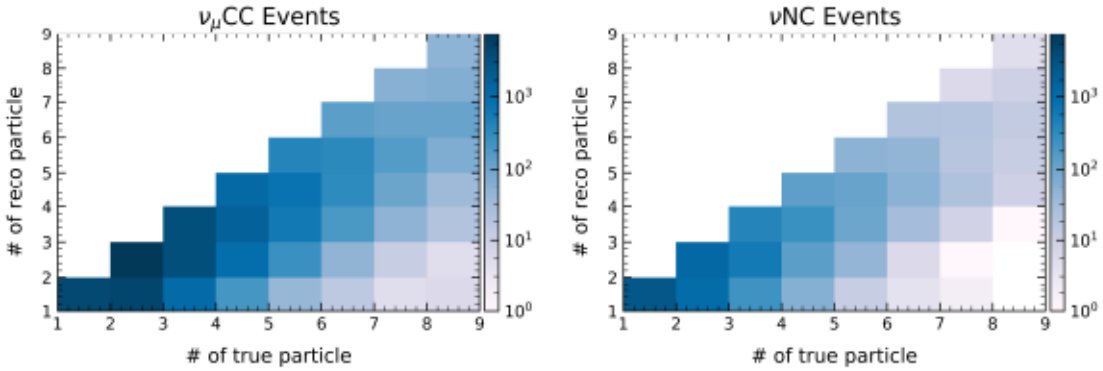


Fig. 5.9 The number of reconstructed particles matched with a primary particle in each event as function of the number of primary true particles in the event for ν_μCC (left) and νNC (right) interactions.

a portion of the track is associated to the energy deposit coming from other particles and, causing an increase of the estimated track length.

5.2.3 Global performance on ν_μ events

The previous considerations were made considering each MCParticle independently. In this section, We describe the performance of the reconstruction considering the full multiplicity of the final state. The number of true particles found in each event with at least 5 associated hits per view is compared with the number of reconstructed associated particles which pass the quality criteria defined in Section 5.2.1. The results are shown in Figure 5.9 for both ν_μCC and νNC events.
Figure 5.9 may be assimilated to a *confusion matrix* for the reconstruction. Most of the entries lie along the diagonal as expected for a good reconstruction. In events with little

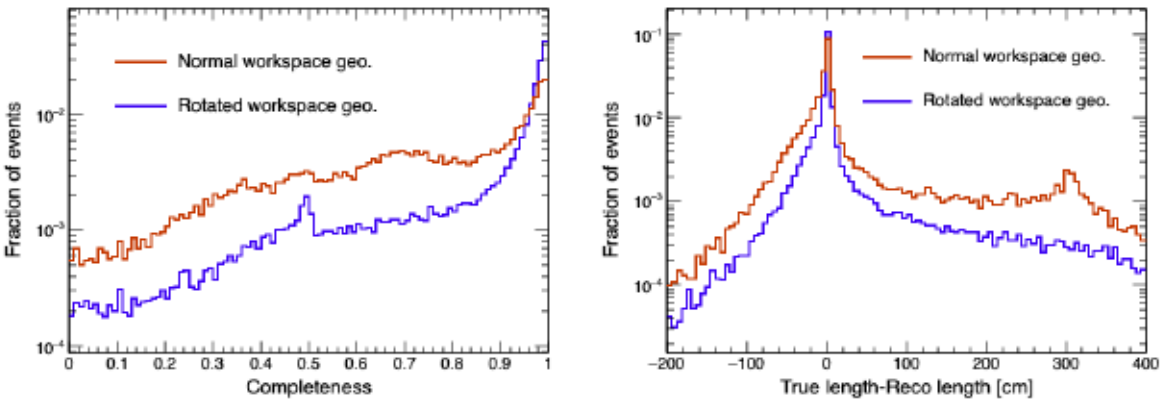


Fig. 5.10 The distribution of *Completeness* (left) and reconstructed length residuals (right) for the reconstructed track matched with the primary muons in ν_μCC events for both the standard and rotated *workspace* geometry.

multiplicity it happens that one of the primary particles is not identified. In events with large multiplicity this effect becomes more evident since those events tend to be dominated by Deep Inelastic interactions which usually produces highly energetic hadrons in the final state. Between the ν_μCC and νNC cases a small difference is observed. This is due by the high efficiency at which the muon tracks can be reconstructed in ν_μCC events.
The previous considerations give an overview of the performances of the reconstruction over the full neutrino events, and remarks once again the necessity to improve the reconstruction performances for secondary particles and nearby the vertex region.

5.2.4 Reconstruction performances with a rotated readout surface

The results on the reconstruction efficiency studies highlight a strong impact on reconstruction of both the reconstruction algorithm used and the particular detector geometry. In the proposed design of the dual-phase DUNE far detector, most of the tracks associated to the primary muon are parallel z and thus more complicated to reconstruct with good quality. Also the reconstruction of other primary particles such as protons and pions is particularly challenging, since most of the information of the event is contained in only one projection. This effect should be mitigated if the CRP were oriented differently with respect to the beam direction. In this section we study the reconstruction efficiency of primary particles in the case of a CRP with orthogonal views, both oriented at 45° with respect to the beam direction. A simple study reveals that the quality of the reconstruction slightly improves with a rotated geometry. This is particularly evident for tracks associated with the primary muons, as shown in Figure 5.10. The calculated track *Completeness* distribution has improved of about a factor 2 and the residuals distribution of the reconstructed length appears more regular, and most of the reconstruction artifact related to reconstruction ambiguities are mitigated in the new geometry.

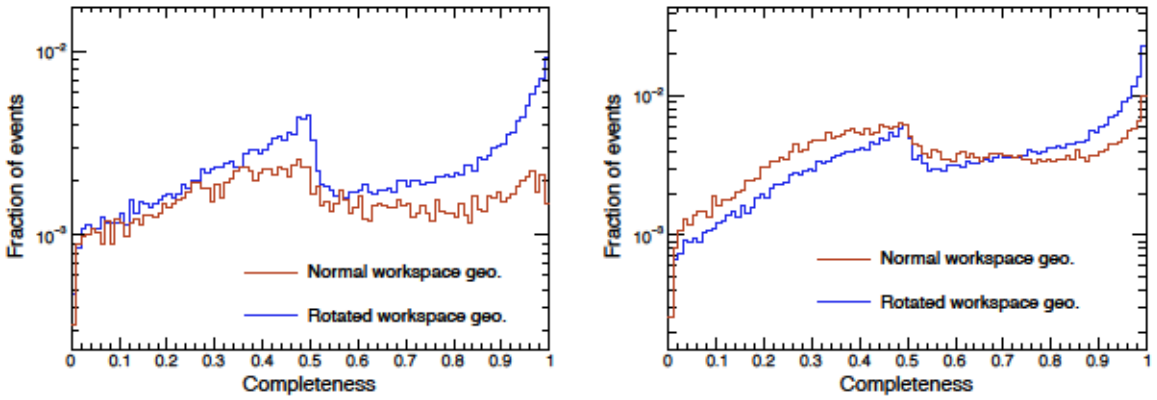


Fig. 5.11 The distribution of *Completeness* for the reconstructed track matched with the primary pions (left) and protons (right) in ν_μCC events for both the standard and rotated *workspace* geometry.

Some improvements are noticeable also for tracks associated with pions and protons, whose *Completeness* is shown respectively in Figure 5.11 left and right. In this case the rotated geometry bears less benefits than for muons. One possible explanation for this may be found analyzing how the Projection Matching Algorithm used for the tridimensional trajectory reconstruction works. The algorithm was originally conceived for single-phase LArTPCs, where the information available on three planes is available. This gives the advantage to rely on a third view to increase the robustness of the matching. On a dual-phase LArTPCs, with only two projections available, the algorithm has less information available to resolve projection ambiguities, with a consequent impact on the performances.
More studies are needed to assess the effective improvements brought by the use of the rotated *workspace* geometry to the reconstruction of tracks in neutrino events. At this stage, a hypothesis of an alternative design for the DUNE dual-phase charge readout seems mainly motivated by advantages in the reconstruction of muons track with better quality. It is important to notice, however, that to implement such a modification at the level of the anode is not straightforward for two main reasons:

- The system has been engineered to have 960x2 orthogonal board-printed strips with 3 m length. The strips length and arrangement is chosen to ensure a correct charge sharing and low noise level on both views.
- In case the strips are printed on the board at a different angle, this would result in a corresponding increase in the number of strips necessary to sample the same surface with the same interstrip pitch. Also the strip length across the CRP would vary, with a resulting increase of noise.

An alternative solution would be changing the orientation of the CRP during the installation phase of the detector. The system has however been designed to fully cover the surface of DUNE with 80 squared CRPs of the same size and altering their orientation would result

Table 5.1 The reconstruction efficiency for different primary particles in the DUNE dual-phase LArTPC with the standard and rotated *workspace* geometry compared with the results obtained for the single-phase DUNE LArTPC with a comparable reconstruction [141].

s Primary	Efficiency (%)		
	Standard geometry	Rotated geometry	Single-phase LArTPC
Muons	96.0	96.5	98.2
Pions	69.4	70.4	≈ 73
Protons	63.9	71.5	≈ 70

in several dead areas, that would be necessary to fill up with smaller pieces of the full CRP. How this whole system would be integrated correctly poses a significant technical challenge, and other disadvantages may arise within this setup.

5.2.5 Summary of the reconstruction performances and outlook

A full reconstruction for tracks at the dual-phase DUNE far detector was tested for the first time on a simulated sample of ν_μCC beam events. The performance of the reconstruction was tested using two quality metrics: the *Completeness* to study the integrity of the reconstructed track when compared to the original GEANT4 particle, and the *Purity* to estimate the level of contamination from other GEANT4 particles inside the reconstructed track.
Table 5.1 shows the integrated efficiencies for muons, pions, and protons in the case of the standard and rotated detector geometry. The values obtained for a single-phase LArTPC using an equivalent reconstruction are also shown.
Table 5.1 highlights overall a good reconstruction efficiency for reconstructed tracks associated to the primary muon. This is an important result because it allows to use reconstructed quantities as proxies of the muon momentum in order to produce a credible estimation of the energy of muon neutrinos. This is described in Section 5.4. However, because of the detector acceptance of beam neutrinos, most of the primary muons are emitted in a direction parallel to one of the planes of readout strips. This poses a significant challenge to the reconstruction of tridimensional trajectory which impacts on the quality of the physical information that can be extracted for higher level reconstruction and analysis. The reconstruction for other primary particles such as protons and pions is not yet sufficiently efficient. Future investigation are necessary to clarify whether three readout planes instead of two are necessary for a better reconstruction.
The work developed in the context of this thesis provides strong evidences that implementation of specific models for view matching for dual-phase LArTPC are required in order to fully exploit the potential of DUNE. In particular two main threads should be followed:

- **Pandora** is a multi-algorithm pattern recognition software which has been already applied with success on single-phase LArTPC data and which now has been adapted to accept dual-phase inputs. As described already in Section 3.5.5, tests are ongoing to assess the performances on Pandora on the ν_μCC sample.

