Background Studies for Electron Antineutrino Oscillations Measurement at the Double Chooz Experiment

Adrien Hourlier

To cite this version:

Adrien Hourlier. Background Studies for Electron Antineutrino Oscillations Measurement at the Double Chooz Experiment. High Energy Physics - Experiment [hep-ex]. Paris Diderot, 2016. English. NNT : . tel-01441945

HAL Id: tel-01441945
https://theses.hal.science/tel-01441945
Submitted on 20 Jan 2017

UNIVERSITÉ SORBONNE PARIS CITÉ

Thèse préparée
à l' UNIVERSITÉ PARIS DIDEROT
ÉCOLE DOCTORALE STEP'UP — ED N°560
Laboratoire Astro-Particules et Cosmologie

Background Studies for $\bar{\nu}_e$ Oscillations Measurement at the Double Chooz Experiment

---

par
Adrien Hourlier
présentée et soutenue publiquement le
9 Septembre 2016
Thèse de doctorat de Physique de l'Univers dirigée
par Dr. Jaime Dawson

Devant un jury composé de:

Jaime Dawson Chargée de recherche (CNRS — APC)	Directrice de thèse
Thomas Patzak Professeur (Université Paris Diderot)	Président du jury
Ines Gil-Botella Professeur (CIEMAT)	Rapporteur
Marco Zito Ingénieur Chercheur (IRFU — CEA)	Rapporteur
Dominique Duchesneau Directeur de Recherche (CNRS — LAPP)	Examinateur
Michael Wurm Professeur (Johannes Gutenberg Universität)	Examinateur
Anatael Cabrera Chargé de recherche (CNRS — APC)	Examinateur

À Aurélie,
Merci d'illuminer ma vie

Acknowledgements

Writing these few words was much more difficult than I anticipated. It is never easy to thank someone, words, sadly, cannot cover the full extent of my gratitude. For too many of the people later mentioned, these simple acknowledgements sound too much like a farewell. I am also aware that I will probably forget people and I would like to apologize in advance.

I would first like to thank my advisor, Jaime. I hope writing will be easier than speaking. Thank you for caring for me all these years. It was a delight to work with you during these past three years, and more if we include the internships. I really enjoyed your enthusiasm, our discussions and your way of thinking outside of the box, finding issues no one else could identify.

I cannot possibly go any further without a thought for Anatael. We spent so many hours, and very nice ramen diners, discussing waveform reconstruction and stopping muon. You always pushed me to better myself and improve the rigor and robustness of my analyses, keeping your smile even in stressful times. I owe a big part of the quality of this thesis to you.

I am grateful to the members of my defense jury. To Thomas Patzak for agreeing to be the president of the jury, and for all your advice on writing and defending my work. To Inés Gil Botella and Marco Zito for their careful reading of this manuscript as my reviewers. To Dominique Duchesneau and Michael Wurm, for their interest they took in my work and for being part of this event.

I wish to thank the full Double Chooz collaboration for a very stimulating work environment, I especially appreciated the reactivity of people always eager to help and provide advice. I found in this collaboration more than colleagues, I found friends. Thank you all.

Thank you Kuze-Sensei, Masaki and the whole Kuze-lab team, thank you for inviting me and welcoming me for a month at TokyoTech, I had a great and productive time working there.

I would like to thank Anselmo for all the work we did on correlated backgrounds, it was a pleasure to work and discuss with you. Thank you also to Ralitsa, and Guillaume, for our hours of discussions over Skype, helping release the stress of deadlines. I am going to miss seeing and drinking with you at collaboration meetings.

Thanks also to Kazu, for being stuck with me during my internship, thank you for your patience and for all your coding tips. I can't wait to see you again at Fermilab.

Many thanks to the Double Chooz group at APC, the people I worked with daily during three years. Hervé, your involvement onsite is invaluable for the whole collaboration, countless trips to Chooz were avoided because of your dedication, and your advice on how to keep track of my work proved very useful to write this manuscript. Didier, you always were of great advice, especially when I was struggling with the servers at Chooz. Many lines of my scripts bare your mark. Davide, you were my advisor, along with Anatael for my first internship on Double Chooz, thank you for introducing me to this experiment, this collaboration, and this laboratory. Michel, you were always very reactive to solve any issue with IN2P3, and my many password reinitializations. Alessandra, your support and your advice on my thesis in general were always very useful. It is true that we were inexperienced on what was expected for a French thesis, and your opinion on the matter was invaluable.

I would like to express a special thought for Anthony and Héctor, with whom I spent too much time drinking coffee and discussing every possible topic, Double Chooz related or not. Thank you also for your help in writing this thesis, your many and very quick comments allowed me to hand-in more thank a draft.

I would also like to thank my office-mates Romain, Paolo and Thibaut for your constant good mood and the pleasure I had to see you everyday.

Thank you Quentin for your advice on the analysis of DCTPC, I have not had time to implement them but I keep them in mind for later use. Thank you Joao for our discussions on atmospheric neutrinos and on the history of oscillations.

I cannot forget the DCTPC team, thank you Josh for all the good times spent on this beautiful detector. It was not always easy and we are still only in the process of understanding it, but it sure was interesting and always gratifying to slowly work towards understanding a new detector. Thank you also Marjon and Allie for your work, we interacted less but it was always a pleasure, and I am looking forward to continue working with all of you.

And I am especially grateful to Janet. You introduced me to neutrino physics during my internship on DAEδALUS and then to the APC for my first contact with Double Chooz, I would probably not have done this thesis without you. We kept in touch all these years through Double Chooz and DCTPC and now you offer me the opportunity to work with you again, as your post-doc on MicroBooNE. Thank you for your continued interest and for your kindness.

Ce travail n'aurait pas été possible sans un dose journalière importante de caféine, ni sans une dose plus occasionnelle et raisonable de bière et de Méléze de la Distillerie des Boussardes, ce qui m'amène à remercier Benjamin et Caroline pour avoir fourni le-dit Méléze qui rend balèze et pour leur précieuse amitié. Un énorme merci aussi à Alban et Philippe pour cette si longue amitiée, et pour avoir fait le déplacement de si loin pour ma soutenance.

Un grand merci également à mes Parents, pour m'avoir toujours soutenu et accompagné, c'est grâce à vous et à votre amour que je suis devenu ce que je suis aujourd'hui.

Enfin, je voudrais remercier celle qui rend ma vie si belle depuis tant d'années, Aurélie. Merci pour tout, et pour ce magnifique 27 juin 2015 qui restera à jamais gravé dans ma mémoire.

Contents

Introduction

I have done a terrible thing, I have postulated a particle that cannot be detected
—Wolfgang Pauli, cited by Frederick Reines in his
foreword to *Spaceship Neutrino* by Christine Sutton

Neutrinos were introduced in 1930 by Wolfgang Pauli to address the question of the apparent non-conservation of the energy observed in β-decays [1]. They were experimentally observed only 60 years ago, in 1956 by Frederick Reines and Clyde Cowan, who detected electron anti-neutrinos emitted by the Savannah River nuclear reactor [2]. Neutrinos have since then driven the understanding of our universe, be it by introducing lepton number conservation, furthering our knowledge of the Standard Model of particle physics, or, at the opposite size scale, allowing us to understand cosmic events such as supernova explosions.

The 2015 Nobel Physics prize was awarded to T.Kajita (Super Kamiokande) and A.B. McDonald (Sudbury Neutrino Observatory) for "the discovery of neutrino oscillations, which shows that neutrinos have mass". By implying that neutrinos have a non-zero mass, neutrino oscillations are an indication of a physics beyond the Standard Model.

Neutrino oscillations are parametrized by two relative mass differences, distinguishing the three mass states of a neutrino, and three mixing angles, describing the repartition of each neutrino flavor on those mass states. An additional parameter δ_{CP} characterizing the CP-violation phase of neutrino oscillation was recently announced to be non-zero at 2σ by the T2K experiment at the NEUTRINO2016 conference this summer [3].

Oscillation signature in the solar neutrino flux observed by the SNO experiment [4] and the one of the atmospheric neutrino flux observed by the Super Kamiokande [5] experiment are driven respectively by the θ_{12} and θ_{23} mixing angles, the Double Chooz experiment was designed to perform a measurement of the last mixing angle θ_{13} via the measurement of the oscillation of the neutrino flux emitted by the two nuclear reactors of the Chooz power plant in the French Ardennes region [6]. The oscillation measurement is performed by two identical detectors at different distances from the reactor cores. The Near Detector provides a reference of the flux and energy spectrum of the neutrinos, and the comparison between that reference and the flux and spectrum observed in the Far detector allows to measure the θ_{13} parameter.

A first indication of a non-zero θ_{13} had been found in 2011 [7], before I started working on the experiment. However, this measurement was performed in a first phase of the experiment, with only one of the two planned detectors in operation. I was lucky enough to be part of the first analysis of Double Chooz with the full detector configuration. In shallow neutrino oscillation experiments, background constraint is one of the major effort of the analysis. My main study was the development of background reduction techniques for Stopping Muons decaying in the detectors. This background is one of the most dangerous for the experiment, particularly to the Near Detector, located at shallower depth.

This thesis is articulated around six chapters, the first chapter describes the path of neutrino discovery, from its first postulation to its experimental evidence, and on to the understanding of its properties, such as its interactions, its oscillating behaviour, and its potential as a messenger in various domains such as monitoring nuclear reactors, probing the composition of the deep earth, or even as a messenger for astrophysics.

The second chapter describes the Double Chooz experiment. Here are presented the experimental concept of the measurement of oscillations with multiple detectors proposed by Double Chooz, the detection process for the $\bar{\nu}_e$, the design of the detectors and the acquisition chain. The energy and vertex reconstruction applied to the events are also presented.

The third chapter describes the tests I performed during the commissioning of the Near Detector to evaluate the linearity of the Flash-ADC used to digitize the analog signals from the PMTs. The understanding of the digitization process and of the PMT pulse shape also allowed us to devise a new method of calorimetry to evaluate the energy deposited in the detector with a better linearity at low energy. This new reconstruction, called RecoZoR, is also presented in this chapter.

The fourth chapter describes the analysis performed by the collaboration to measure θ_{13}. Since the beginning of Double Chooz data taking, several analyses were successively performed, increasing both the statistics and systematic uncertainty of the θ_{13} measurement. During the course of my PhD, I took part in several of these analyses, here I will present the last single-detector results from Double Chooz, and the first two-detector analysis of the experiment.

The fifth chapter describes a side experiment called Double Chooz Time Projection Chamber (DCTPC), a gaseous neutron TPC that was installed in order to provide information on the direction, rate and spectral shape of the Fast Neutron flux in both of the experimental halls of Double Chooz. I was involved in all the stages of this experiment, from the installation, the running of the TPC, the energy reconstruction and the data analysis.

Finally, the sixth chapter describes the methods I developed in order to tag and reject Stopping Muons that decay in the Double Chooz detectors. Here are described as well the studies I performed to evaluate the performance of these methods in terms of rejection capabilities and efficiency on the neutrino candidates in the rest of the Double Chooz selection.

Chapter 1

The Neutrino, an Evasive Particle

The path of neutrino understanding is full of theoretical puzzles and technological challenges. Not only do they evade detection, they also evade comprehension.

They are neutral leptons, described as massless by the Standard Model ($\mathcal{SM}$). As such, they are only sensitive to weak interactions, making their interaction with other particles very unlikely. Neutrino experiments over the past two decades have evidenced a behaviour called *neutrino oscillations* that proves that neutrino are in fact massive, which is not predicted by the $\mathcal{SM}$.

In this chapter, we will discuss the discovery of the neutrino and the long path to understanding this particle, both theoretically and experimentally, with an emphasis on the oscillation phenomenon.

1.1 Neutrino Discovery

The discovery of the neutrino is closely linked to our growing knowledge of nuclear science. In the beginning of the XXth century, β-decay was understood as a parent nucleus X decaying into a daughter nucleus Y and the emission of a β-particle : $X = Y + e^{\mp}$. The sign of the electron determines the name of the decay being β^+ or β^-. This understanding required the electron to have a fixed energy. However, in 1914, a continuous electron spectrum was observed by James Chadwick [8].