- **Deep-learning** offers the possibility to create a reconstruction that is able to learn relevant patterns directly on data. With this approach usual low level clustering and projection matching would be bypassed and a full classification of the entire event or of the single particles might be provided.

5.3 Shower reconstruction

Electrons and photons are usually easy distinguishable from other types of particles due to their shower-like topology. In this case reconstruction aims at separating track-like and shower-like energy deposits, and at identifying the primary vertex and orientation of the shower cone. The former can be achieved using the CNN model for pixel semantic classification already described in Section 3.6 and used also for the dual-phase reconstruction of neutrino events.
Regretfully, unlike for the case of tracks, there is not currently in place for dual-phase a reliable algorithm to reconstruct tridimensional showers in neutrino events. Efforts on Pandora aim mostly at including soon this possibility within the reconstruction chain.

5.4 Energy reconstruction

The oscillation probability depends on the baseline and on the energy of the neutrino, therefore it is paramount to measure the neutrino energies with great accuracy and resolution. In an experiment such as DUNE, bias in the energy scale or a poor resolution can washout the relevant patterns present in the probability spectra and be compatible with a different combination of the oscillation parameters than the one chosen by nature. Previous studies [142, 143] have shown that a $\approx 20\%$ underestimation of the hadrons missing energy could results in a $\approx 1\sigma$ bias in the extracted δ_{CP} value. More detailed analyses carried out in the context of DUNE [144] have highlighted the need to constrain the lepton energy scale within the 1% level. It is also pointed out that, due to the different elasticity distribution for neutrinos and antineutrinos, even a 3% bias in the hadronic scale has important effects on the estimation of δ_{CP}. Moreover, calibration uncertainties have to be constrained around the 1-2%. This poses the significant challenge to finely monitor local charge variations across the large DUNE surface. ProtoDUNE will pursue strategies in order to understand, control and eventually reduce possible uncertainties of calibration.
For the work of this thesis, our main goal is to show a method for the energy reconstruction at the DUNE far detector to be deployed in the context of the Long baseline (LBL) oscillation analysis, which uses the dual-phase reconstruction previously described. The energy is reconstructed using a combination of muon momentum and calorimetric approach for the hadronic and electromagnetic contributions. The momentum of the primary muon is approximated using the reconstructed range or Multiple Coulomb Scattering (MCS). Possible improvements of this method rely on the possibility to reconstruct the momentum of individual particles inside the hadronic shower. This aspect is however not explored in this thesis because of the sub-optimal reconstruction efficiency and quality for particles others that

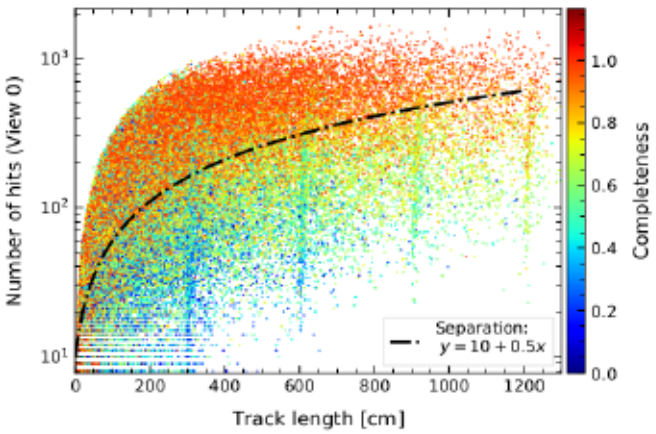


Fig. 5.12 Distribution of the *Completeness* of each reconstructed track associated to a primary muon as function of the reconstructed length and the number of hits on view 0. The straight vertical lines pattern corresponds to tracks crossing the boundaries of the CRPs.

muons and the absence of effective particle identification strategies in dual-phase that would work with the full reconstruction previously proposed. For hadronic and electromagnetic cascades the calorimetric approach is instead used, which consists in summing all the visible charge contained in the reconstructed hits found in the event and correct it using Monte Carlo to estimate the true hadrons energy.
In the following section we consider three classes of events: ν_μCC events with primary muon contained into the fiducial volume, ν_μCC events with primary muons escaping the fiducial volume, and ν_eCC events. We prefer to separate the case of ν_μCC with contained and escaping muons because of the different inelasticity distribution at the *workspace* geometry used.

5.4.1 ν_μCC energy reconstruction using track momentum and hadronic energy

Sample selection

The energy of muon neutrinos is reconstructed with higher resolution by separating the hadronic contribution from the leptonic contribution of the main muon. Primary muons are often well identified in ν_μCC event with long straight tracks compared to the rest of the primary particles. Since a precise analysis of the identification performances of muons in a LArTPC is beyond the scope of this report, we just naively assume that the longest reconstructed track found in the event is a muon. However this might not always be true and the longest track might be matched with protons, pions, and less often with other particles. This in general occurs when muons have lower momentum than the associated hadronic system, or when the main muon track is broken into smaller pieces or not reconstructed or identified at all.

Table 5.2 Definition of the sample used for the ν_μCC energy studies, including also the percentage of true particles associated to the longest reconstructed track and the size of the total sample after the quality cut (see text).

	Contained (%)	Not contained (%)
Muons	75.53	91.98
Pions	7.68	2.70
Protons	9.92	3.30
Others	6.86	2.02
Total	11264	17942
Selected	5147 (45.69 %)	7365 (41.05 %)

For the study of ν_μCC events we consider a sample of 32820 events with vertex contained within the fiducial volume of the TPC, defined by applying a cut to reduce the dimension of the *workspace* geometry volume of Section 5.1.1 to $|x| < 550$, $|x| < 250$, and $50 < z < 1100$. This choice was done to ensure a better containment of the hadronic part of the event.
The sample is then further divided into events with the longest track contained within the fiducial volume, and events with longest track escaping the fiducial volume. The complete sample definition is reported in Table 5.2, with also the fraction of the different true particle associated to the longest track. Since there is no event selection or particle identification strategy in place, we ignore the contribution from the mismatched muons for the resolution estimation.
One last quality selection is applied to the leftover muon tracks. This cut is performed in order to select only tracks with the highest *Completeness* in the sample. In Section 5.2 it is shown how track completeness is strongly correlated with the track direction parallel to z. For this cut we simply compare the number of hits along view 0, which are a good proxy of the projected length along the direction perpendicular to z and the total reconstructed length. Tracks travelling parallel to z are expected to have quite long reconstructed length and a relative smaller number of hits collected on view 0. Figure 5.12 shows the *Completeness* distribution of the reconstructed tracks associated to the primary muon as function of the reconstructed length and the number of reconstructed hits associated to the track on view 0. The separation line of Figure 5.12 is used to select tracks with large *Completeness* by using only reconstructed information. The number of tracks after the quality cut is also shown in Table 5.2.

Calibration of the muon momentum using the reconstructed longest track

The absolute momentum of fully contained reconstructed tracks is estimated by the reconstructed range l_r of the track. The calibration is performed using the Monte Carlo simulation over the sample of contained events of Table 5.2, after selecting only the true contained muons tracks. The reconstructed ranges are then divided according to their true momentum p_t into bins from 0.2 to 2 GeV as shown in Figure 5.13 (left), and a Gaussian fit is performed

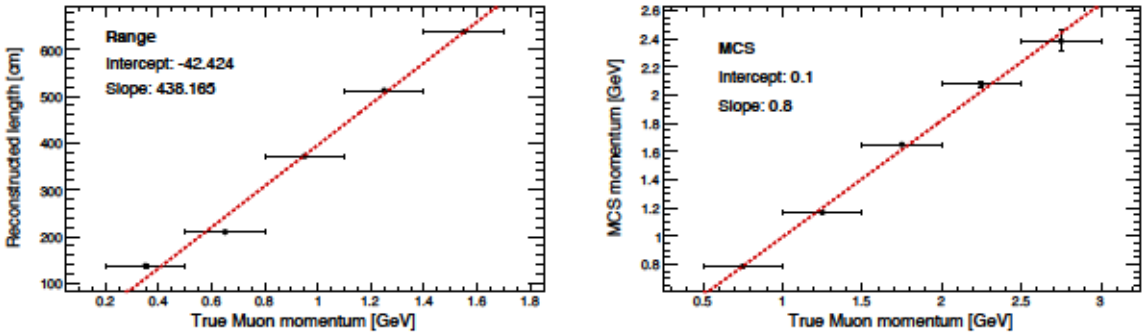


Fig. 5.13 Calibration of the momentum in ν_μCC events using the reconstructed range (left) and the MCS (right). Fit errors are negligible for this analysis.

on the distribution of each bin. In Figure 5.13 (left) the x are the center value of the true momentum bin and the y the corresponding means of the Gaussian fit. The reconstructed momentum p_{range} is then calculated as:

$$p_{range} = \frac{l_r - a}{b},$$

with a and b being respectively the intercept and the slope taken from the linear fit of Figure 5.13 (left).

For events with an exiting longest reconstructed track, the momentum is estimated using the average Multiple Coulomb Scattering (MCS) angle. This is calculated using a dedicated algorithm originally implemented for the MicroBooNE experiment [145]. The algorithm divides the track into segments of equal length and fitting them with a straight line. The scattering angle is then estimated as the angle between successive fitted segments. This is larger at lower momenta and smaller at higher momenta, for this reason the algorithm only performs well for muons below 4 GeV. It is possible to recover the true particle momentum p_{MCS} from the momentum reconstructed by the algorithm in similar fashion to the case of the contained tracks, as shown in Figure 5.13 (right).