This result seemed in contradiction with the principle of energy conservation, such that Niels Bohr suggested that the energy could be conserved only statistically [9]. 16 years later, in 1930, Wolfgang Pauli proposed to have a new particle he called *a neutron*, that would restore the energy conservation. This particle would be neutral to ensure charge conservation, of spin 1/2 to conserve angular momentum, would be much lighter than a proton ($< 0.01\mathrm{m_p}$), and with a penetrating power "perhaps ten times larger [...] than a gamma-ray" [1].
After Chadwick's discovery in 1932 of a neutral particle of mass comparable to the one of a proton that he calls *neutron* [10], Enrico Fermi renamed Pauli's particle *neutrino* when he formulated the three-body β-decay (eq. 1.1,1.2) in 1934 [11]. Pontecorvo extended Fermi's theory to muon decays in 1947 [12], prefiguring the lepton families. β-decays are now understood

as three body decays, the spectrum of electrons is thus continuous and the missing energy is carried away by neutrinos.

$$
\begin{aligned}
{}^{A}_{Z}X &= {}_{\mathrm{Z-1}}^{\ \ \mathrm{A}}\mathrm{Y} + e^{+} + \nu_e && \beta^{+}\text{decay} \qquad (1.1)\\
{}^{A}_{Z}X &= {}_{\mathrm{Z+1}}^{\ \ \mathrm{A}}\mathrm{Y} + e^{-} + \overline{\nu}_e && \beta^{-}\text{decay} \qquad (1.2)
\end{aligned}
$$

Since neutrinos are neutral leptons, they are not submitted to strong or electromagnetic interactions, but only to weak interactions. As such, the detection of neutrinos was long thought impossible. However, Frederick Reines and Clyde Cowan used inverse β-decay (IBD, eq. 1.3) from anti-neutrinos emitted by the Savannah River nuclear reactor to achieve the first detection of neutrino in 1956 [13, 2].

$$
p + \overline{\nu}_e \rightarrow n + e^{+} \tag{1.3}
$$

Their detector was segmented in layers of Cadmium chlorine dissolved in water, and liquid scintillator equipped with photo multiplier tubes (PMT). The coincidence between the positron scintillation signal and the scintillation of gammas from the neutron capture on cadmium was used to tag a neutrino event. By comparing data with the nuclear reactor on and off, as shown in Figure 1.1, they were able to show the detection of the neutrino and to measure an interaction cross section $\sigma_{\mathrm{exp}} = (1.2^{+0.7}_{+0.4}) \times 10^{-43}\mathrm{cm}^2$ compatible with the theorized value of $\sigma_{\mathrm{th}} = (1.0 \pm 0.17) \times 10^{-43}\mathrm{cm}^2$ [14].

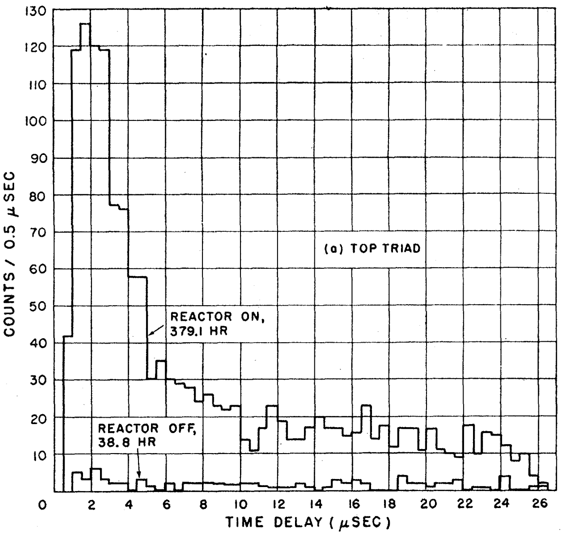


Figure 1.1: Time delay between the positron and delayed neutron capture on Cadmium for data acquired with the reactor on and off [14].

An experiment conducted by Davis and Harmer in 1959 at the Savannah River nuclear reactor as well, aiming to observe the reaction $\bar{\nu}_e{}^{37}_{17}\mathrm{Cl} \rightarrow {}^{37}_{18}\mathrm{Ar}e^{-}$ gave negative results. This can be explained by the introduction in 1953 by Konopinski and Mahmoud of the Lepton number L. The lepton number is $+1$ for an electron and neutrino, and -1 for a positron and anti-neutrino.

The lepton number is conserved throughout interaction, therefore, a $\bar{\nu}$ can only produce a e^+ when interacting on a nucleus, and therefore, by charge conservation, can only change a proton into a neutron.

In the year following the discovery of the neutrino, Lee and Yang showed that the weak interaction could violate the conservation of parity [15]. This behaviour was observed a few months later, in 1957, by Wu *et al* by observing an asymmetry in the β emission of the decay of polarized ^{60}Co[16] as shown in Figure 1.2. The asymmetry was measured by comparing the number of electrons emitted upward for a given temperature of the sample to the number of electrons emitted upward for the warm sample. One can see that the electrons are emitted preferentially opposite to the direction of the magnetic field. In fact the Wu experiment showed not only that the conservation of parity P is violated, but that invariance under charge symmetry C is violated as well. These results were confirmed by Goldhaber a year later [17].

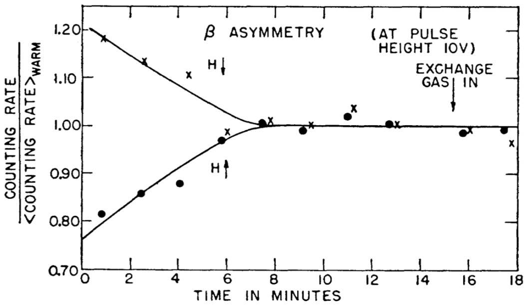


Figure 1.2: Wu observed the direction of emission of electrons in the decay of cooled down ^{60}Co under two opposite orientations of magnetic field. When the sample is warm (in the end of the experiment) electrons are emitted isotropically, however, at the beginning of the experiment, the electrons are preferentially emitted in the direction opposite to the magnetic field. This result was the first evidence that the P symmetry was violated.

In 1962, Danby *et al.* exposed a spark chamber to the first neutrino beam at the Brookhaven National Laboratory. The neutrino beam was created from 15 GeV protons impinging on a Beryllium target creating, among other products, pions and kaons that in turn decay into muons and neutrinos:

$$\pi^{\pm} \rightarrow \mu^{\pm}\nu(\bar{\nu}) \tag{1.4}$$

$$K^{\pm} \rightarrow \mu^{\pm}\nu(\bar{\nu}) \tag{1.5}$$

The charged leptons were stopped in a 13.5 m iron shielding, and only neutrinos were left to interact in the spark chamber. Neutrinos were found to specifically create muons through charged current interaction, which led to the idea of a different family of neutrino than those detected by Cowan and Reines [18].

The fact that the two lepton families are distinct, with a ν_e always producing a e, and a ν_μ always producing a μ lead to the conclusion that muons and electrons should have separate lepton numbers L_e and L_μ.

When in 1975 a third family of lepton, the τ-lepton was discovered, a third neutrino was thought to exist, the ν_τ which would couple to the τ with a new lepton number L_τ as the ν_μ couples to the μ and the ν_e couples to the e. However, because of the very short life-time of the τ, the experimental evidence of the existence of the ν_τ was very challenging. A definitive confirmation that a third neutrino family should exist was given by the LEP experiments ALEPH, DELPHI, L3 and OPAL at CERN in which the decays of bosons $W^\pm$ and Z^0 were studied in detail between 1990 and 2000. The study of the decay of the Z^0 boson allowed to fix at three the number of active neutrino families [19]. The decay width of the Z^0, as the sum of the widths of all independent decay processes, depends on the number of neutrino flavors. Thus the decay width of the Z^0 can be written $\Gamma_Z = \Gamma_c + \Gamma_{inv}$ where Γ_c is the contribution of all the processes that include charged fermions, and Γ_{inv} the contribution of all invisible processes, including the $Z^0 \to \nu_x \overline{\nu}_x$ decay.

$$\frac{\Gamma_{inv}}{\Gamma_\nu^{\mathcal{SM}}} = 2.9840 \pm 0.0082 \tag{1.6}$$

The invisible contribution is found to be three times bigger than the contribution predicted by the $\mathcal{SM}$ for a single neutrino flavor ($\Gamma_\nu^{\mathcal{SM}}$). The LEP results only constrain the number of neutrino flavors to three under the mass of $m_Z/2 \approx 45.5$ GeV/c^2. There could still be a flavor above that energy.

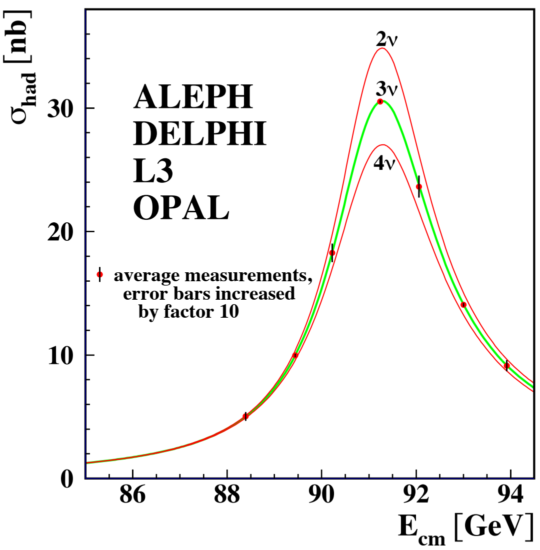


Figure 1.3: Measurements of the hadron production cross-section around the Z resonance by a joint analysis from LEP experiments ALEPH, DELPHI, L3 and OPAL. The curves indicate the predicted cross-section for two, three and four neutrino species with SM coupling and negligible mass.

The third family of neutrino was finally discovered in 2001 by the DONUT experiment at Fermilab [20]. The DONUT experiment looks for ν_τ interaction in emulsion targets interleaved with scintillator and steel absorbers. ν_τ interaction is identified as the only lepton created in a vertex originating inside the detector and decaying within 2 mm of its creation point. Figure 1.4 shows one of the four events identified as ν_τ. In this event, the two tracks on the side are interpreted as hadronic showers, and the central track shows a kink after a flight of 4535µm, the daughter particle is identified as an electron. This event is compatible with a $\tau^- \rightarrow e^- \, \bar{\nu}_\tau \, \nu_e$ or the CP conjugate.

Each charged lepton flavor is then linked to the corresponding neutrino through the conservation of the lepton number.

$$L_\alpha = \begin{array}{c} 1 \, \{\alpha, \nu_\alpha\} \\ -1 \, \{\bar{\alpha}, \bar{\nu}_\alpha\} \\ 0 \text{ for the rest} \end{array} \tag{1.7}$$

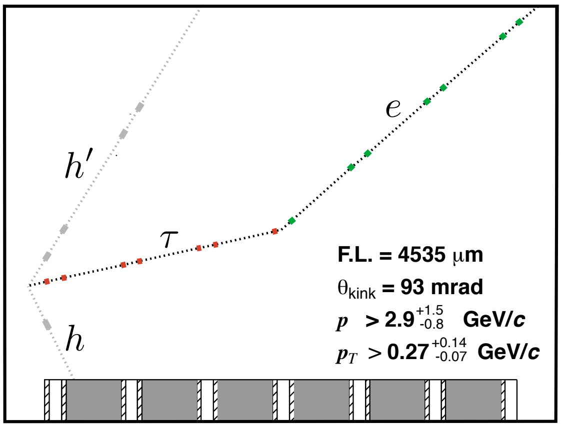


Figure 1.4: One of the four ν_τ events identified by the DONUT experiment.

1.2 The Neutrino in the Standard Model

The Standard Model of particle physics ($\mathcal{SM}$) is an evolving model that describes the three fundamental interactions (strong, electro-magnetic and weak interactions) and their effects on elementary particles. The gravity is not described by the $\mathcal{SM}$ for now. Elementary particles are grouped in two categories: Fermions, with half integer spin, and Bosons, with integer spin. Vectors bosons are carriers of the strong interaction (eight gluons), the electro-magnetic interaction (photon) and the weak interaction (W and Z bosons). An additional scalar boson, the Higgs boson, is needed to provide masses to the particles.

Fermions are furthermore separated in two categories: the quarks, which are sensitive to all four interactions, and the leptons which are not sensitive to strong interaction, but only to electro-magnetic, weak, and gravitational interactions. Quarks cannot be observed as free entities, since they are always confined into composite structures of finite dimensions called hadrons. In particular, their mass can only be measured as effective masses within bound system. Figure 1.5 shows an overview of the organization of particles in the Standard Model.