About one third of the contained longest reconstructed particle is actually the broken piece of a longer exiting track. This common case is easily confused with real stopping muons and has the consequence of spoiling the energy resolution. The effect is mitigated by comparing the momenta reconstructed with both the range p_{range} and MCS p_{MCS} techniques. The two estimates are expected to agree for true stopping muon tracks and to disagree if the reconstructed track is just the broken portion of exiting muon. For this latter case, p_{range} will in fact underestimate the true momentum. For the analysis of the energy, we define the ratio $R = p_{range}/p_{MCS}$ and we impose p_{MCS} to be the momentum of a stopping muon if $R < 0.7$.

The momentum resolution is estimated from the fractional residuals from the best fit of Figure 5.13 (left) and Figure 5.13 (right) for respectively contained and exiting reconstructed tracks. Resolutions and biases are found from a Gaussian fit on the residual distribution to be around 6% for contained tracks (Figure 5.14 left) and about 24% for exiting tracks (Figure 5.14 right). Bias in both cases is at the per mille level.

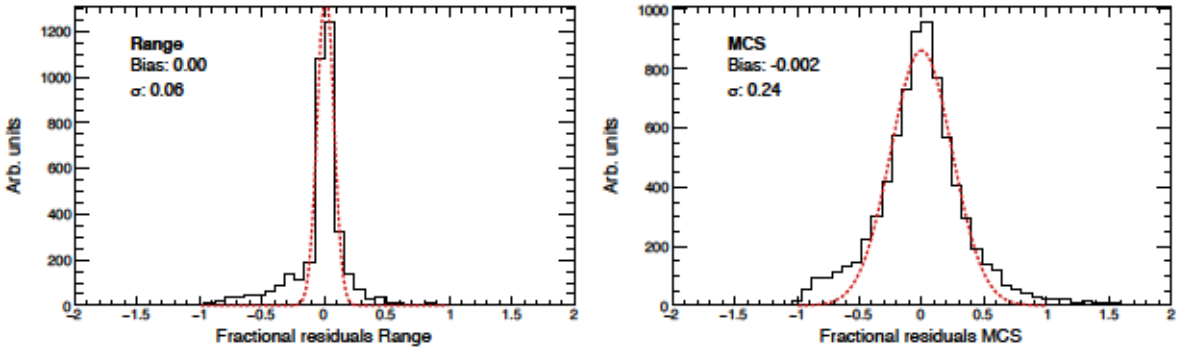


Fig. 5.14 Distribution of the fractional residuals of the reconstructed muon momentum using the reconstructed range (left) and the MCS (right). Fit errors are negligible for this analysis

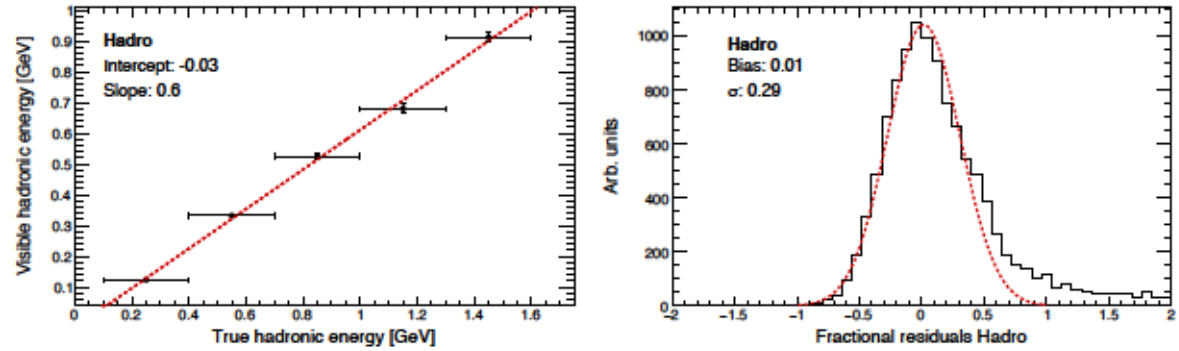


Fig. 5.15 Left: reconstructed hadronic energy in ν_μCC events. Fit errors are negligible for this analysis. Right: fractional residuals for the energy estimation of the hadronic system.

The small accumulation on the left tail of Figure 5.14 (right) is due to tracks whose momentum is underestimated. This happens for tracks with very short projections on view 0. The MCS algorithm highly depends on the trajectory fit of the tracks which usually is not perfect along those directions. There is also a contribution from high energy muons, which escape the detector close to the end point of the fiducial volume along z. For muons above 4 GeV, the resolution of the detector is not sufficient to use the scattering angle to determine the particle momentum.

Calibration of the hadronic energy

The hadronic contribution to the neutrino energy is estimated from the total charge of the event after removing the contribution of the longest track and making the necessary corrections for gain, electron lifetime and quenching. The reconstructed hadronic energy tends to be lower than the true value of the hadronic system contribution estimated as the difference between the neutrino and lepton energies. This happens because of the missing energy from neutral particles and other types of hadronic interactions such as pion absorption. Figure 5.15 (left)

shows an attempt to recover the true energy from the reconstructed energy with the technique previously used for the momentum reconstruction. A correction made in this way however only accounts for average energy losses and not for event-by-event fluctuations. Furthermore, it will assume a linear dependence of the visible energy with respect to the hadronic energy deposited, which is often not the case. This assumption introduces a 1% positive bias in the residuals distribution of Figure 5.15 (right). This happens mainly because our calibration tends to overestimate a little the true hadronic energy below 1 GeV. In these conditions, the overall hadronic energy resolution of the detector is estimated to be around 29%.

Resolution on ν_μCC events

The muon momenta and the hadronic energy measured as previously described are combined together to obtain the total energy of ν_μCC events. The distribution of the fractional residuals between the true ν_μCC energy and our estimation is used to obtain the bias and resolution of the neutrino energy reconstruction. The overall distributions are shown in Figure 5.16 (left) and Figure 5.17 (left) respectively for ν_μCC with a contained and exiting muons. The total resolutions for the two cases are estimated to be 17% and 23%, and the average bias is constrained at the per mille level. In Figure 5.16 (right) and Figure 5.17 (right) the bias and resolution fluctuation are also estimated for bins of true ν_μ energy in a range between 0 and 8 GeV for contained and exiting muon.
The resolution on ν_μCC events energy with a contained muon is largely dominated by the hadronic part. This becomes even more important at high energy, as it is possible to notice in Figure 5.16 (right). The reason for this behavior is a consequence of the *workspace* geometry used for this analysis, which represents about 1:8 of the total far detector geometry. As shown in Figure 5.2 (right), the inelasticity distribution observed with the *workspace* geometry is very different than the one expected at the far detector cryostat. Since a muon has to have a relatively low momentum, $\lesssim 3GeV$, to be contained inside the *workspace* geometry, most of the energy is in general shared with the hadronic system causing an increase in the observed resolution.
Almost an opposite trend is noticeable in Figure 5.17 (right), where the resolution of neutrinos with exiting muons is more dominated by the resolution on the muon momentum and it is slightly poorer at low energies. This is again a consequence of the inelasticity distribution, which tends to be slightly larger at low energies for ν_μCC events with exiting muons. This happens because at low energies, a muon is more likely to be contained inside the detector, and thus the only few not-contained events are related to events close to the fiducial volume boundaries with muons escaping from one of the borders.
To conclude our estimate of the ν_μCC energy resolution, we try to indicate possible sources of systematic uncertainties and calculate their impact on the final measure, in particular on the estimated bias. We consider only three relevant sources of uncertainties for this analysis which affect the neutrino energy estimation:

1. **Range momentum**. The uncertainties on the reconstructed track length are estimated as the means of the residuals distribution over the track start and end position. The two

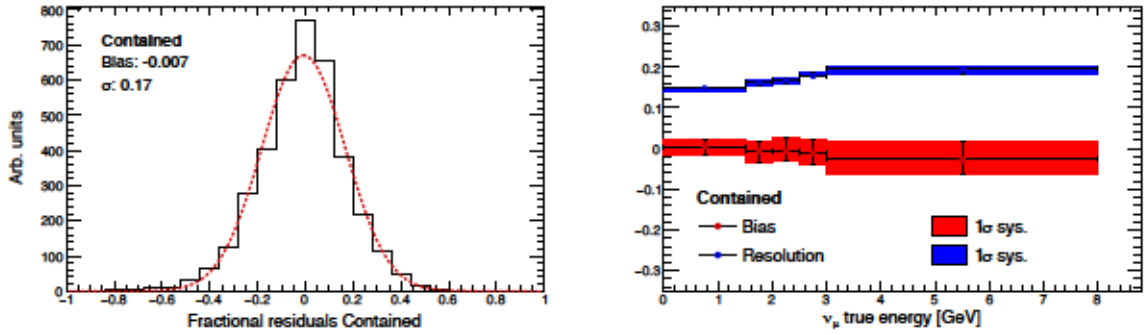


Fig. 5.16 Left: distribution of the fractional residuals between the true energy and the estimated energy in ν_μCC events with a contained muon tracks. Fit errors are negligible for this analysis. Right: estimate of the fractional bias and resolution for bins of true ν_μCC events with a contained muon.

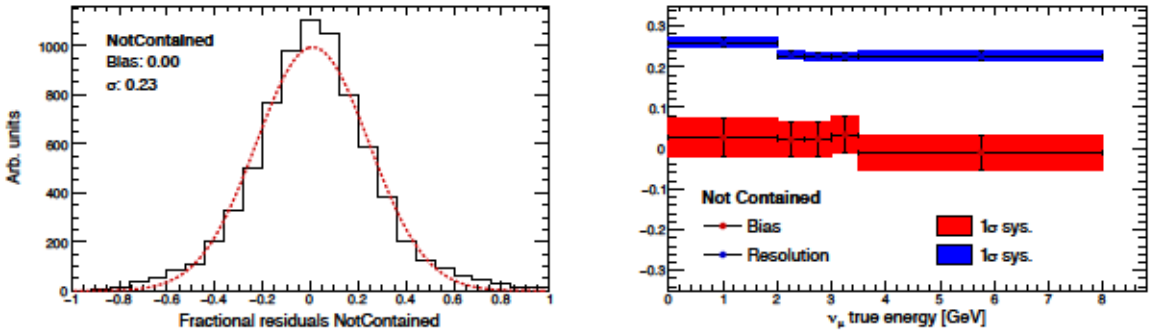


Fig. 5.17 Left: distribution for the fractional residuals between the true energy and the estimated energy in ν_μCC events with an exiting muon tracks. Fit errors are negligible for this analysis. Right: fractional bias and resolution for bins of true ν_μCC events with an exiting muon track.

means are summed in quadrature yielding to a total error over the range of 3.70 cm, that impacts with an uncertainty of $\pm 1\%$ over the estimated value for the momentum.