The $\mathcal{SM}$ is a gauge theory under the local symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$, where C, L and Y respectively stand for the color, the left-handed chirality, and the hyper-charge. Lie

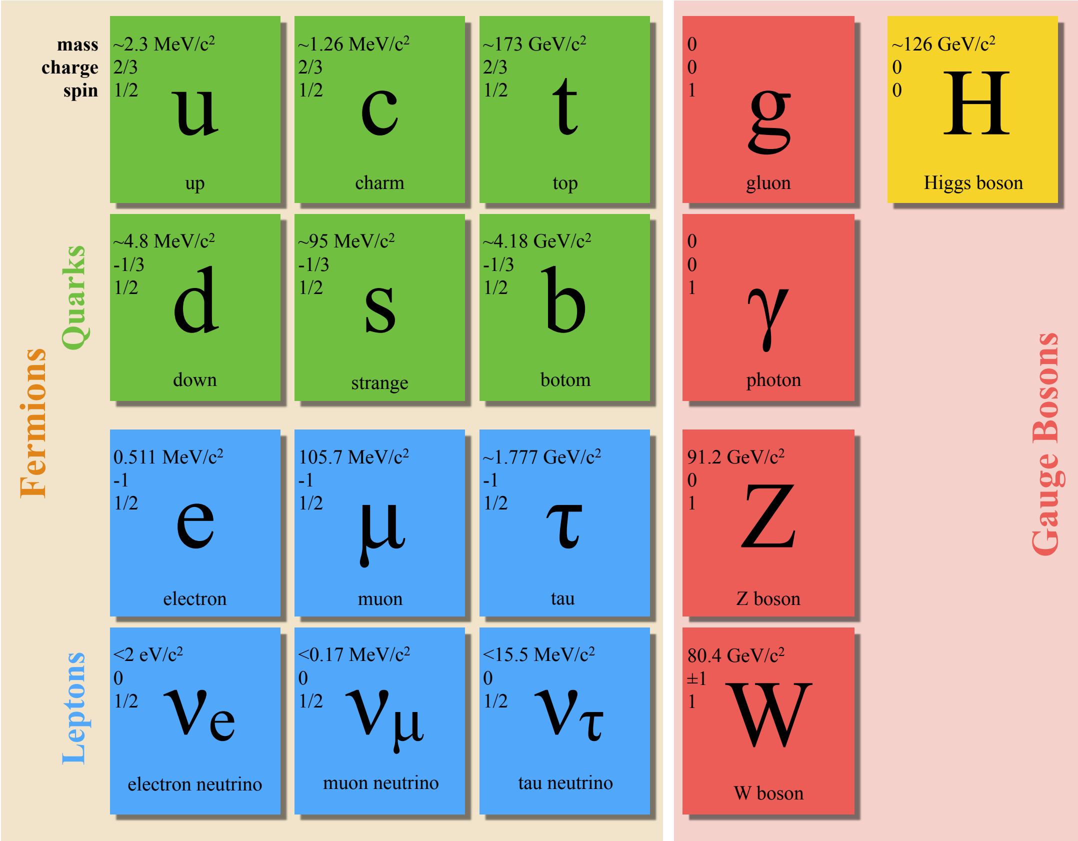
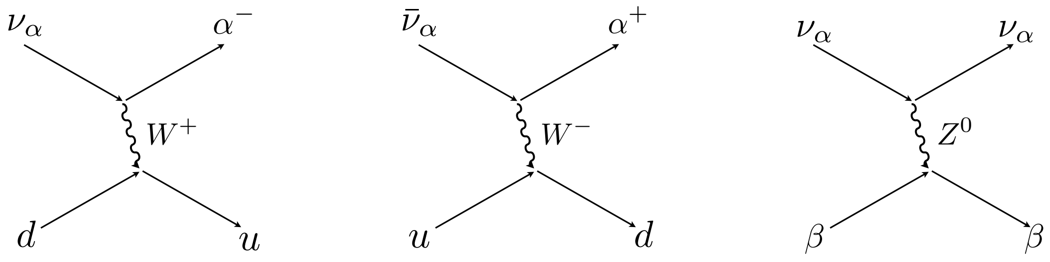


Figure 1.5: Overview of the Standard Model of elementary particles and gauges bosons. Anti-particles are not shown here for brevity.

algebras determine that a $SU(n)$ group has $n^2 - 1$ generators, while $U(n)$ groups have n^2 generators. Therefore $SU(3)_C$ has eight generators, the gluons, $SU(2)$ has three generators, the $W^{\pm}$ and Z^0 bosons, and $U(1)$ has one generator, the photon γ. $SU(3)_C$ can be separated from our study because the symmetry of quantum chromodynamics (QCD) is left unbroken by the electro-weak interaction from $SU(2)_L \times U(1)_Y$, hence the separation between quarks and leptons.

Even in the lepton sector, neutrinos have a particular role as they are neutral, and thus are not sensitive to the electro-magnetic interaction. Their interactions are only guided by the $SU(2)$ symmetry group, i.e., they can only interact through the exchange of $W^{\pm}$ or Z^0 bosons. Figure 1.6 shows the three possible interaction of neutrinos in the $\mathcal{SM}$. The exchange of a $W^{\pm}$ boson is called Charge Current (CC), and the exchange of a Z^0 boson is called Neutral current (NC).



(a) CC (b) CC (c) NC

Figure 1.6: Examples of Charged Current interactions (CC) and Neutral Current (NC) for neutrino and anti-neutrino of flavor α.
In a CC, the outgoing lepton is specific to the flavor of the neutrino, and the sign is specific of it being neutrino/anti-neutrino.
NC is not flavor specific and corresponds to elastic scattering of neutrinos or anti-neutrinos on a given particle.

NC is not flavor-specific, neutrinos of any flavor can interact with any quark or lepton. It is typically through this interaction that ν_μ or ν_τ scatter on the electrons of the conventional matter in Cerenkov detectors.

CC however is flavor-specific, i.e. a ν_α interacting through CC can only produce a α. It is then possible to differentiate between neutrinos of two flavors, in the same beam for instance. For ν_e undergoing elastic scattering on electrons, it is interesting to note that CC and NC will both happen, while only NC is available for ν_μ and ν_τ, thus increasing the ν_e cross section.

Following the results of Wu and Goldhaber on the Charge Partiy symmetry breaking of the electro-weak interaction, fermions are described as left-handed doublets, that can interact with gauge bosons, and right-handed singlets that cannot. However, the existence of a right-handed neutrino has never been observed, as such, neutrinos do not have a right-handed component in the $\mathcal{SM}$, as a consequence, anti-neutrinos are always right-handed.

1.3 Missing Neutrinos

1.3.a The Solar Neutrino Problem

Once neutrinos were discovered, and it was found that no shielding could prevent their presence in sensitive experiments, it became important to characterize precisely their properties and sources. Neutrinos were particularly a dangerous background to proton-decay experiments. The Sun is a bright source of neutrinos. The two main neutrino production processes in the Sun are the so-called *pp chain* and *CNO cycle*:

$$p + p \rightarrow d + e^+ + \nu_e \qquad (pp) \tag{1.8}$$

Although the *pp* chain is the dominant neutrino production process, other sub-dominant processes occur, but the maximum energy of the neutrinos in these processes make them easier to detect.

$$\begin{aligned} p + e^- + p &\rightarrow d + \nu_e && (pep) \\ {}^3_2\mathrm{He} + p &\rightarrow {}^4_2\mathrm{He} + e^+ + \nu_e && (hep) \\ {}^7_4\mathrm{Be} + e^- &\rightarrow {}^7_3\mathrm{Li} + \nu_e && ({}^7\mathrm{Be}) \\ {}^8_5\mathrm{B} &\rightarrow {}^8_4\mathrm{Be}^* + e^+ + \nu_e && ({}^8\mathrm{B}) \end{aligned} \tag{1.9}$$

Organic elements are produced in the Sun through the CNO cycle:

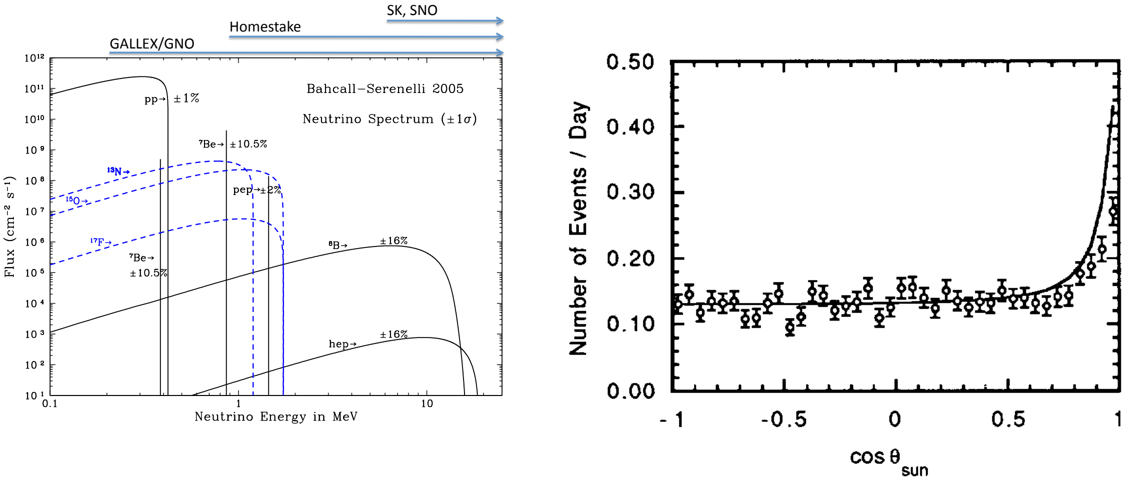
$$\begin{aligned} p + {}^{12}\mathrm{C} &\rightarrow {}^{13}\mathrm{N} + \gamma \\ {}^{13}\mathrm{N} &\rightarrow {}^{13}\mathrm{C} + e^+ + \nu_e \\ p + {}^{13}\mathrm{C} &\rightarrow {}^{14}\mathrm{N} + \gamma \\ p + {}^{14}\mathrm{N} &\rightarrow {}^{15}0 + \gamma \\ {}^{15}\mathrm{O} &\rightarrow {}^{15}\mathrm{N} + e^+ + \nu_e \\ p + {}^{15}\mathrm{N} &\rightarrow {}^{12}\mathrm{C} + \alpha \end{aligned} \tag{1.10}$$

The neutrino generation in the pp-chain and in the CNO-cycle can be summarized as $4p \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e$, releasing an energy of 26.731 MeV. With this energy release, it is possible to compute the solar neutrino flux expected on earth as:

$$\Phi_{\nu_e} = 2 \times \frac{L_\odot}{Q \times 4\pi R^2} = 6.385 \times 10^{10}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}, \tag{1.11}$$

with $L_\odot = 3.846 \times 10^{26}$W the Sun luminosity, and $R = 149.598 \times 10^6$km the astronomical unit.

Figure 1.7(a) shows the spectra of the different neutrino production processes in the sun, normalized to their expected fluxes. At the top of the figure are shown the energy ranges of sensitivity of the different solar experiments. Figure 1.7(b) shows the angular distribution of events in the Kamiokande detector with respect to the direction of the Sun. The Kamiokande detector is a Cerenkov detector, ν_e from the sun interact through elastic scattering (here it could be CC or NC) on electrons, and the recoiling electron produces a Cerenkov light cone that can be used to reconstruct the energy and direction of the incoming neutrino. An excess of events is observed in the direction of the Sun ($\cos\theta_{sun} = 1$) with respect to the other directions, corresponding to solar neutrinos. However, a deficit of $\sim 50\%$ is observable when comparing to the prediction.



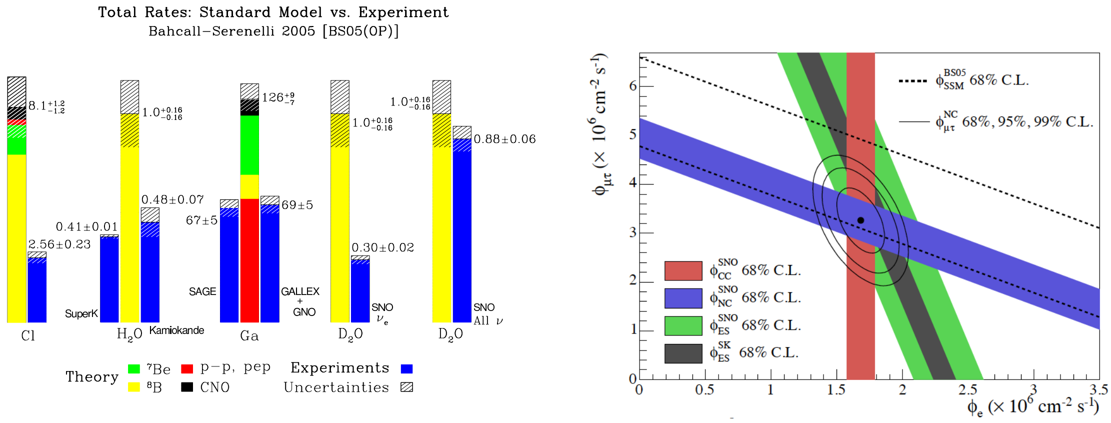
(a) (b)

Figure 1.7: 1.7(a): Spectra of the various neutrino production processes in the Sun. The black lines correspond to the pp-chain decays, and the blue lines to the CNO cycle [21].
1.7(b): Reconstructed angle of events in the Kamiokande detector with respect to the direction of the Sun. $\cos(\theta_{sun}) = 1$ corresponds to the direction of the Sun. A deficit is observed between the data points and the prediction (solid line) [22].

This observation is consistent with other flavor-specific solar neutrino experiments as presented in Figure 1.8(a). However, the SNO experiment, located at the Sudbury Neutrino Observatory measures the expected flux when accounting for all neutrino interactions. The particularity of the SNO experiment is the use of heavy water, allowing for a measurement of the neutrino-deuteron reaction forming two protons (CC on the neutron, possible only for ν_e), or resulting to a proton and a neutron (NC on the neutron or proton, possible for all flavors), but also the elastic scattering (ES) on electrons, producing Cerenkov light, and available for all flavors. Therefore, the SNO experiment is sensitive to all three flavors.

Figure 1.8(b) shows a common fit of the SNO and Super-Kamiokande experiments, comparing the flux observed by CC, hence a pure ν_e flux Φ_e, to fluxes measures by NC and ES. The last two detection processes are available to all flavor, so the measured flux is the sum of Φ_e and the flux of neutrino with different flavor $\Phi_{\mu\tau}$. Super-Kamiokande only sees the ES process. The width of the bands corresponds to a 68% confidence level (C.L.). A global fit, the result of which is shown by the ellipses, is performed to estimate the fraction of ν_e that have changed flavors into ν_μ or ν_τ. The prediction based on the Solar Standard Model (SSM) is shown by the two parallel dotted lines (68% C.L.) and is perfectly compatible with the best fit value from SNO and Super-Kamiokande. The concentric ellipses show the areas in which the true values has a probability of being of 68% to 99% from the center outward.

The SNO experiment has provided the proof of the existence of a flavor-changing behaviour in the propagation of solar neutrinos.



(a) (b)

Figure 1.8: 1.8(a): All flavor-selective experiments see a neutrino flux compatible with 1/3 to 1/2 of the expected flux from SSM. SNO, by using NC, is sensitive to all 3 flavors and see all expected neutrinos.
1.8(b): Fluxes of ^{8}B neutrinos from SNO and Super-Kamiokande for different detection processes. The global fit (ellipse) is compatible with SSM prediction (dotted lines)[4].

1.3.b The Atmospheric Anomaly

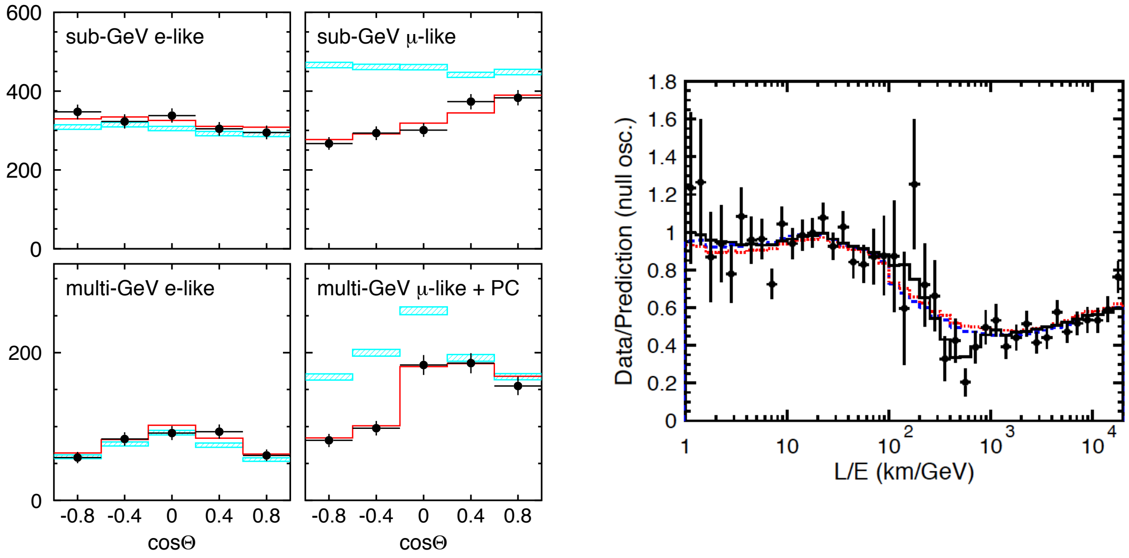
Cosmic rays, mainly high energy protons and heavy nuclei, imping on the high atmosphere continuously. They interact with nuclei from the atmosphere and produce secondary showers containing all kinds of hadrons, which eventually decay in pions and kaons. Those pions and kaons in turn decay into muons and muon-neutrinos, and the muons decay into electrons, muon-neutrinos and electron-neutrinos.

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{1.12}$$
$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{1.13}$$

A ratio of ν_μ/ν_e of 2:1 (within systematic uncertainties of the order of 10%) is expected from these showers, but in the middle of the 1980s, The IMB and Kamiokande experiments reported ratios of respectively 0.54 and 0.60 [23]. However, the number of ν_e detected was well within the uncertainties, hence a disappearance of the ν_μ flux had been observed.

The rate of high energy muons (above ~GeV) from atmospheric showers is uniformly distributed throughout the surface of the Earth are uniform, thus, so should the ν_μ rate be. Therefore, the flux of down-going ν_μ and up-going ν_μ in a Cerenkov detector such as Super-Kamiokande should be identical.

Figure 1.9(a) shows the angular distribution for e-like and μ-like events, separated into two energy ranges, below and above 1 GeV. The ν_e flux is perfectly consistent with the predictions, however, the ν_μ flux shows a clear deficit for up-going events ($\cos\theta = -1$). These neutrinos had to journey through the full Earth's diameter, while the down-doing neutrinos only had to cross the atmosphere. For the same energy range, a disappearance of ν_μ occurs as a function of the distance travelled. Since the ν_e flux is not affected at first order, the disappeared ν_μ must have turned into ν_τ.



(a) (b)

Figure 1.9: 1.9(a): Angular distribution of e-like and μ-like events in Super-Kamiokande for sub-GeV and multi-GeV energy ranges. $\cos\theta = 1$ corresponds to down-going neutrinos (only crossing the atmosphere), and $\cos\theta = -1$ corresponds to up-going neutrinos, crossing the atmosphere and the full Earth's diameter. The blue boxes correspond to the prediction and the black dots to the data. The red line shows the best fit assuming $\nu_\mu \leftrightarrow \nu_\tau$ oscillations [24].
1.9(b): Super Kamiokande oscillation result: the ratio of detected neutrino over the un-oscillated expectation exhibits an oscillation pattern [5].

Figure 1.9(b) shows the ratio of the observed ν_μ flux over the predicted ν_μ flux as a function of the L/E ratio where L is the distance travelled (from the muon track angle), and E is the neutrino energy. The sinusoidal pattern, with a first minimum at 500 km/GeV is called the atmospheric neutrino oscillation.

1.4 Neutrino Oscillations

Neutrino oscillations were first introduced by Pontecorvo in 1957 to describe the possibility of a $\nu \leftrightarrow \bar{\nu}$ oscillation, as an analogy to the $K^0 \leftrightarrow \bar{K}^0$ oscillation in the quark sector. The $\nu \leftrightarrow \bar{\nu}$ has yet to be observed today, but is served as foundation for the flavor oscillation, introduced by Maki, Nakagawa and Sakata in 1962 [25], and evidenced by the SNO and Super-Kamiokande experiment as described earlier in this chapter.

1.4.a Oscillations in the vacuum

The oscillation formalism requires that neutrinos have a mass, that at least two of these masses are distinct, and that the flavor eigenstates do not coincide with the mass eigenstates.

Neutrino Mixing Parametrization

Under these hypotheses the neutrino can be described equivalently in two bases: the basis of the electro-weak flavor eigenstates $\{|\nu_\alpha\rangle\}_{\alpha=e,\mu,\tau}$, and the basis of the mass eigenstates $\{|\nu_i\rangle\}_{i=1,2,3}$.

The flavor and mass eigenstates are two bases of the same vectorial space, therefore, there is a unitary $N \times N$ matrix, called U_{PMNS}, for Pontecorvo, Maki, Nakagawa, and Sakata, that allows to write:

$$|\nu_\alpha\rangle = \sum_i U^*_{\alpha i}|\nu_i\rangle. \tag{1.14}$$

In the three-neutrino scheme, the PMNS matrix is analogous to a 3D rotation matrix and is parametrized by three Euler angles θ_{ij}, called mixing angles, and a CP-violation phase δ_{CP}:

$$U_{\text{PMNS}} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}, \tag{1.15}$$

$$U_{\text{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.16}$$

where c_{ij} stands for $\cos(\theta_{ij})$ and s_{ij} stands for $\sin(\theta_{ij})$.

The first matrix, parametrizing the $\nu_\mu \leftrightarrow \nu_\tau$ oscillations with the θ_{23} mixing angle, is known as the atmospheric sector. The second matrix, parametrizing the $\nu_e \leftrightarrow \nu_\tau$ oscillations with the θ_{13} mixing angle, is called the reactor sector; it is this angle that Double Chooz measures. The third matrix, parametrizing the $\nu_e \leftrightarrow \nu_\mu$ oscillation with the θ_{12} mixing angle is the solar sector.

When this parametrization was first established, θ_{13} was the last unknown mixing angle, and was thought to be very small as only a upper limit had been place by the CHOOZ experiment [26]. The effects of δ_{CP} cannot be observed, should one of the angles be zero. δ_{CP} was therefore arbitrarily placed in the reactor matrix. If one considers neutrinos as Majorana particles, a fourth matrix needs to be added $M_{\text{Majorana}} = \text{diag}(1, e^{i\phi_1/2}, e^{i\phi_2/2})$ with two additional phases ϕ_1 and ϕ_2.

This description can easily be extended to a N-neutrino scheme, with N flavor and mass states. The $N \times N$ unitary matrix can be parametrized by $N(N-1)/2$ Euler angles and $N(N+1)/2$ CP-violating phases. If the neutrino is a Dirac particle, only $(N-1)(N-2)/2$ phases are physical. If, however, the neutrino is a Majorana particle, the matrix contains $N(N-1)/2$ CP-violating phases.

Oscillation in the General N-Neutrino Framework

The mass states $|\nu_i\rangle$ are the eigenstates of the Hamiltonian $H_0|\nu_i\rangle = E_i|\nu_i\rangle$, where E_i is the special relativity energy $E_i = \sqrt{p_i^2 + m_i^2}$, with p_i is the momentum of the neutrino, and m_i is the mass of the neutrino state $|\nu_i\rangle$.

A single mass eigenstate propagates through space as a plane wave, described by the Schrödinger equation:

$$i\bar{h}\frac{\partial}{\partial t}|\nu_i(t)\rangle = H_0|\nu_i(t)\rangle \qquad \text{with} \qquad |\nu_i(t)\rangle = e^{-iE_it}|\nu_i\rangle \tag{1.17}$$

A flavor eigenstate can therefore be written:

$$|\nu_\alpha(t)\rangle = \sum_i U^*_{\alpha i} e^{-iE_it}|\nu_i\rangle, \tag{1.18}$$

and a mass eigenstate can be described as a linear combination of the flavor eigenstates:

$$|\nu_i\rangle = \sum_\beta U_{\beta i}|\nu_\beta\rangle. \tag{1.19}$$

By combining Equations 1.18 and 1.19, the flavor eigenstates can be expressed as:

$$|\nu_\alpha\rangle = \sum_\beta \left(\sum_i U^*_{\alpha i} e^{-iE_it} U_{\beta i}\right)|\nu_\beta\rangle. \tag{1.20}$$

Even created as a pure electro-weak flavor state, the neutrino is a superposition of all flavor states, and the relative weights of each state evolves over time. The probability of transition from a state $|\nu_\alpha\rangle$ to $|\nu_\beta\rangle$ can be computed by

$$P_{\nu_\alpha\to\nu_\beta}(t) = |A_{\nu_\alpha\to\nu_\beta}(t)|^2, \tag{1.21}$$

where A is the probability amplitude

$$A_{\nu_\alpha\to\nu_\beta}(t)\langle\nu_\beta|\nu_\alpha(t)\rangle = \sum_i U^*_{\alpha i} U_{\beta i}\ e^{-iE_it}, \tag{1.22}$$

thus the probability $P_{\nu_\alpha\to\nu_\beta}(t)$ is

$$P_{\nu_\alpha\to\nu_\beta}(t) = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U^*_{\alpha j} U_{\beta j}\ e^{-i(E_i-e_j)t}. \tag{1.23}$$

Neutrinos are described as massless in the $\mathcal{SM}$, and no experiment was able to reach the sensitivity required to measure their mass so far, setting limits below the 1 eV scale. It is therefore reasonable to assume that neutrinos are ultra-relativistic ($m_i \ll p_i$), especially so in Double Chooz where the mean neutrino energy is $\langle E\rangle \sim 4\,\text{MeV}$. The energy E_i can thus be re-written:

$$E_i \simeq E + \frac{m_i^2}{2E}, \tag{1.24}$$

and $t \simeq L$ in the natural unit system ($c = 1$). Then, Equation 1.23 becomes:

$$P_{\nu_\alpha\to\nu_\beta}(L,E) = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U^*_{\alpha j} U_{\beta j} e^{-i\frac{\Delta m^2_{ij}L}{2E}}. \tag{1.25}$$

where $\Delta m_{ij}^2 = m_i^2 - m_j^2$. Finally, by decomposing the complex exponential, we find

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(L,E) = \delta_{\alpha\beta} - 4 \sum_{i=2}^{N} \sum_{j=1}^{i-1} \mathrm{Re} \left[U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta i}^* \right] \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right) \\ \pm 2 \sum_{i=2}^{N} \sum_{j=1}^{i-1} \mathrm{Im} \left[U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta i}^* \right] \sin^2 \left(\frac{\Delta m_{ij}^2 L}{2E} \right) \end{aligned} \tag{1.26}$$

where $\delta_{\alpha\beta} = 1 \Leftrightarrow \alpha = \beta$. The probability $P_{\nu_\alpha \to \nu_\beta}$ depends on L/E through sine functions, hence the denomination of *oscillations*. The wavelength of each sinusoid component varies as $E/\Delta m^2$. The ± sign in the third line corresponds to a difference neutrino/anti-neutrino, the + corresponding to the $\nu \leftrightarrow \nu$ oscillation, and the − to the $\bar{\nu} \leftrightarrow \bar{\nu}$ oscillation.