2. **MCS momentum**. The MCS algorithm provide an estimate of the momentum of $\approx 3\%$. The value has been calculated by comparing the estimated momenta using range and MCS in a pure sample of stopping muons inside the MicroBooNE detector [145]. We consider the uncertainty found by MicroBooNE valid also in our case.

3. **Hadronic energy**. Within our assumptions, the main systematic uncertainty on the hadronic energy estimation comes from calibration. For this study we use the values of ICARUS as indicative of the uncertainties level in a LArTPC, however DUNE dual-phase specific errors must be used for future updates of this work. The average energy to produce an electron ion-pair is $23.6^{+0.5}_{-0.3}$ eV [71]. This translates into an error at the level of 2%. ICARUS measured the uncertainty on the electron lifetime at the level of 4.9% and on the recombination factor at the level of 2% [74]. Finally the uncertainty on the electronics calibration factor has been estimated to be 1.3% in the ICARUS detector [74]. The quadrature sum of the four uncertainties give a total uncertainty level at the order of 5.8%.

The total contribution of the systematic to the measure of the neutrino energy has been calculated by adding to the estimated values of energy in the residual distributions of Figure 5.16 (left) and Figure 5.17 (left) an appropriate systematic contributions and comparing biases and resolution with the original distributions. The overall systematic uncertainty on the neutrino energy is at the level of 3.5% for the contained sample and 4.4% for the not-contained one. Most of the contribution to the systematic uncertainty on the energy of the contained events sample is brought by the hadronic system. This is again a consequence of the detector *workspace* geometry used that tends to favor contained events with high *inelasticity*. This is also noticeable in Figure 5.16 (right), where both the resolution and the errors on the bias tends to be larger in the last bins. For the not-contained sample, the hadronic part alone contributes for $\approx 2.4\%$ of the total systematic error. In this case no noticeable fluctuation of the error on the bias is visible as function of the true neutrino energy in Figure 5.17 (right).
The total energy resolution along with the associated systematic error is expected to improve when the full DUNE-FD is considered for three main reasons:

- More ν_μCC events are expected to be contained even at low *inelasticity*.
- It is possible to optimize the fiducial volume boundaries along z to improve the resolution of ν_μCC events with escaping muons with shorter tracks.
- The fiducial volume may be tuned to ensure a better containment of the hadronic showers.

It is not possible to extrapolate the value of resolution calculated with this analysis to the full scale far detector geometry in an accurate way. A simple test was performed by pre-selecting all neutrino vertices contained in certain areas of the detector to study the resolution and

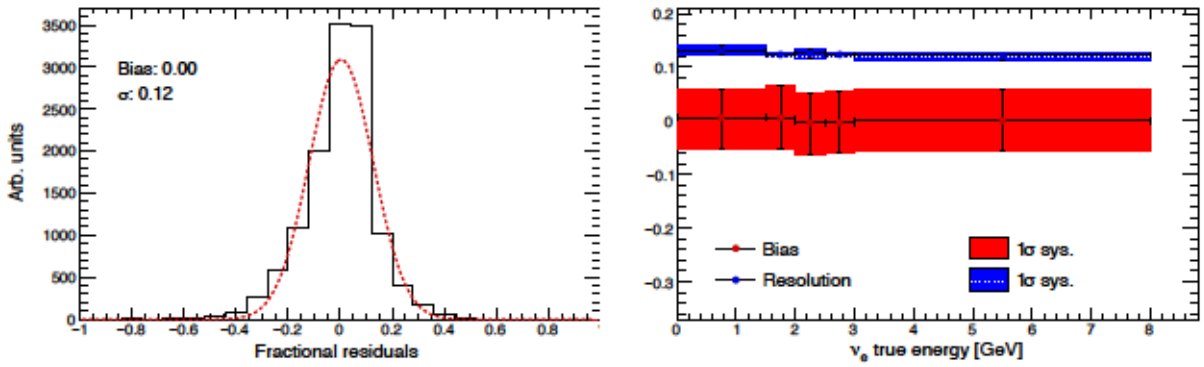


Fig. 5.18 Left: distribution for the fractional residuals between the true energy and the estimated energy in ν_eCC events. Fit errors are negligible for this analysis . Right: fractional bias and resolution for bins of true ν_eCC events.

compare the values found with those previously calculated. However no improvement larger than 1% was observed. The reason for this is the quality selection cut on the reconstructed tracks that tends to discards most of the tracks which travel along the horizontal direction, making more difficult to exploit any improvement to the resolution that may arise by reducing or increasing the fiducial volume cut along z. A complete study of the detector acceptance and the impact of the fiducial volume on the energy resolution will be addressed in future studies.

5.4.2 ν_eCC events

The same study described in the previous section can be repeated for the ν_eCC sample. We use the calorimetric method to evaluate the total visible energy from the electromagnetic cascade associated to the primary lepton and the real energy is estimated using a Monte Carlo based correction. The same procedure is applied to the hadronic energy, which is then summed to the lepton contribution to access the primary neutrino energy.
The distribution of the fractional residuals between the estimated and true energy of the ν_eCC sample is shown in Figure 5.18 (left). The energy resolution calculated with this method is estimated to be 12% with a bias smaller than 1%. In Figure 5.18 (right) both resolution and bias looks very stable across the true ν_eCC energy range considered. The error bands in Figure 5.18 (right) are calculated considering the same 5.8% of calorimetric uncertainty previously described on both the electrons and hadrons energy. This impact severely the error associated on the energy bias up to 10%. Such a large error it is not compatible with DUNE and it is expected to decrease once further measures will be taken with ProtoDUNE. Finally, it is important to remark that for this study, the reconstruction effects specific for the dual-phase appear mostly at the hit level. Since a suitable reconstruction for electromagnetic showers is not yet in place for a dual-phase detector, cheated information were used to cluster together all hits belonging to the primary lepton. For this reason, the resolution of energy is

expected to worsen a little, since reconstruction algorithms usually are not able to cluster together all the energy contribution from a single shower.

5.4.3 Impact of the gain

A simple study to consider the impact of the gain in gas Argon has been carried out using electrons simulated in fixed values of energy bins at 1, 2, 3, 5 and 7 GeV. There are 1000 particle in each energy bin. The simulation is performed in the far detector *workspace* geometry and direction and starting position of the particles has been chosen to maximize the showers containment.
In order to be identified as hit, a given peak of the waveform must have an amplitude larger than 4 ADC counts. This threshold has been arbitrary chosen during the hit-finder tuning for the reconstruction of the neutrino sample. The visible energy of the event is reconstructed from the charge value calculated from the reconstructed hits and calibrated in similar a way to the electromagnetic part of ν_eCC samples. The contribution of systematic errors has also been evaluated in a similar fashion to what previously described for the ν_μCC and ν_eCC analysis.
The samples are produced at 4 different values of detector gain: 5, 10, 20, and 50, as shown in Figure 5.19. for each energy bin, the distribution of the fractional residuals between the reconstructed and true energy of the electrons has been fitted with a Normal distribution and the resolution estimated from its standard deviation. The variations of the resolution as function of the energy are then fitted using the calorimetric equation:

$$\frac{\sigma}{E} \approx \frac{A}{\sqrt{E}} \oplus B.$$

A is the *stochastic* term, accounting for the random fluctuations of the energy deposits in matter, and B is the constant term.
From the fitted values of A and B it is possible to appreciate a steady increase of the measured resolution is observable as function of the gain increase. With the given simulation settings, the minimal gain required to detect a *mip* with a signal-to-noise ration larger than 10:1 is 20. A suboptimal configuration of the gain has immediately a strong impact on the resolution of the electromagnetic shower, especially while approaching the sub-GeV region. This remarks the importance of reaching the nominal signal-to-noise configuration for the DUNE far detector. On the contrary, an increase of gain above 20 has in general a mild impact on the resolution for the electrons in the 1-10 GeV range. Improvements are mainly on the stochastic term, which favors mostly the low energy range below 1 GeV.
The simple study previously described is mostly intended to provide a general idea of the impact of the gain on the expected calorimetric performances of a dual-phase large LArTPC. The validation of the trend here highlighted may be carried out at ProtoDUNE-DP electron beam tests.

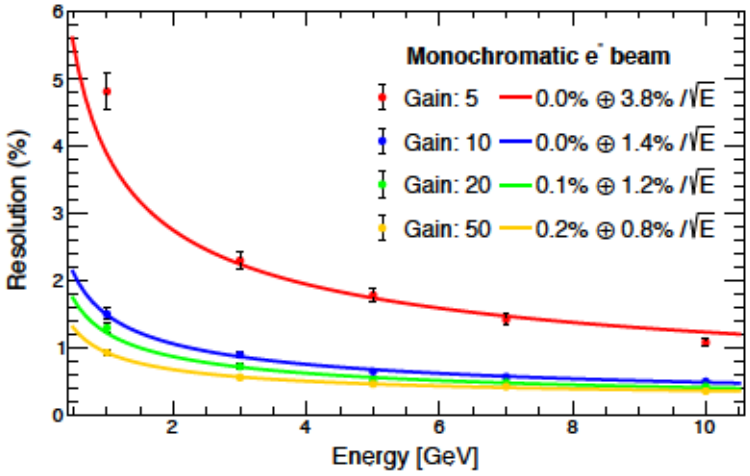


Fig. 5.19 Energy resolution of electrons simulated at fixed values of energy as function of the true energy. The samples are generated for different values of gain from 5 to 50.

5.4.4 Towards the oscillation analysis: smearing matrices

In the DUNE CDR [46] and several other works [142–144], detector effects are taken into account using smearing (or migration) matrices, M_{ij}, which represents the probability distribution functions for a neutrino interaction at the true energy in the i-th bin to be reconstructed with an energy in the j-th bin.
For an ideal detector, the smearing matrix would be the unit matrix. In a more realistic configuration, the form of the matrix depends on the intrinsic detector resolution and on the reconstruction procedure. The smearing matrices obtained for the ν_μCC, ν_eCC samples studied for this analysis are shown in Figure 5.20.