Mass Hierarchy

The oscillation depends on a parameter Δm_{ij}^2 which is the difference of square masses of the $|\nu_i\rangle$ states. In the three-neutrino scheme, three mass-squared differences are considered:

$$\begin{aligned} \Delta m_{21}^2 &= m_2^2 - m_1^2 \simeq 7.59^{+0.19}_{-0.21} \times 10^{-5} \mathrm{eV}^2, \\ \Delta m_{32}^2 &= m_3^2 - m_2^2 \simeq 2.43 \pm 0.13 \times 10^{-3} \mathrm{eV}^2, \\ \text{and } \Delta m_{31}^2 &= m_3^2 - m_1^2 \simeq 2.44^{+0.09}_{-0.10} \times 10^{-3} \mathrm{eV}^{\ 2}. \end{aligned}$$

Only two are independent, Δm_{21}^2 and $|\Delta m_{32}^2|$ have been measured respectively by KamLAND [27] and MINOS [28]. Those results show that either $m_1 < m_2 \ll m_3$ or $m_3 \ll m_1 < m2$, as shown in the cartoon Figure 1.10. The two possible configurations are called respectively Normal Hierarchy or Inverted Hierarchy.

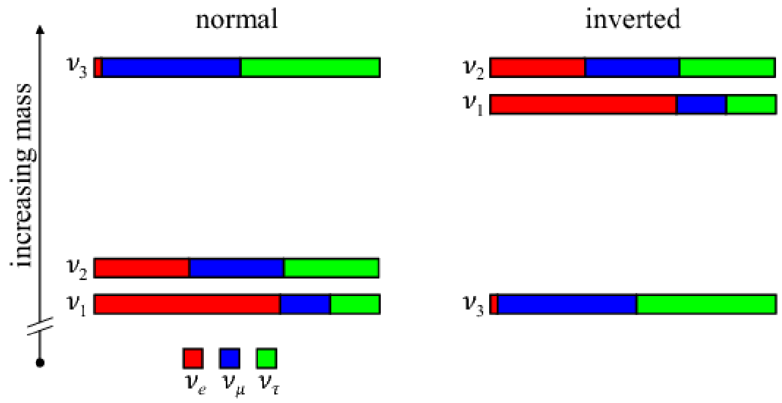


Figure 1.10: Sketch illustrating the mass hierarchy degeneracy in neutrino oscillations. The vertical axis corresponds to the relative mass of each $|\nu_i\rangle$ mass states, and the colors represent the fraction of the flavor state $|\nu_\alpha\rangle$ in each mass state [29].

Case of a Two-Neutrino Oscillation

The two orders of magnitudes between the mass square differences allow most experiments to be in one domain or another, but be dominated by a single oscillation, analogous to a two-neutrinos scheme.

In the two-neutrino case we consider the two flavor eigenstates $|\nu_\alpha\rangle$ and $|\nu_\beta\rangle$, and the two mass eigenstates $|\nu_1\rangle$ and $|\nu_2\rangle$. The mixing matrix is only parametrized by a single angle θ, and no CP-violation phase.

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{1.27}$$

By applying this parametrization to Equation 1.26, we find:

$$P_{\nu_\alpha \to \nu_\beta}(L, E) = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \tag{1.28}$$

$$P_{\nu_\alpha \to \nu_\alpha}(L, E) = 1 - P_{\nu_\alpha \to \nu_\beta}(L, E) = 1 - \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \tag{1.29}$$

Here, the maximum amplitude of the oscillation is given by the value of $\sin^2(2\theta)$.

Equation 1.28 describes the probability of a ν_β being detected in an originally pure ν_α flux. This phenomenon is studied in so-called appearance experiments. Equation 1.29 is called Survival Probability, it is the probability to detect a ν_α from an originally pure ν_α flux. This phenomenon is studied in so-called disappearance experiments. Reactor experiments such as Double Chooz use this method to quantify the mixing angle θ.

Oscillations in Matter

As neutrinos journey through matter, these oscillation probabilities are affected. As electrons are a common component of ordinary matter while muons and taus are not, the ν_e component of the travelling wave packet can undergo CC and NC while the other component can only undergo NC. As a result, all flavors are not affected the same way, by the additional potential added to the vacuum Hamiltonian. Therefore, an effective mixing matrix with an additional phase can be written, depending on the electron density and the neutrino energy. This effect, called the MSW[1] effect [30] does not enter the scope of this study, focused on reactor neutrino where matter effects are negligible.

1.5 Determination of the Oscillation Parameters

In the three-neutrino paradigm, oscillations are parametrized by three mixing angles θ_{ij}, two independent mass squared differences Δm^2_{ij}, and a CP-violation phase δ_{CP}. This is far too many parameters to be determined at once. However, neutrino experimentalists have been

---

[1] MSW stands for Mikheyev, Smirnov and Wolfenstein

lucky enough that two orders of magnitude separate the mass squared differences.

Indeed, in the oscillation probability, terms in $\sin^2\left(\frac{\Delta m^2}{4E}L\right)$ drive the oscillations with oscillation lengths $L_{\text{osc}} = \frac{4\pi E}{\Delta m^2}$. Three possibilities arise from here:

- $E/L \gg \Delta m^2$, i.e. $L \ll L_{\text{osc}}$, the oscillation driven by Δm^2 have not developed yet, the experiment is not sensitive to this oscillation.
- $E/L \ll \Delta m^2$, i.e. $L \gg L_{\text{osc}}$, the finite energy resolution and spatial reconstruction of the detector make it impossible to distinguish the oscillation, the experiment only sees an average constant value $\frac{1}{2}\sin^2(2\theta)$.
- $E/L \sim \Delta m^2$, i.e. $L \sim L_{\text{osc}}/4$, the oscillation is close to its maximum amplitude, the experiment is maximally sensitive to this particular oscillation.

It is therefore possible to study each oscillation independently from the others, in three sectors:

- the solar sector,
- the atmospheric sector,
- the reactor sector.

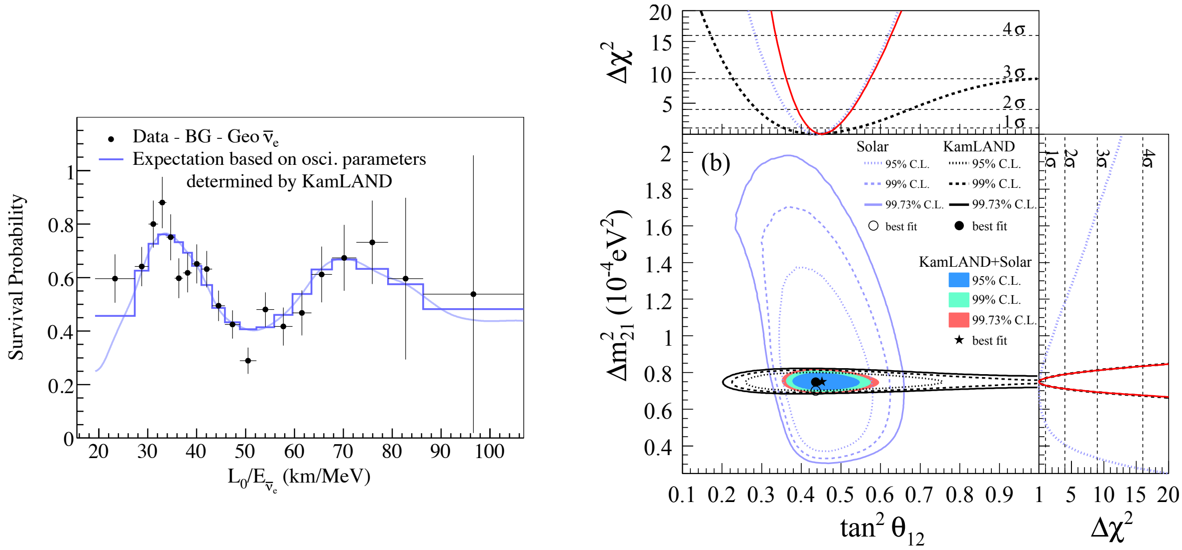
1.5.a The Solar Sector

The deficit from solar neutrinos, first evidenced by the Homestake experiment, and confirmed by many independent observations including the Gallex, Sage or Super Kamiokande experiments, has been one of the first observations of the neutrino oscillations. While solar experiments such as SNO [4] provide a good constraint on the amplitude of the oscillation $\sin^2(2\theta_{12})$, they poorly constrain the oscillation length through Δm_{21}.

The KamLAND experiment [27] studies reactor neutrino disappearance from all Japanese reactors over a mean baseline of 175 km. The L/E ratio of this configuration allow KamLAND to be sensitive to the solar oscillation (Δm^2_{21} and θ_{12}). The various distances from all the reactors allowed to scan a large L/E phase space, and to obtain the oscillation profile of the survival probability shown in Figure 1.11(a). The averaging of the oscillation is visible at high L/E, limiting the sensitivity to $\sin^2(2\theta_{12})$, but a precise measurement of Δm^2_{21} is possible. The measurement of Δm^2_{21} demonstrated that vacuum oscillation could not explain the solar neutrino deficit, and that matter effects were needed (MSW effect).

By combining the measurements from KamLAND to the other solar experiments as shown in Figure 1.11(b), a precise determination of both parameters is possible [27]:

$$\Delta m^2_{21} = 7.58^{+0.22}_{-0.26} \times 10^{-5}\text{eV}^2 \text{ and } \sin^2(2\theta_{12}) = 0.304^{+0.022}_{-0.016} \tag{1.30}$$



(a) (b)

Figure 1.11: 1.11(a): Oscillation pattern of the survival probability as a function of the L/E ratio [27]. 1.11(b): Global fit of KamLAND results with the other solar neutrino experiments [31].

1.5.b The Atmospheric Sector

Atmospheric neutrinos can be studied by comparing the rates of neutrinos according to their azimuthal angle, and therefore the distance they had to travel since their production in the atmosphere, distance ranging from ~ 20 km to $\sim 13\,000$ km, as it is the case for experiments such as Super-Kamiokande. Since atmospheric oscillations are $\nu_\mu \leftrightarrow \nu_\tau$ oscillations, they can also be studied in accelerator-based experiments such as T2K and MINOS.

The MINOS experiment uses neutrinos from the NuMI beamline located at the FNAL to study atmospheric oscillations over a 735 km baseline. MINOS is composed of a near detector, located at the FNAL, and a far detector, MINOS+, located in the Soudan Mine.

Figure 1.12, from [32], shows the results of the allowed regions in the $(\sin^2(2\theta_{23}), \Delta m^2_{32})$ space for Super-Kamiokande, MINOS and T2K.

The best fit values for a global analysis are [31]:

$$|\Delta m^2_{32}| = 2.35^{+0.12}_{-0.09} \times 10^{-3}\text{eV}^2 \text{ and } \sin^2(2\theta_{23}) = 0.42^{+0.08}_{-0.03} \tag{1.31}$$

Given the high energies of atmospheric neutrinos, and the relatively low electron density of the Earth, the MSW effects are negligible, even over long distances. Therefore, the degeneracy of the mass hierarchy remains un-lifted and therefore only an absolute value can be measured.