5.4.5 Summary on the energy resolution and comparison with single-phase

In this section, a method to estimate the energy of beam neutrinos has been proposed for the dual-phase design of the DUNE far detector. This study has been addressed using a full-simulation and including reconstruction effects for the first first time. The goal is to provide a realistic estimate of the energy resolution and bias in preparation for a dual-phase LBL oscillation analysis.
The results found for ν_μCC interactions are limited by the particular acceptance of the dual-phase *workspace* geometry used, in particular for the inelasticity of the neutrino interaction which is not representative of what expected using a full geometry. It is however reasonable to believe that the results obtained are more conservative that what is expected for the full-scale DUNE.
Our values for resolution and bias are compatible with the ones obtained for the single-phase DUNE far detector in a previous study [146]. The results are compared in Table 5.3. For

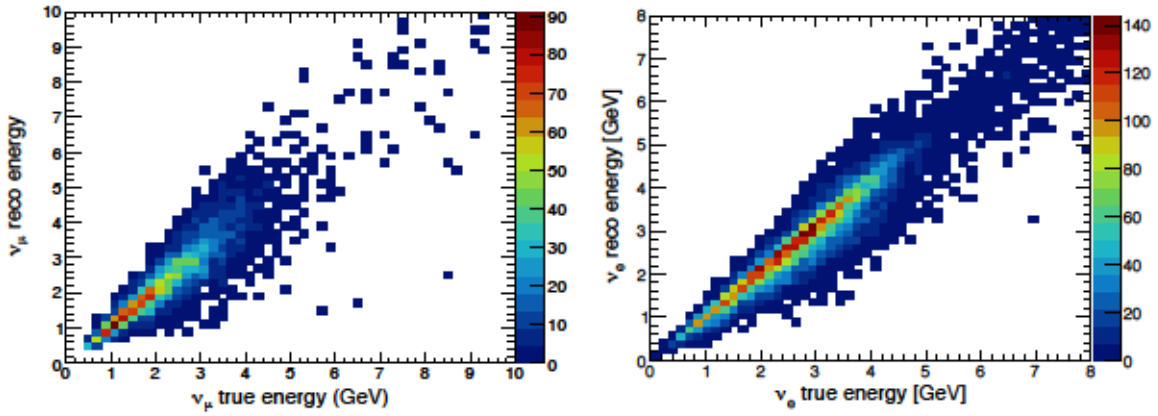


Fig. 5.20 The smearing matrices for the ν_μCC (left), ν_eCC (right) samples. Entries are not normalized.

Table 5.3 Summary of the reconstructed energy resolution and biases for the ν_μCC and ν_eCC samples for both single (SP) and dual-phase (DP).

		ν_μCC contained		ν_μCC not contained		ν_eCC	
		DP (%)	SP (%)	DP (%)	SP (%)	DP (%)	SP (%)
Leptons	Resolution	6	5	24	23	6	8
	Bias	0	1	0	0	0	0
Hadrons	Resolution	29	23	29	23	34	49
	Bias	1	4	1	4	0	5
All	Resolution	17	18	23	20	12	13
	Bias	0	1	0	4	0	0

ν_μCC, the single-phase detector seems to perform in general slightly better, due perhaps to the better reconstruction capabilities of the muon momentum. Dual-phase seems to bring a small contribution to the energy resolution for electromagnetic showers and hadrons, but more complete studies are needed in this sense, using a full reconstruction of showers in order to state and more realistic gain settings in order to state more realistically the impact of this design.

The current error on the bias is based on previous studies and chosen to be conservative. Both ProtoDUNE single and dual-phase are expected to provide more insight to constraint this uncertainty under the 2% level necessary.

5.5 Conclusions and outlook

In this chapter, the dual-phase LArSoft simulation has been used to study beam neutrinos at the DUNE dual-phase far detector for the first time. Also a fully automated reconstruction for hits and tracks has been tested for the first time on the neutrino sample.
The reconstruction has been found to be as efficient as and equivalent single-phase reconstruction, however most of the events are oriented in such a way that their two-dimensional projections are difficult to match correctly, reducing the quality of the reconstructed tracks. This affects mainly the tracks associated to the primary muons in ν_μCC events. A test with a newly rotated readout plane has shown removing part of the reconstruction troubles and improving the quality of muons tracks.
The reconstruction proposed has been applied to the energy reconstruction of ν_μCC and ν_eCC events at the DUNE far detector dual-phase, achieving values for the energy resolution and bias comparable with the CDR expectations and previous studies carried out in the single-phase detector.
Future improvements to the work described in this chapter should pursue the the development of dedicated algorithms able to perform more effectively the matching between the two collection views of the dual-phase. Machine learning in this sense may offer viable solution to squeeze out most information from the topology of the event, without the necessity to rely on ambiguous projections. Since energy reconstruction performances are comparable to what observed with the single-phase detector, a follow-up of this of the work should rather address the development of strategies for the neutrino flavor selection, in order to produce the spectra required for the LBL oscillation analysis.

Conclusions

The work presented in this thesis represents the first study towards the integration of the dual-phase design within DUNE and the evaluation of its potential and impact on the experiment. The first step has been the development of a dual-phase simulation and dedicated reconstruction within the LArSoft framework shared by other liquid Argon experiments and the single-phase design of DUNE.

- For simulation, the work has focused on introducing into the framework some dual-phase specificities, in particular related to the signal formation in dual-phase and the simulation of noise based on data-driven input. Thanks to the flexibility of the LArSoft framework to different geometries, the work done is easy to export to other dual-phase detectors.

- An almost entirely automated reconstruction chain able to reconstruct beam neutrinos is now in place. With this thesis, we contributed to interface specific algorithms of signal processing and hit finding designed for dual-phase, with higher level reconstruction algorithms of bidimensional clustering and tridimensional projection matching.

- An attempt to separate locally shower-like particle from track-like particle has been made using a Convolutional Neural Network able to classify each hit of the event based on the context in its immediate proximity. The classification accuracy for each class is larger than 92% with an inter-class *confusion* smaller than 8%. The robustness of the model has been tested against variations of the noise level, the charge scale, and the number of dead wires and the observed variations in the classification scores have been understood. The classification model trained with pions and protons has been introduced in the full reconstruction chain and tested on neutrinos event, with satisfactory performances overall.

The $3\times1\times1$ m^3 dual-phase prototype, exposed to cosmic rays in late 2017, was the first ton-scale DP-LArTPC. The important contribution given by this thesis on the evaluation of the performances of the detector comes from the study of the noise observed and the analysis of the detector gain. It is worth to remind that the analysis has been carried out using for the first time the dual-phase automated reconstruction tools developed for this thesis.

- The noise analysis has revealed a significant contribution of coherent noise brought inside the detector from the slow control cables, in particular the connectors to the temperature probes and the high-voltage board. Coherent patterns have been identified

along with the relevant frequencies and a coherent noise filter has been found effective in reducing the observed noise fluctuations to a value very close to the expectations from design.

- The analysis of the detector performances is highly affected by the sub-optimal high-voltage settings of the $3\times1\times1$ m^3 caused by a faulty connector on the extraction grid. Nevertheless it was possible to study the detector gain, the energy resolution, and the uniformity of the gain on the charge readout plane. Regretfully it has not been possible to study fluctuations of the gain over an extended period of time.

- A scan of the extraction and amplification of the detector has been performed showing results in quite good agreement with our predictions of the detector effects, however only qualitative trends can be compared and the $6\times6\times6$ m^3 prototype should address similar scan in the nominal high-voltage conditions.

The potential of the dual-phase design on the DUNE long baseline oscillation analysis has been studied for the first time. One important point of the test is to use a full simulation to properly account for all the dual-phase effects related to the signal formation on the two collection views. Neutrino events are reconstructed using the tools developed in this thesis. In this Chapter, the dual-phase LArSoft simulation has been used to study beam neutrinos at the DUNE dual-phase far detector for the first time. Also a fully automatic reconstruction for hits and tracks has been tested for the first time on the neutrino sample and applied to the reconstruction of the energy of neutrinos:

- The reconstruction efficiency for muons has been found to be 95%, similar to the one observed with a similar reconstruction in the single-phase simulation of DUNE. The quality of the reconstructed track is however poorer than in single phase due to the direction of the muons which tends to be parallel to one of the readout views. Reconstruction efficiency is lower for protons and pions compared to the case of single-phase.

- A method for the energy reconstruction of ν_μCC and ν_eCC events is proposed. The resolution achieved are 17% for ν_μCC events with a contained muon, 23% for ν_μCC events with an exiting muon, and 12% for ν_eCC events. The bias is always below 1%. The results obtained are comparable with the ones achieved for the single-phase with a similar approach.

A possible outlook for this work is mostly related to reconstruction. Although straight tracks can be reconstructed fairly well, more complex topologies requires further algorithms expressively developed for the dual-phase case. ProtoDUNE dual-phase is currently in operation at CERN and it will offer the opportunities to further understand the technology, improve the simulation and explore reconstruction strategies also with data.

References

[1] Wolfgang Pauli. *Letter to the Tübingen conference*. 1930.

[2] Enrico Fermi. "Tentativo di una Teoria Dei Raggi β". In: *Nuovo Cim.* 11.1 (1934), pp. 1–19. ISSN: 0029-6341. DOI: 10.1007/BF02959820. URL: http://link.springer.com/10.1007/BF02959820.

[3] F. Reines and C. L. Cowan. "Detection of the Free Neutrino". In: *Phys. Rev.* 92.3 (1953), pp. 830–831. ISSN: 0031-899X. DOI: 10.1103/PhysRev.92.830. URL: https://link.aps.org/doi/10.1103/PhysRev.92.830.

[4] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". In: *Science* 124.3212 (1956), pp. 103–104. ISSN: 0036-8075. DOI: 10.1126/science.124.3212.103. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.124.3212.103.

[5] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9.1 (1962), pp. 36–44. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[6] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". In: *Phys. Rev. Lett.* 35 (22 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[7] K. Kodama et al. "Observation of tau neutrino interactions". In: *Phys. Lett. B* 504.3 (2001), pp. 218–224. ISSN: 03702693. DOI: 10.1016/S0370-2693(01)00307-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S0370269301003070.

[8] "Precision electroweak measurements on the Z resonance". In: *Physics Reports* 427.5 (2006), pp. 257 –454. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2005.12.006. URL: http://www.sciencedirect.com/science/article/pii/S0370157305005119.

[9] Bruce T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". In: *Astrophys. J.* 496.1 (1998), pp. 505–526. ISSN: 0004-637X. DOI: 10.1086/305343. URL: http://stacks.iop.org/0004-637X/496/i=1/a=505.

[10] Y. Fukuda et al. "Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days". In: *Phys. Rev. Lett.* 81.6 (1998), pp. 1158–1162. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.81.1158. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1158.

[11] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[12] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[13] N. Agafonova et al. "Observation of tau neutrino appearance in the CNGS beam with the OPERA experiment". In: *Progress of Theoretical and Experimental Physics* 2014.10 (Oct. 2014). ISSN: 2050-3911. DOI: 10.1093/ptep/ptu132. eprint: http://oup.prod.sis.lan/ptep/article-pdf/2014/10/101C01/4414189/ptu132.pdf. URL: https://doi.org/10.1093/ptep/ptu132.

[14] B. Pontecorvo. "Mesonium and anti-mesonium". In: *Zh.Eksp.Teor.Fiz.* 33 (1957), p. 429. URL: http://inspirehep.net/record/2884?ln=en.

[15] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Prog. Theor. Phys.* 28.5 (1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. URL: https://academic.oup.com/ptp/article-lookup/doi/10.1143/PTP.28.870.

[16] T. K. Kuo and James Pantaleone. "Neutrino oscillations in matter". In: *Rev. Mod. Phys.* 61.4 (1989), pp. 937–979. ISSN: 00346861. DOI: 10.1103/RevModPhys.61.937. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[17] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98.3 (2018), p. 030001. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[18] F. Kaether et al. "Reanalysis of the Gallex solar neutrino flux and source experiments". In: *Phys. Lett. B* 685.1 (2010), pp. 47–54. ISSN: 0370-2693. DOI: 10.1016/J.PHYSLETB.2010.01.030. URL: https://www.sciencedirect.com/science/article/pii/S0370269310000729?via{\%}3Dihub.

[19] J. N. Abdurashitov et al. "Measurement of the solar neutrino capture rate with gallium metal. III. Results for the 2002–2007 data-taking period". In: *Phys. Rev. C* 80.1 (2009), p. 015807. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.80.015807. URL: https://link.aps.org/doi/10.1103/PhysRevC.80.015807.

[20] B. Aharmim et al. "Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory". In: *Phys. Rev. C* 88.2 (2013), p. 025501. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.88.025501. URL: https://link.aps.org/doi/10.1103/PhysRevC.88.025501.

[21] A. Gando et al. "Reactor on-off antineutrino measurement with KamLAND". In: *Phys. Rev. D* 88.3 (2013), p. 033001. ISSN: 1550-7998. DOI: 10.1103/PhysRevD.88.033001. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.033001.

[22] M. G. Aartsen et al. "Measurement of Atmospheric Neutrino Oscillations at 6–56 GeV with IceCube DeepCore". In: *Phys. Rev. Lett.* 120 (7 2018), p. 071801. DOI: 10.1103/PhysRevLett.120.071801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.071801.

[23] M. A. Acero et al. "New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment". In: *Phys. Rev. D* 98 (3 2018), p. 032012. DOI: 10.1103/PhysRevD.98.032012. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.032012.

[24] P. Adamson et al. "Measurement of Neutrino and Antineutrino Oscillations Using Beam and Atmospheric Data in MINOS". In: *Phys. Rev. Lett.* 110 (25 2013), p. 251801. DOI: 10.1103/PhysRevLett.110.251801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.251801.

[25] K. Abe et al. "Search for *CP* Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with 2.2×10^{21} Protons on Target". In: *Phys. Rev. Lett.* 121 (17 2018), p. 171802. DOI: 10.1103/PhysRevLett.121.171802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.171802.

[26] D. Adey et al. "Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay". In: *Phys. Rev. Lett.* 121 (24 2018), p. 241805. DOI: 10.1103/PhysRevLett.121.241805. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.241805.

[27] G. Bak et al. "Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO". In: *Phys. Rev. Lett.* 121 (20 2018), p. 201801. DOI: 10.1103/PhysRevLett.121.201801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.201801.

[28] Y. and Abe et al. "Improved measurements of the neutrino mixing angle $\theta 13$ with the Double Chooz detector". In: *Journal of High Energy Physics* 2014.10 (2014), p. 86. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2014)086. URL: https://doi.org/10.1007/JHEP10(2014)086.

[29] Ivan Esteban et al. "Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of $\theta_2 3$, $\delta_C P$, and the mass ordering". In: *JHEP* 01 (2019), p. 106. DOI: 10.1007/JHEP01(2019)106. arXiv: 1811.05487 [hep-ph].

[30] Francesco Capozzi, Eligio Lisi, and Antonio Marrone. "Probing the neutrino mass ordering with KM3NeT-ORCA: Analysis and perspectives". In: *J. Phys.* G45.2 (2018), p. 024003. DOI: 10.1088/1361-6471/aa9503. arXiv: 1708.03022 [hep-ph].

[31] Steven Wren. "Neutrino Mass Ordering Studies with PINGU and IceCube/DeepCore". In: *Proceedings, Prospects in Neutrino Physics (NuPhys2015): London, UK, December 16-18, 2015*. 2016. arXiv: 1604.08807 [physics.ins-det].

[32] Fengpeng An et al. "Neutrino Physics with JUNO". In: *J. Phys.* G43.3 (2016), p. 030401. DOI: 10.1088/0954-3899/43/3/030401. arXiv: 1507.05613 [physics.ins-det].

[33] James M. Cline. "Baryogenesis". In: *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*. 2006. arXiv: hep-ph/0609145 [hep-ph].

[34] Antonio Riotto. "Theories of baryogenesis". In: *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy, June 29-July 17, 1998*. 1998, pp. 326–436. arXiv: hep-ph/9807454 [hep-ph].

[35] S. T. Petcov. "Leptonic CP violation and leptogenesis". In: *Int. J. Mod. Phys.* A29 (2014). [,179(2014)], p. 1430028. DOI: 10.1142/S0217751X14300282,10.1142/9789814590112_0009. arXiv: 1405.2263 [hep-ph].

[36] Ubaldo Dore, Pier Loverre, and Lucio Ludovici. "History of accelerator neutrino beams". In: (2018). arXiv: 1805.01373 [physics.acc-ph].

[37] Hiroshi Nunokawa, Stephen J. Parke, and Jose W. F. Valle. "CP Violation and Neutrino Oscillations". In: *Prog. Part. Nucl. Phys.* 60 (2008), pp. 338–402. DOI: 10.1016/j.ppnp.2007.10.001. arXiv: 0710.0554 [hep-ph].

[38] William Marciano and Zohreh Parsa. "Intense neutrino beams and leptonic CP violation". In: *Nucl. Phys. Proc. Suppl.* 221 (2011), pp. 166–172. DOI: 10.1016/j.nuclphysbps.2011.03.114. arXiv: hep-ph/0610258 [hep-ph].

[39] S.K. Agarwalla et al. "The mass-hierarchy and CP-violation discovery reach of the LBNO long-baseline neutrino experiment". In: *J. High Energy Phys.* 2014.5 (2014), p. 94. ISSN: 1029-8479. DOI: 10.1007/JHEP05(2014)094. URL: http://link.springer.com/10.1007/JHEP05(2014)094.

[40] K. Abe et al. "Observation of Electron Neutrino Appearance in a Muon Neutrino Beam". In: *Phys. Rev. Lett.* 112.6 (2014), p. 061802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.112.061802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.061802.

[41] K. Abe et al. "Combined Analysis of Neutrino and Antineutrino Oscillations at T2K". In: *Phys. Rev. Lett.* 118.15 (2017), p. 151801. DOI: 10.1103/PhysRevLett.118.151801. arXiv: 1701.00432 [hep-ex].

[42] K. Abe et al. "Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —". In: (2011). arXiv: 1109.3262 [hep-ex].

[43] K. Abe et al. "Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —". In: (2011). arXiv: 1109.3262. URL: http://arxiv.org/abs/1109.3262.

[44] Mathieu Lamoureux. "Upgrade of the T2K near detector ND280: effect on oscillation and cross-section analyses". In: *PoS* NuFact2017 (2018), p. 056. DOI: 10.22323/1.295.0056. arXiv: 1803.02645 [physics.ins-det].

[45] K. Abe et al. "Physics potentials with the second Hyper-Kamiokande detector in Korea". In: *PTEP* 2018.6 (2018), p. 063C01. DOI: 10.1093/ptep/pty044. arXiv: 1611.06118 [hep-ex].

[46] R. Acciarri et al. "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)". In: (2016). arXiv: 1601.05471 [physics.ins-det].

[47] Milind V. Diwan. "The Case for a super neutrino beam". In: *Frascati Phys. Ser.* 35 (2004). [,89(2004)], pp. 89–109. arXiv: hep-ex/0407047 [hep-ex].

[48] M. Bass et al. "Baseline Optimization for the Measurement of CP Violation, Mass Hierarchy, and θ_{23} Octant in a Long-Baseline Neutrino Oscillation Experiment". In: *Phys. Rev.* D91.5 (2015), p. 052015. DOI: 10.1103/PhysRevD.91.052015. arXiv: 1311.0212 [hep-ex].

[49] Andre De Gouvea et al. "Physics with Beam Tau-Neutrino Appearance at DUNE". In: *Phys. Rev.* D100.1 (2019), p. 016004. DOI: 10.1103/PhysRevD.100.016004. arXiv: 1904.07265 [hep-ph].

[50] B. Abi et al. *The DUNE Far Detector Technical Design Report, Volume 2: Dune Physics (In preparation).* 2019.

[51] R. Acciarri et al. "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)". In: (2015). arXiv: 1512.06148 [physics.ins-det].

[52] James Strait et al. "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)". In: (2016). arXiv: 1601.05823 [physics.ins-det].

[53] R. Acciarri et al. "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)". In: (2016). arXiv: 1601.02984 [physics.ins-det].

[54] B. Abi et al. "The DUNE Far Detector Interim Design Report Volume 1: Physics, Technology and Strategies". In: (2018). arXiv: 1807.10334 [physics.ins-det].

[55] B. Abi et al. "The DUNE Far Detector Interim Design Report, Volume 2: Single-Phase Module". In: (2018). arXiv: 1807.10327 [physics.ins-det].