1.5.c The Reactor Sector

θ_{13} was the last measured mixing angle, with a definitive non-zero measurement reported in 2012 by the latest generation of θ_{13} experiments Double Chooz [7], Daya Bay [33], and RENO [34]. The previous best limit had been obtained by the CHOOZ experiment in 2003 [26], but did not rule out the possibility that θ_{13} could be null. For this reason, the CP violation phase

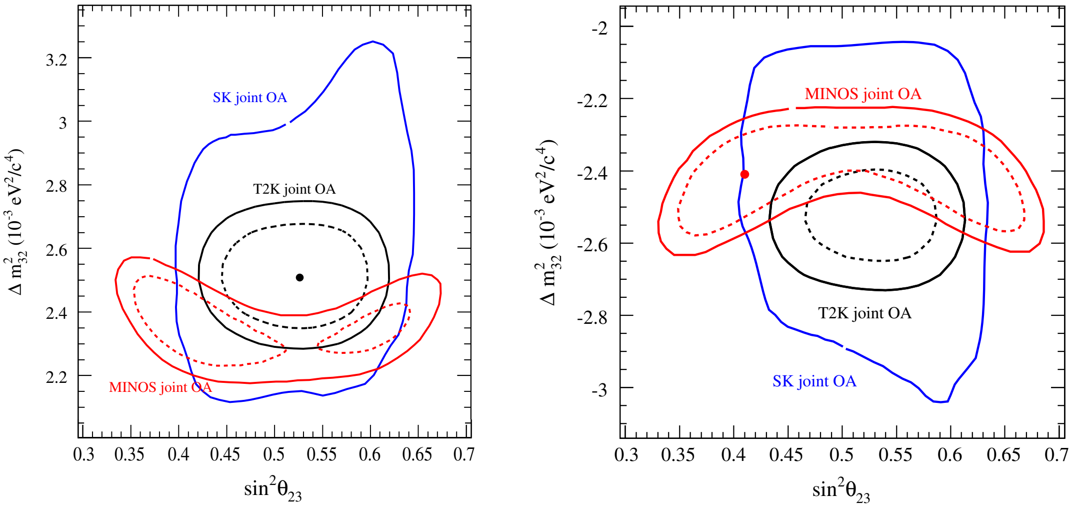


Figure 1.12: Determination of atmospheric parameters by MINOS, T2K and Super-Kamiokande [32].

δ_{CP} was parametrized together with the reactor sector, as they were the last two unknown parameters. However, θ_{13} does not have any particular role with respect to δ_{CP} compared to the other mixing angles.

Reactor neutrinos are ideal to study θ_{13}, indeed, given their energy and the value of Δm^2_{31}, the maximum of oscillation is located at a distance of one kilometer from the reactor cores, and the oscillation probability is well approximated by a two-flavor oscillation.

Figure 1.13 shows the survival probability for 3 MeV $\bar{\nu}_e$. The blue and orange lines corresponds respectively to the reactor and solar contributions to the full oscillation in red. The location of the detectors of all reactor θ_{13} experiments are located around the first minimum, in a region dominated by θ_{13}. While the Far Detector is maximally sensitive to θ_{13}, the Near Detector provides an absolute normalization of the $\bar{\nu}_e$ flux, cancelling the uncertainties associated to the flux prediction. The detectors are designed to be identical, thus cancelling the correlated detection efficiency systematics.

CHOOZ

The CHOOZ experiment took place in the late 1990s at the Chooz nuclear power plant. The detector was located in the same laboratory than the one where now sits the Far Detector of Double Chooz. CHOOZ was an Inverse Beta Decay (IBD) experiment, using the well known process that allowed Reines and Cowan to first discover the neutrino, except in this case, the neutron capture occurred on Gadolinium instead of Cadmium.

The detector design is presented in Figure 1.14(a). It is built in three concentric volumes filled with liquid scintillator. The innermost volume, the neutrino target, is doped with Gadolinium; neutrons captures are observed in this volume only. A containment region surrounds the neutrino target, allows for a full collection of the energy of the events in the target. The walls

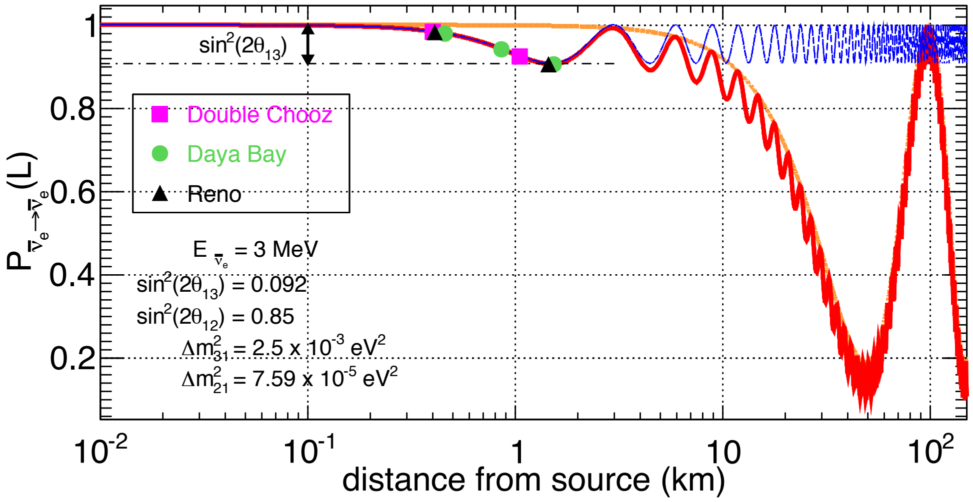


Figure 1.13: Oscillation pattern of the survival probability of reactor $\bar{\nu}_e$ as a function of the distance from the reactor for a 3 MeV $\bar{\nu}_e$. The points show the position of the detectors of the various reactor experiments. Each experiment relies on a near+Far Detector layout. The red full line corresponds to the full oscillation pattern, the orange dotted line is the solar oscillation part, dominated by θ_{12}, and the blue dotted part is the reactor oscillation part, dominated by θ_{13}.

of this containment region are equipped with PMTs. Optically insulated from the first two volumes, the third volume serves as a veto to reject outside events.

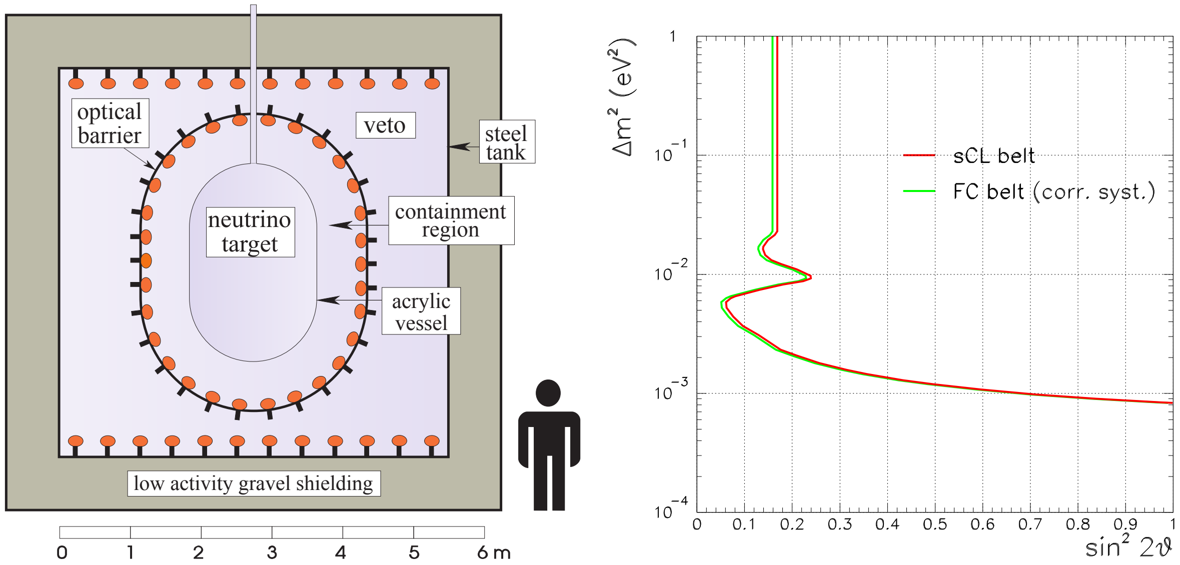
CHOOZ started taking data in 1997, before the commissioning of the reactors, and thus almost 40% of the run time was without reactor, allowing a precise background measurement. Sadly, the run time was limited by the degradation of the transparency of the scintillator after a year of data taking. CHOOZ results are statistics limited, and only an upper limit was set at $\sin^2(2\theta_{13}) < 0.14$ at 90% C.L. for $\Delta m^2_{31} = 2.5 \times 10^{-3}\text{eV}^2$.

Long Baseline Experiments

Long baseline experiments utilize accelerator ν_μ of $\sim 10\,\text{GeV}$ detected over distances of $\sim 100\,\text{km}$. Those apparition experiments are sensitive to $\sin^2(\theta_{13})$ through:

$$P_{\nu_\mu \to \nu_e}(L,E) = \sin^2(2\theta_{13})\sin^2(2\theta_{23})\sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right) \tag{1.32}$$
$$+ \text{ CP-violation corrections}$$
$$+ \text{ MSW effect corrections}$$

Accelerator-based experiments suffer from degeneracies linked to the mass hierarchy, and the unknown CP-violation. A measurement of θ_{13} from reactor experiments could help lift these degeneracies.



(a) (b)

Figure 1.14: Detector and final exclusion plot of the CHOOZ experiment. CHOOZ was sensitive to the region at the upper right of the lines and found no evidence of oscillations, this region is therefore excluded.

The T2K experiment uses the Super-Kamiokande detector to measure ν_e apparition in a ν_μ beam from the J-PARC facility at Tokai, Japan, over a baseline of 295 km. T2K presented the first indication for a non-zero θ_{13} by MINO value in 2011 [35]. Figure 1.15 shows the simultaneous measurement of δ_{CP} and $\sin^2(2\theta_{13})$ performed by T2K.

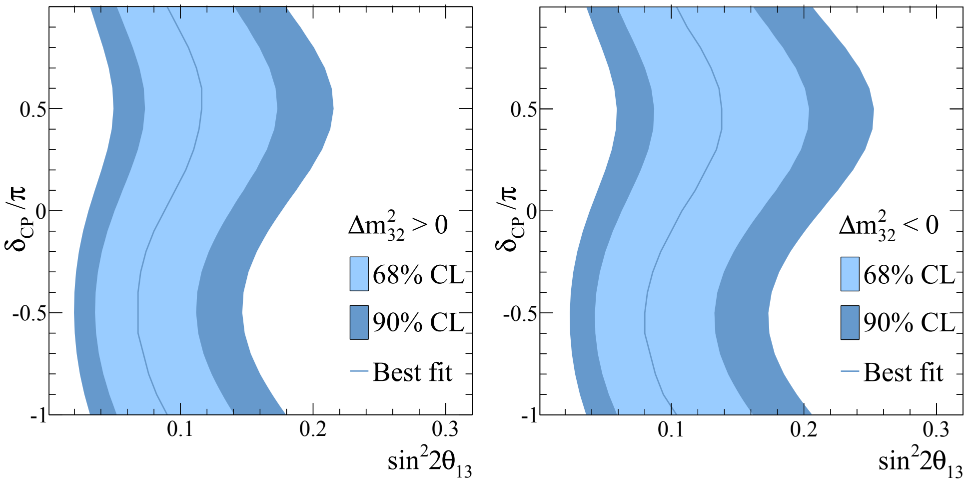


Figure 1.15: θ_{13} and $\delta_{\rm CP}$ fit performed by T2K for the Normal Hierarchy (left), and the Inverted Hierarchy (right) [36].

New Generation of Reactor Experiments

Indications from experiments such as T2K, MINOS, or KamLAND show that the value of $\sin^2(2\theta_{13})$ seems close to the limit of 0.14 set by CHOOZ. However, due to the too many unconstrained parameters, long baseline experiments cannot reach the level of sensitivity of a reactor-based experiment.

The reactor experiment Double Chooz, studied in more details in the following chapters, was designed from the CHOOZ experiment. The CHOOZ laboratory is now the location of the Far Detector of Double Chooz. Improving on the CHOOZ experiment, Double Chooz proposed a two-detector concept: a Near Detector, identical to the Far Detector, is meant to characterize the flux and spectral shape of the neutrinos emitted by the reactors before they oscillate. Comparing the observations in the Near and Far Detectors will allow to reduce the systematic error linked to the flux prediction. Moreover, an additional volume was added in the design of the detectors, filled with non-scintillating oil It separates the PMTs, an intense source of radioactivity background, from the scintillating detection volumes, allowing for a reduction of the backgrounds and their systematic uncertainties. Double Chooz also used a new scintillator, which proved much more stable in time than the one in CHOOZ, as the experiment has been running for five years and no significant loss of transparency has been found so far.

Based on the same principle as Double Chooz, the RENO [37] and Daya Bay [38] experiments vary the number of detectors and reactors. Located close to more reactors makes the control of reactor-based systematics more delicate, but significantly increases the statistics, effectively allowing Daya Bay to present the best θ_{13} measurement. Figure 1.16 shows the lay-out of the reactor and detector of the three reactor θ_{13} experiments.