[56] B. Abi et al. "The DUNE Far Detector Interim Design Report, Volume 3: Dual-Phase Module". In: (2018). arXiv: 1807.10340 [physics.ins-det].

[57] C. Rubbia. *The Liquid Argon Time Projection Chamber: A New Concept for Neutrino Detectors*. 1977. URL: http://inspirehep.net/record/857394/?ln=en.

[58] D. R. Nygren. "The Time Projection Chamber: A New 4 pi Detector for Charged Particles". In: *eConf* C740805 (1974), p. 58.

[59] W.J. Willis and V. Radeka. "Liquid-argon ionization chambers as total-absorption detectors". In: *Nucl. Instruments Methods* 120.2 (1974), pp. 221–236. ISSN: 0029554X. DOI: 10.1016/0029-554X(74)90039-1. URL: https://linkinghub.elsevier.com/retrieve/pii/0029554X74900391.

[60] F. Arneodo et al. "The ICARUS 50-l LAr TPC in the CERN neutrino beam". In: *New detectors. Proceedings, 36th Workshop of the INFN Eloisatron Project, Erice, Italy, November 1-7, 1997*. 1998, pp. 3–12. arXiv: hep-ex/9812006 [hep-ex].

[61] S. Amerio et al. "Design, construction and tests of the ICARUS T600 detector". In: *Nucl. Instrum. Meth.* A527 (2004), pp. 329–410. DOI: 10.1016/j.nima.2004.02.044.

[62] C. Anderson et al. "The ArgoNeuT Detector in the NuMI Low-Energy beam line at Fermilab". In: *JINST* 7 (2012), P10019. DOI: 10.1088/1748-0221/7/10/P10019. arXiv: 1205.6747 [physics.ins-det].

[63] R. Acciarri et al. "Design and Construction of the MicroBooNE Detector". In: *JINST* 12.02 (2017), P02017. DOI: 10.1088/1748-0221/12/02/P02017. arXiv: 1612.05824 [physics.ins-det].

[64] B. Abi et al. "The Single-Phase ProtoDUNE Technical Design Report". In: (2017). arXiv: 1706.07081 [physics.ins-det].

[65] C. Adams et al. "Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation". In: *JINST* 13.07 (2018), P07006. DOI: 10.1088/1748-0221/13/07/P07006. arXiv: 1802.08709 [physics.ins-det].

[66] Andrea Scarpelli. "ProtoDUNE and a Dual-phase LArTPC". In: *PoS* NOW2018 (2019), p. 036. DOI: 10.22323/1.337.0036. arXiv: 1902.04780 [physics.ins-det].

[67] B. Baller. *LArTPC calculator*. URL: MicroBooNEdoc:3564.

[68] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994. ISBN: 978-3-540-57280-0. DOI: 10.1007/978-3-642-57920-2. URL: http://link.springer.com/10.1007/978-3-642-57920-2.

[69] L. Landau. "On the energy loss of fast particles by ionization". In: *J.Phys.(USSR)* 8 (1944), pp. 201–205. URL: http://inspirehep.net/record/45074/.

[70] P.V. Vavilov. "Ionization losses of high-energy heavy particles". In: *Zh.Eksp.Teor.Fiz.* 32 (1957), pp. 749–751. URL: http://inspirehep.net/record/48374/.

[71] M. Miyajima et al. "Average energy expended per ion pair in liquid argon". In: *Phys. Rev. A* 9.3 (1974), pp. 1438–1443. ISSN: 0556-2791. DOI: 10.1103/PhysRevA.9.1438. URL: https://link.aps.org/doi/10.1103/PhysRevA.9.1438.

[72] George Jaffé. "Zur Theorie der Ionisation in Kolonnen". In: *Ann. Phys.* 347.12 (1913), pp. 303–344. ISSN: 00033804. DOI: 10.1002/andp.19133471205. URL: http://doi.wiley.com/10.1002/andp.19133471205.

[73] J. Thomas and D. A. Imel. "Recombination of electron-ion pairs in liquid argon and liquid xenon". In: *Phys. Rev. A* 36.2 (1987), pp. 614–616. ISSN: 0556-2791. DOI: 10.1103/PhysRevA.36.614. URL: https://link.aps.org/doi/10.1103/PhysRevA.36.614.

[74] S. Amoruso et al. "Study of electron recombination in liquid argon with the ICARUS TPC". In: *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 523.3 (2004), pp. 275–286. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2003.11.423. URL: https://www.sciencedirect.com/science/article/pii/S0168900204000506?via{\%}3Dihub.

[75] J. B. Birks et al. *The Theory and Practice of Scintillation Counting: International Series of Monographs in Electronics and Instrumentation.* ISBN: 9781483156064.

[76] R. Acciarri et al. "A Study of Electron Recombination Using Highly Ionizing Particles in the ArgoNeuT Liquid Argon TPC". In: *JINST* 8 (2013), P08005. DOI: 10.1088/1748-0221/8/08/P08005. arXiv: 1306.1712 [physics.ins-det].

[77] Tadayoshi Doke, Kimiaki Masuda, and Eido Shibamura. "Estimation of absolute photon yields in liquid argon and xenon for relativistic (1 MeV) electrons". In: *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 291.3 (1990), pp. 617–620. ISSN: 01689002. DOI: 10.1016/0168-9002(90)90011-T. URL: https://linkinghub.elsevier.com/retrieve/pii/016890029090011T.

[78] S. Kubota et al. "Recombination luminescence in liquid argon and in liquid xenon". In: *Phys. Rev. B* 17.6 (1978), pp. 2762–2765. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.17.2762. URL: https://link.aps.org/doi/10.1103/PhysRevB.17.2762.

[79] D W Swan. "Drift velocity of electrons in liquid argon, and the influence of molecular impurities". In: *Proc. Phys. Soc.* 83.4 (1964), pp. 659–666. ISSN: 0370-1328. DOI: 10.1088/0370-1328/83/4/319. URL: http://stacks.iop.org/0370-1328/83/i=4/a=319?key=crossref.0c05fba46c0ddd7c69392e51697e8442.

[80] W. Walkowiak. "Drift velocity of free electrons in liquid argon". In: *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 449.1-2 (2000), pp. 288–294. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(99)01301-7. URL: https://www.sciencedirect.com/science/article/pii/S0168900299013017?via{\%}3Dihub.

[81] S. Amoruso et al. "Analysis of the liquid argon purity in the ICARUS T600 TPC". In: *Nucl. Instrum. Meth.* A516 (2004), pp. 68–79. DOI: 10.1016/j.nima.2003.07.043.

[82] E. Buckley et al. "A study of ionization electrons drifting over large distances in liquid argon". In: *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 275.2 (1989), pp. 364–372. ISSN: 0168-9002. DOI: 10.1016/0168-9002(89)90710-9. URL: https://www.sciencedirect.com/science/article/pii/0168900289907109.

[83] A. F. Borghesani et al. "Electron transmission through the Ar liquid-vapor interface". In: *Phys. Lett. A* 149.9 (1990), pp. 481–484. ISSN: 03759601. DOI: 10.1016/0375-9601(90)90221-9. URL: https://www.sciencedirect.com/science/article/abs/pii/0375960190902219.

[84] E Gushchin, A Kruglov, and I Obodovskil. "Emission of hot electrons from liquid and solid argon and xenon". In: *Sov. Phys. JETP* 82 (Jan. 1982), pp. 860–862.

[85] A. I. Bolozdynya. "Two-phase emission detectors and their applications". In: *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 422.1-3 (1999), pp. 314–320. ISSN: 01689002. DOI: 10.1016/S0168-9002(98)00965-6. URL: https://www.sciencedirect.com/science/article/pii/S0168900298009656.

[86] Takahiko Aoyama. "Generalized gas gain formula for proportional counters". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 234.1 (1985), pp. 125 –131. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(85)90817-4. URL: http://www.sciencedirect.com/science/article/pii/0168900285908174.

[87] Wu Shuoxing. "Study of alternative double phase LAr TPC charge readout systems". PhD thesis. 2012. DOI: 10.3929/ETHZ-B-000172413. URL: https://www.research-collection.ethz.ch/handle/20.500.11850/172413.

[88] W. H. Furry. "On Fluctuation Phenomena in the Passage of High Energy Electrons through Lead". In: *Phys. Rev.* 52.6 (1937), pp. 569–581. ISSN: 0031-899X. DOI: 10.1103/PhysRev.52.569. URL: https://link.aps.org/doi/10.1103/PhysRev.52.569.

[89] D. S. Akerib et al. "The Large Underground Xenon (LUX) Experiment". In: *Nucl. Instrum. Meth.* A704 (2013), pp. 111–126. DOI: 10.1016/j.nima.2012.11.135. arXiv: 1211.3788 [physics.ins-det].

[90] C. E. Aalseth et al. "DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS". In: *Eur. Phys. J. Plus* 133 (2018), p. 131. DOI: 10.1140/epjp/i2018-11973-4. arXiv: 1707.08145 [physics.ins-det].

[91] P Mitra. "Current to Conductors Induced by a Moving Point Charge". In: *American Journal of Physics* 38 (July 2019), pp. 112–112. DOI: 10.1119/1.1976205.

[92] C Rubbia et al. "Underground operation of the ICARUS T600 LAr-TPC: first results". In: *Journal of Instrumentation* 6.07 (2011), P07011–P07011. DOI: 10.1088/1748-0221/6/07/p07011. URL: https://doi.org/10.1088%2F1748-0221%2F6%2F07%2Fp07011.

[93] F. Arneodo et al. "Performance of a liquid argon time projection chamber exposed to the CERN West Area Neutrino Facility neutrino beam". In: *Phys. Rev. D* 74.11 (2006), p. 112001. ISSN: 1550-7998. DOI: 10.1103/PhysRevD.74.112001. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.112001.

[94] A. Ankowski et al. "Energy reconstruction of electromagnetic showers from pi0 decays with the ICARUS T600 Liquid Argon TPC". In: *Acta Phys. Polon.* B41 (2010), pp. 103–125. arXiv: 0812.2373 [hep-ex].

[95] A. Rubbia. *Experiments for CP violation: A Giant liquid argon scintillation, Cerenkov and charge imaging experiment?* 2004. URL: http://inspirehep.net/record/644226.

[96] D. Angus et al. "The LAGUNA design study-towards giant liquid based underground detectors for neutrino physics and astrophysics and proton decay searches". In: *European strategy for future neutrino physics. Proceedings, Workshop, Geneva, Switzerland, October 1-3, 2009*. 2010, pp. 226–229. arXiv: 1001.0077 [physics.ins-det].

[97] A. Breskin et al. "A concise review on THGEM detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 598.1 (2009). Instrumentation for Collding Beam Physics, pp. 107 –111. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2008.08.062. URL: http://www.sciencedirect.com/science/article/pii/S0168900208012047.

[98] Filippo Resnati. "Modeling, design and first operation of the novel double phase LAr LEM-TPC detector". PhD thesis. 2012. DOI: 10.3929/ETHZ-A-007582292. URL: http://inspirehep.net/record/1677473/.

[99] C. Cantini et al. "Performance study of the effective gain of the double phase liquid Argon LEM Time Projection Chamber". In: *J. Instrum.* 10.03 (2015), P03017–P03017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/03/P03017. URL: http://stacks.iop.org/1748-0221/10/i=03/a=P03017?key=crossref.95c258cc35849060b41ad266e597f0e8.

[100] C Cantini et al. "Long-term operation of a double phase LAr LEM Time Projection Chamber with a simplified anode and extraction-grid design". In: *Journal of Instrumentation* 9.03 (2014), P03017–P03017. DOI: 10.1088/1748-0221/9/03/p03017. URL: https://doi.org/10.1088%2F1748-0221%2F9%2F03%2Fp03017.

[101] I. De Bonis et al. "LBNO-DEMO: Large-scale neutrino detector demonstrators for phased performance assessment in view of a long-baseline oscillation experiment". In: (2014). arXiv: 1409.4405 [physics.ins-det].

[102] B. Aimard et al. "A 4 tonne demonstrator for large-scale dual-phase liquid argon time projection chambers". In: *JINST* 13.11 (2018), P11003. DOI: 10.1088/1748-0221/13/11/P11003. arXiv: 1806.03317 [physics.ins-det].

[103] *MicroTCA standard*. URL: https://www.picmg.org/openstandards/microtca/.

[104] *White Rabbit Project*. URL: https://white-rabbit.web.cern.ch/.

[105] A. Badertscher et al. "Construction and operation of a Double Phase LAr Large Electron Multiplier Time Projection Chamber". In: (2008). arXiv: 0811.3384. URL: http://arxiv.org/abs/0811.3384.

[106] E. L. Snider and G. Petrillo. "LArSoft: Toolkit for Simulation, Reconstruction and Analysis of Liquid Argon TPC Neutrino Detectors". In: *J. Phys. Conf. Ser.* 898.4 (2017), p. 042057. DOI: 10.1088/1742-6596/898/4/042057.

[107] C. Green et al. "The Art Framework". In: *J. Phys. Conf. Ser.* 396 (2012), p. 022020. DOI: 10.1088/1742-6596/396/2/022020.

[108] Costas Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual". In: (2015). arXiv: 1510.05494 [hep-ph].

[109] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[110] Ralph Engel et al. "Towards a Next Generation of CORSIKA: A Framework for the Simulation of Particle Cascades in Astroparticle Physics". In: *Comput. Softw. Big Sci.* 3.1 (2019), p. 2. DOI: 10.1007/s41781-018-0013-0. arXiv: 1808.08226 [astro-ph.IM].

[111] "Improvements in the Geant4 Hadronic Physics". In: *Journal of Physics: Conference Series* 331.3 (2011), p. 032002. DOI: 10.1088/1742-6596/331/3/032002. URL: https://doi.org/10.1088%2F1742-6596%2F331%2F3%2F032002.

[112] B. Aimard et al. "Performance study of a $3\times1\times1$ m^3 dual phase liquid Argon TPC exposed to cosmic rays (In Preparation)". In: ().

[113] B. Baller. *TrajCluster users guide*. 2018. URL: https://larsoft.org/single-record/?pdb=98.

[114] The MicroBooNE Collaboration. *Reconstruction Performance Studies with MicroBooNEData in Support of Summer 2018 Analyses*. Tech. rep. 2018. URL: MICROBOONE-NOTE-1049-PUB.

[115] Irene Nutini. "The LArIAT Experiment". In: *J. Phys. Conf. Ser.* 888 (2017), p. 012168. DOI: 10.1088/1742-6596/888/1/012168.

[116] R. Sulej and D. Stefan. *The Projection Matching Algorithm*. 2016. URL: https://larsoft.org/single-record/?pdb=102.

[117] M. Antonello et al. "Precise 3D track reconstruction algorithm for the ICARUS T600 liquid argon time projection chamber detector". In: *Adv. High Energy Phys.* 2013 (2013), p. 260820. DOI: 10.1155/2013/260820. arXiv: 1210.5089 [physics.ins-det].

[118] J. S. Marshall and M. A. Thomson. "The Pandora Software Development Kit for Pattern Recognition". In: *Eur. Phys. J.* C75.9 (2015), p. 439. DOI: 10.1140/epjc/s10052-015-3659-3. arXiv: 1506.05348 [physics.data-an].

[119] R. Acciarri et al. "The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector". In: *Eur. Phys. J.* C78.1 (2018), p. 82. DOI: 10.1140/epjc/s10052-017-5481-6. arXiv: 1708.03135 [hep-ex].

[120] R. Acciarri et al. "Michel Electron Reconstruction Using Cosmic-Ray Data from the MicroBooNE LArTPC". In: *JINST* 12.09 (2017), P09014. DOI: 10.1088/1748-0221/12/09/P09014. arXiv: 1704.02927 [physics.ins-det].

[121] Yann LeCun, Y Bengio, and Geoffrey Hinton. "Deep Learning". In: *Nature* 521 (May 2015), pp. 436–44. DOI: 10.1038/nature14539.

[122] F. Chollet. *Deep Learning with Python*. 1st ed. Greenwich, CT, USA: Manning Publications Co., 2017. ISBN: 1617294438 9781617294433.

[123] Alexander Radovic et al. "Machine learning at the energy and intensity frontiers of particle physics". In: *Nature* 560.7716 (2018), pp. 41–48. DOI: 10.1038/s41586-018-0361-2.

[124] A. Aurisano et al. "A Convolutional Neural Network Neutrino Event Classifier". In: *JINST* 11.09 (2016), P09001. DOI: 10.1088/1748-0221/11/09/P09001. arXiv: 1604.01444 [hep-ex].

[125] R. Acciarri et al. "Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber". In: *JINST* 12.03 (2017), P03011. DOI: 10.1088/1748-0221/12/03/P03011. arXiv: 1611.05531 [physics.ins-det].

[126] R. Reed and Marks R. *Neural Smithing: Supervised Learning in Feedforward Artificial NeuralNetworks*. A Bradford Book. MIT Press, 1999.

[127] Mathworks. *Convolutional Neural Network, 3 things you need to know*. URL: https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html.

[128] Vinod Nair and Geoffrey E. Hinton. "Rectified Linear Units Improve Restricted Boltzmann Machines". In: *ICML*. 2010.

[129] C. M. Bishop. *Pattern recognition and Machine Learning*. Springer-Verlag New York, Inc., 2006.

[130] R. Sulej. *Image Pattern algorithms*. 2017. URL: https://larsoft.org/single-record/?pdb=138.

[131] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[132] Song Mei, Andrea Montanari, and Phan-Minh Nguyen. "A mean field view of the landscape of two-layer neural networks". In: *Proceedings of the National Academy of Sciences* 115.33 (2018), E7665–E7671. ISSN: 0027-8424. DOI: 10.1073/pnas.1806579115. eprint: https://www.pnas.org/content/115/33/E7665.full.pdf. URL: https://www.pnas.org/content/115/33/E7665.

[133] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[134] François Chollet et al. *Keras*. https://keras.io. 2015.

[135] R. Acciarri et al. "Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC". In: *JINST* 12.08 (2017), P08003. DOI: 10.1088/1748-0221/12/08/P08003. arXiv: 1705.07341 [physics.ins-det].

[136] Caspar Schlösser. "Commissioning and Charge Readout Calibration of a 5 Ton Dual Phase Liquid Argon TPC". MA thesis. 2017. DOI: 10.3929/ethz-b-000167186. URL: https://www.research-collection.ethz.ch/handle/20.500.11850/167186.

[137] R. Veenhof. "Garfield, a drift chamber simulation program". In: *Conf. Proc.* C9306149 (1993). [,66(1993)], pp. 66–71.

[138] *ANSYS*. URL: https://econengineering.com/software-solution/ansys/.

[139] B. Aimard et al. "Study of scintillation light collection, production and propagation in a 4 tonne dual phase LArTPC (In Preparation)". In: ().

[140] Devis Lussi. "Study of the response of the novel LAr LEM-TPC detector exposed to cosmic rays and a charged particle beam". PhD thesis. 2013. DOI: 10.3929/ETHZ-A-010008070. URL: http://inspirehep.net/record/1677412?ln=en.

[141] The DUNE collaboration. *Far Detector Task Force Final Report*. Tech. rep. 2017. URL: https://docs.dunescience.org/cgi-bin/private/ShowDocument?docid=3384.

[142] Artur M. Ankowski et al. "Comparison of the calorimetric and kinematic methods of neutrino energy reconstruction in disappearance experiments". In: *Phys. Rev.* D92.7 (2015), p. 073014. DOI: 10.1103/PhysRevD.92.073014. arXiv: 1507.08560 [hep-ph].

[143] Artur M. Ankowski et al. "Missing energy and the measurement of the CP-violating phase in neutrino oscillations". In: *Phys. Rev.* D92.9 (2015), p. 091301. DOI: 10.1103/PhysRevD.92.091301. arXiv: 1507.08561 [hep-ph].

[144] Alexander Friedland and Shirley Weishi Li. "Understanding the energy resolution of liquid argon neutrino detectors". In: *Phys. Rev.* D99.3 (2019), p. 036009. DOI: 10.1103/PhysRevD.99.036009. arXiv: 1811.06159 [hep-ph].

[145] P. Abratenko et al. "Determination of muon momentum in the MicroBooNE LArTPC using an improved model of multiple Coulomb scattering". In: *JINST* 12.10 (2017), P10010. DOI: 10.1088/1748-0221/12/10/P10010. arXiv: 1703.06187 [physics.ins-det].

[146] N Grant. *Neutrino energy reconstruction in the DUNE far detector*. Tech. rep. 2019. URL: https://docs.dunescience.org/cgi-bin/private/ShowDocument?docid=2278.