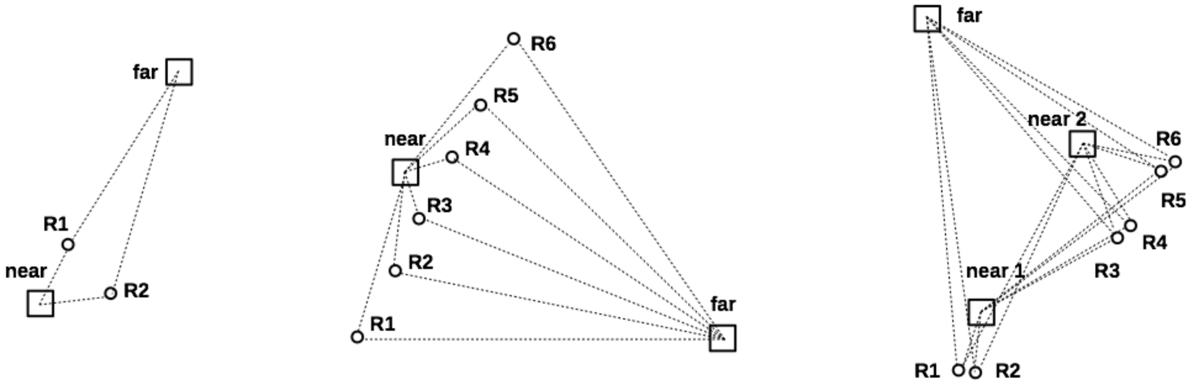
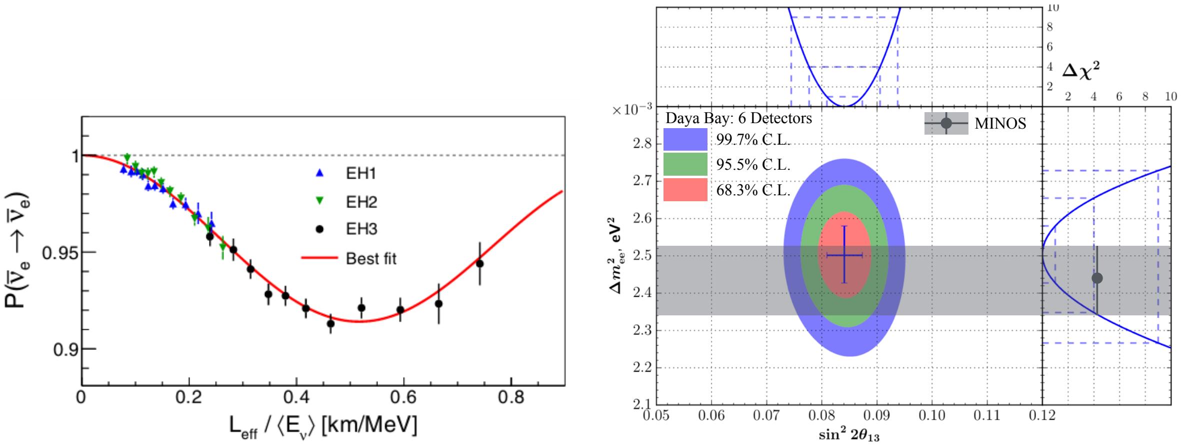


Figure 1.16: Reactor and detector layouts for θ_{13} reactor neutrino experiments:
left : Double Chooz, two reactors, two detectors,
middle : RENO, six reactors, two detectors,
right : Daya Bay, six reactors, three detector halls.

Figure 1.17(a) shows the ratio of Daya Bay data over the no-oscillation hypothesis (i.e. survival probability) for all its experimental halls and relative baselines L/E. The red line corresponds to a simultaneous fit of Δm^2 and $\sin^2(2\theta_{13})$. Figure 1.17(b) presents the 68.3%, 95.5%, and 99.7% confidence intervals on these two parameters [39]. The grey area is the 68.3%

C.L. contour of the current best measurement by MINOS [28].



(a) (b)

Figure 1.17: 1.17(a): Ratio of Daya Bay data over the un-oscillated prediction for all laboratories and relative baselines [39].
1.17(b): Confidence interval obtained by fitting the data points in Figure 1.17(a). In grey is shown MINOS 1σ confidence region.

The Reactor Anomaly

The multiple detector layout is needed for reactor θ_{13} experiments to provide an absolute flux normalization of the prediction. Indeed, a single detector experiment, such as CHOOZ, or the first phase of Double Chooz, could rely on simulations of the reactor cores and provide a reasonable neutrino flux prediction, as it will be shown in Chapter 2. However, previous short baseline measurements close to nuclear research or commercial reactors have found the observed flux to be consistently lower than expected. Figure 1.18 summarizes the ratio of the short baseline measurements measured flux over an un-oscillated prediction. The red line corresponds to the expectation from simulations, taking into account a three-neutrino oscillation paradigm. Most of the measurements are 1σ below the line, their average measured flux is shown by the blue line. This observation is called the Reactor Anomaly. The observed deficit is

$$\frac{\mathrm{N_{obs}}}{\mathrm{N_{exp}}} \simeq 0.94. \tag{1.33}$$

A first possible explanation for this inconsistency between data and prediction is an issue with the simulation of the reactors. A second hypothesis is the existence of a very short baseline oscillation ($\mathrm{L_{osc}} \sim 1\,\mathrm{m}$). Such an oscillation could be created by a fourth neutrino generation, with a high value of Δm^2. However, LEP experiments have shown that only three neutrino flavors couple to the weak interaction, an additional flavor should therefore not couple with the $W^\pm$ or Z^0 bosons, and neither with other interactions. Such neutrino are referred to as Sterile Neutrinos. Very short baseline oscillations would average out at longer baselines and we would only see a constant deficit of $\frac{1}{2}\sin^2(2\theta_{\mathrm{st}})$ on the un-oscillated prediction. Figure 1.19 shows

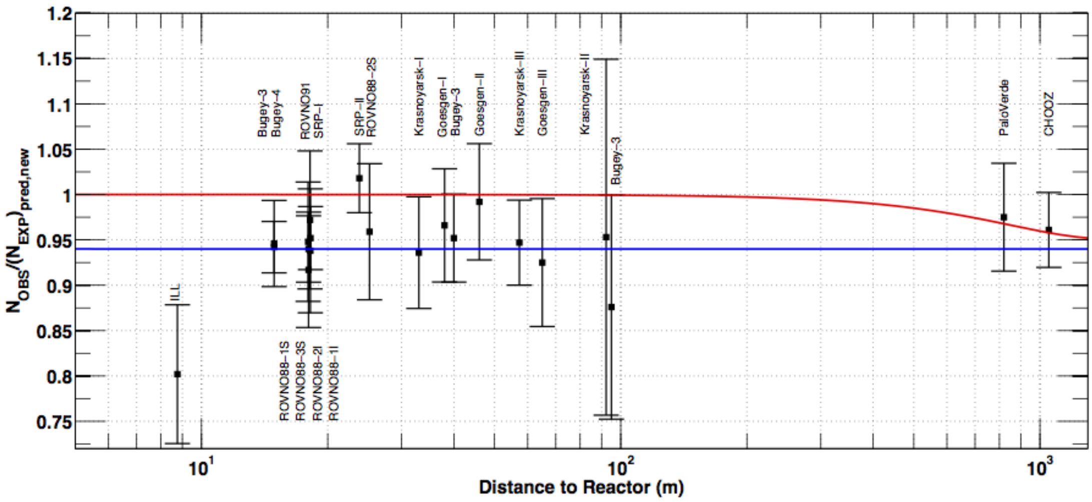


Figure 1.18: Illustration of the reactor anomaly. Experimental results are compared with the no oscillation hypothesis. The red line shows the evolution expected for a three neutrino oscillation paradigm. The blue line shows the average flux observed by these experiments.

a possible survival probability, in case there exists a sterile neutrino, that would reconcile the reactor experiments with the predictions.

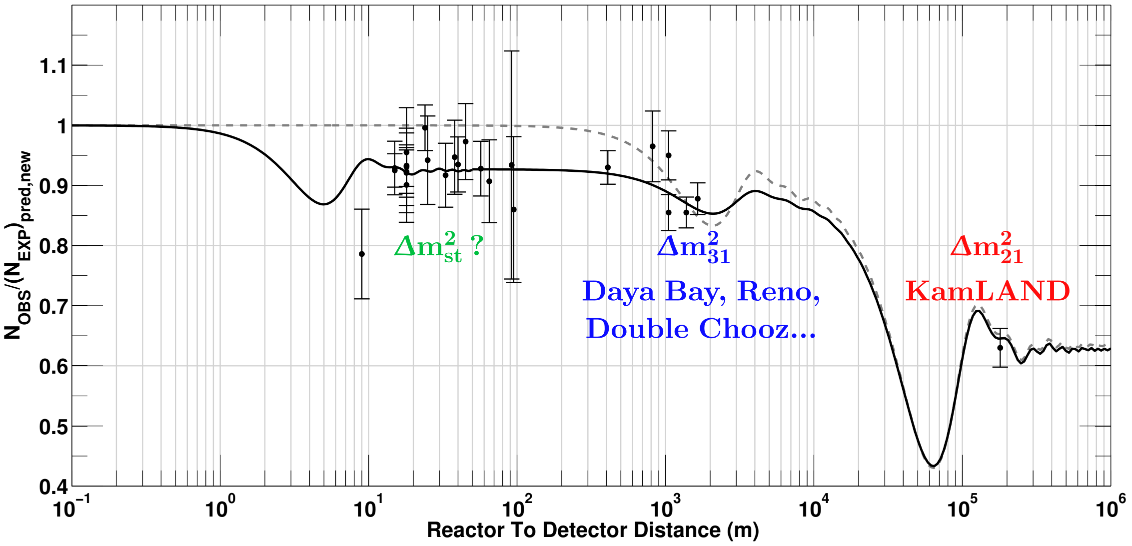


Figure 1.19: Possible interpretation of the reactor anomaly through the introduction of a sterile neutrino.

As of today, there are no clear evidence for a sterile neutrino, however, the reactor anomaly has not been solved yet. To address the absolute normalization issue seen by short baseline experiments, reactor θ_{13} experiments use the measurement from Bugey-4 [40], performed at $\sim 15\,\mathrm{m}$ from a reactor, to normalize their predictions. The use of a Near Detector also provides a normalization that further reduces the uncertainties introduced by the prediction.

1.6 Beyond Neutrino Oscillations

1.6.a Mass Generation

The symmetry groups of the $\mathcal{SM}$ presented above only dictates the number of generators (i.e. vector bosons), but they are described as massless particles. However, the range of an interaction depends on the mass of the boson, and the electro-weak interaction is found experimentally to be a short-range interaction, so the W and Z bosons should have a non-zero mass . The Brout-Englert-Higgs mechanism describes a symmetry breaking that predicts that the W and Z bosons are massive through the introduction of a scalar field $\Phi(x)$. This field can be written as a Higgs doublet:

$$\Phi(x) = \begin{pmatrix} \Phi^+(x) \\ \Phi^0(x) \end{pmatrix}, \tag{1.34}$$

where $\Phi^+(x)$ and $\Phi^0(x)$ are respectively the charged and neutral complex scalar fields. With the addition of the Higgs field, the $\mathcal{SM}$ electro-weak Lagrangian is written as:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{matter}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{1.35}$$

where each term corresponds respectively to the $SU(2)_L$ and $U(1)_Y$ gauge bosons, the kinetic energy of fermions, the Higgs fields, and the Yukawa coupling describing the fermion masses. The potential linked to the Higgs field is expressed:

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 = \lambda\left(\Phi^\dagger\Phi + \frac{\mu^2}{2\lambda}\right)^2 - \frac{\mu^4}{4\lambda} \tag{1.36}$$

where λ is the Higgs self-coupling parameter, and μ is the Higgs quadratic coupling parameter linked to the Higgs mass by $m_{\text{Higgs}} = \sqrt{-2\mu^2}$. Thus it follows that $\mu^2 < 0$ in order to have spontaneous symmetry breaking. The last term is constant and can be neglected, the potential is minimum for

$$\Phi^\dagger\Phi = -\frac{\mu^2}{2\lambda} \tag{1.37}$$

The minimum corresponds to the vacuum, the state with the lowest energy. The expectation value for Φ is called the vacuum expectation value (VEV) $\langle\Phi\rangle$. As only a scalar field can have a non-zero VEV, it follows:

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \sqrt{-\mu^2/\lambda} \end{pmatrix} \tag{1.38}$$

In general, a fermion in the $\mathcal{SM}$ can acquire a mass through the Yukawa coupling to the Higgs:

$$-\mathcal{L}_{\text{Yukawa}} = m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L), \tag{1.39}$$

therefore, a field needs both chiralities L and R (left and right handed) to have a mass. However, we saw that neutrinos only possess a left-handed component, which means they cannot acquire a mass in the $\mathcal{SM}$.

1.6.b Dirac or Majorana?

Dirac Mechanism

If a right-handed neutrino is added as a singlet to the $\mathcal{SM}$, the Yukawa coupling can be applied naturally as:

$$\mathcal{L}^D = -m_\nu^D(\bar{\nu}_R\nu_L + \bar{\nu}_L\nu_R) \tag{1.40}$$

In this case, neutrinos would play the same role as any other fermion in the $\mathcal{SM}$. The right handed neutrino being introduced as a singlet, it will not take part in any $\mathcal{SM}$ interaction (gauge bosons interact with the left-handed fields) and is therefore a sterile neutrino.
However, under this description, the Yukawa coupling for the neutrino would be of the order of magnitude of $G^\nu \sim 10^{-12}$ which is several orders of magnitude lower than that of the electron ($G^e \sim 0.3 \times 10^{-5}$) and therefore seems unphysical.

Majorana Mechanism

As neutrinos are the only neutral fermion of the $\mathcal{SM}$, it becomes possible for them to be their own anti particle. This description of neutral particle being their own anti-particle under the CPT symmetry, was proposed by Ettore Majorana, in 1937 [41].

The Lagrangian term for the Majorana coupling is written

$$\mathcal{L}^M = -\frac{1}{2}m_\nu^M(\bar{\nu}_L^c\nu_L + \bar{\nu}_L\nu_L^c), \tag{1.41}$$

where X^c stands for the Charge conjugate of X.

See-Saw Mechanism

Neither Dirac nor Majorana masses provide a satisfactory explanation of the smallness of the neutrino masses. However, a Lagrangian can be written using the two descriptions:

$$\mathcal{L}^{D+M} = -\frac{1}{2}\left(\bar{\nu}_L^c, \bar{\nu}_R\right)\begin{pmatrix} m_L^M & m^D \\ m^D & m_R^M \end{pmatrix}\begin{pmatrix} \nu_L \\ \nu_R^c \end{pmatrix} \tag{1.42}$$

where m_L^M and m_R^M are the masses of the left and right-handed Majorana neutrinos, and m^D is the Dirac mass obtained by the Yukawa coupling. The eigenvalues of this matrix are noted m_1 and m_2:

$$m_{2,1} = \frac{m_L^M + m_R^M}{2} \pm \sqrt{\left(\frac{m_L^M - m_R^M}{2}\right)^2 + (m^D)^2} \tag{1.43}$$

While no mass hierarchy has been fixed and no absolute mass has been measured, some assumptions are possible. Let us assume $m_L^M = 0$ to respect the $\mathcal{SM}$ symmetries and $m_R^M \gg m^D$ since ν_R is generated by a beyond-$\mathcal{SM}$ mechanism, and thus at higher energy scales, we find:

$$m_1 \simeq m_R^M \tag{1.44}$$

$$m_2 \simeq \frac{(m^D)^2}{m_R^M} \tag{1.45}$$

The eigenvalues lead to light Majorana neutrinos, mostly active ν_L with small mixing with sterile ν_R for the m_2 solution, and only heavy sterile Majorana neutrinos for the m_1 solution.

It is to be noted that, the heavier the sterile neutrinos are, the lighter the active ones are, hence the *See-Saw* denomination.

1.6.c Absolute Neutrino Mass Determination

Neutrino oscillations are only sensitive to the difference between the neutrino masses squared. Other studies had to be developed to measure the absolute neutrino masses. As neutrino interaction is driven by their flavor states, it is interesting to express the mass of a flavor state as a mixing of the mass eigenstates:

$$m^2(\nu_\alpha) = \sum_i |U_{\alpha i}|^2 m^2(\nu_i) \tag{1.46}$$

β decays provide a well controlled way of measuring the mass of the ν_e in a direct way. In the case where $m_{\nu_e} = 0$, the end point of the β spectrum corresponds to the E_0 value of the decaying nucleus. *A contrario*, if $m_{\nu_e} \neq 0$, the maximum energy of the electron is $E_0 - m_{\nu_e}$. The sketch on Figure 1.20 shows the principle of such a detection. The end point of the Tritium decay spectrum is brought down by $m_{\bar{\nu}_e}$ if the $\bar{\nu}_e$ is indeed massive.

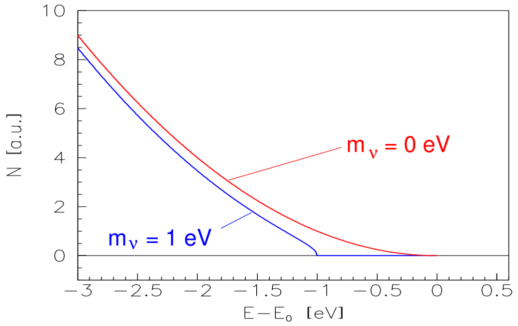


Figure 1.20: Sketch of the direct neutrino mass measurement with a β-decay of ^{3}H for $m_{\bar{\nu}_e} = 0\,$eV and $m_{\bar{\nu}_e} = 1\,$eV [42].

However, the expected neutrino mass is lower than 1 eV, so a nucleus with a small maximum energy E_0 is required to remain sensitive to a variation of this magnitude. In addition, the end point of the spectrum means a low count rate, and therefore, any background in that energy range become dominant; A short life-time is also an advantage to boost the count rate.

Tritium is a good candidate for such measurement, with a maximum energy of $E_0 = 18.6\,$keV, a simple electronic shell and a short life-time of 12.3 years. The current best precision on Tritium spectroscopy comes from the Troisk experiment with an upper limit at $m_{\bar{\nu}_e} < 2.12\,$eV [43]. The upcoming KATRIN experiment expects to reach a sensitivity of $m_{\bar{\nu}_e} < 0.35\,$eV [42].

1.6.d Neutrino-less Double-Beta Decays

Neutrino-less double-beta decays ($0\nu2\beta$) allow to probe the Dirac or Majorana nature of the neutrino. Some nuclei cannot undergo simple β-decay, but can undergo two simultaneous

β-decays:

$$^{A}_{Z}\mathrm{X} \rightarrow {}^{A}_{Z+2}\mathrm{Y} + 2e^- + 2\bar{\nu}_e \tag{1.47}$$

This reaction, however quite rare, is allowed in the $\mathcal{SM}$ as it does not violate the lepton number. Typical half-life for such decays are 10^{19} years.

In the hypothesis where neutrinos are Majorana particles, they are their own anti-particle, and therefore the two virtual neutrinos can annihilate with each other. The final observable state would then be:

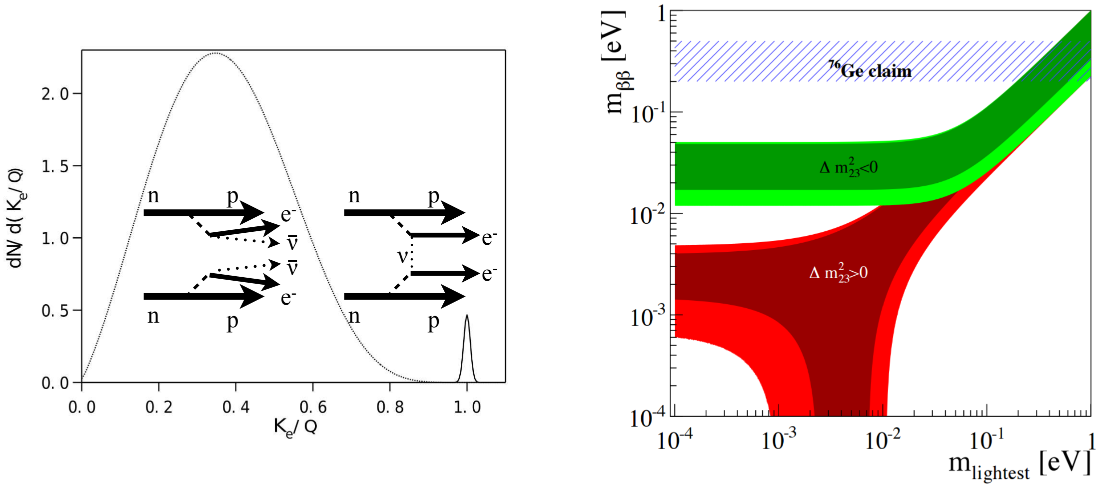
$$^{A}_{Z}\mathrm{X} \rightarrow {}^{A}_{Z+2}\mathrm{Y} + 2e^-. \tag{1.48}$$

This reaction is not allowed in the $\mathcal{SM}$ as it violates the lepton number by two units. In this reaction, the two electron carry out the full energy of the decay, and create a mono-energetic ray in the β energy spectrum as shown in the sketch in Figure 1.21(a). Observing such a decay would bring the demonstration that neutrinos are Majorana particles.

The $0\nu2\beta$ half-life is inversely proportional to $m_{\beta\beta}$, the effective Majorana mass:

$$m_{\beta\beta} = |\sum_i U_{ei}^2 m_i| \tag{1.49}$$

Figure 1.21(b) shows the correlation between $m_{\beta\beta}$ and the mass of the lightest mass state. The distribution behaves differently for the Inverted Hierarchy ($\Delta m_{31} < 0$) and the Normal Hierarchy ($\Delta m_{31} > 0$).



(a)

(b)

Figure 1.21: 1.21(a): Sketch of the $0\nu2\beta$ decay measurement principle [44]. In a $0\nu2\beta$ decay, the full energy is taken by the two electrons, therefore an accumulation of events appears at the end-point of the standard β-decay.
1.21(b): Distribution of $m_{\beta\beta}$ as a function of the mass of the lightest mass eigenstate for both hierarchies [45].

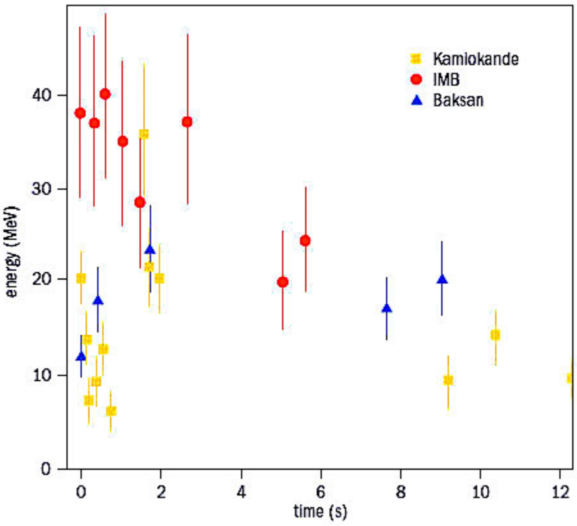


Figure 1.22: Coincident detection of 19 neutrino events over 13 s in the Kamiokande, IMB, and Baksan detectors from the explosion of the supernova SN1987A, consistent with the photon time of arrival on earth[46].

1.6.e Neutrinos as Astrophysical Messengers

Neutrinos are full of promises as messengers for astrophysical events. Indeed photons are absorbed by gas clouds, charged particles such as muons or protons are deflected by magnetic fields and therefore lose information on their source. Neutrinos, however, do none of the above, as they barely interact and the universe in average is a low-density material. As such, neutrino telescopes have been developed over the last decades (Antares, IceCube), and projects are still under development (ORCA, KM3NeT). Because of the high energy of cosmic neutrinos and the low rate, such detectors need to be very large, the volume of Antares is 0.1 km^3, and IceCube and KM3NeT are at the km^3 level.

Neutrinos are produced in numbers during the core collapse of a red giant to form a supernova, but well before the actual explosion. A neutrino burst was detected in 1987 during the explosion of the supernova SN1987A, 19 neutrino events, shown in Figure 1.22, were detected in three detectors worldwide in a 13 s time window, three hours before the arrival of the first photons. A supernova alert called SNEWS system is now in operation since 2004 with seven neutrino detectors: Borexino, Daya Bay, KamLAND, HALO, IceCube, LVD, and Super-Kamiokande.

1.6.f Neutrinos to probe the Earth

Closer to us, Neutrinos can bring information on the composition of the Earth. Geo-neutrinos are emitted from β-decays of daughter nuclei of ^{235}U and ^{232}Th decay chains. Experiments such as Borexino can, in addition to detecting of solar ν_e, detect the geo-neutrinos $\bar{\nu}_e$ through IBD with a clean signature as they are far from any nuclear reactor [47]. Figure 1.23 shows the $\bar{\nu}_e$ spectrum observed by Borexino. Such neutrinos are produced in the Earth's crust or in the mantle, providing insight on the rock compositions for depths at which drilling is impossible.

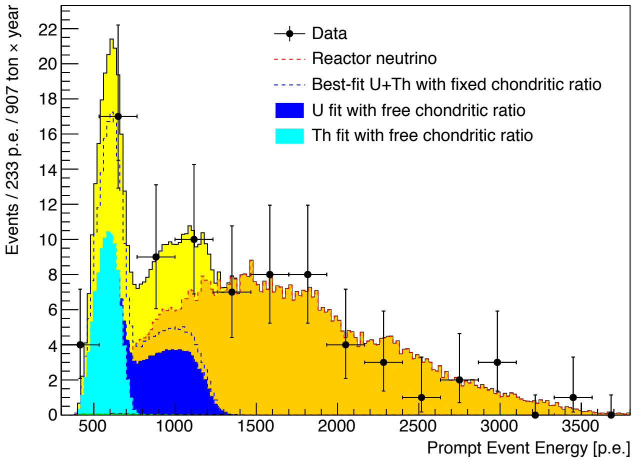
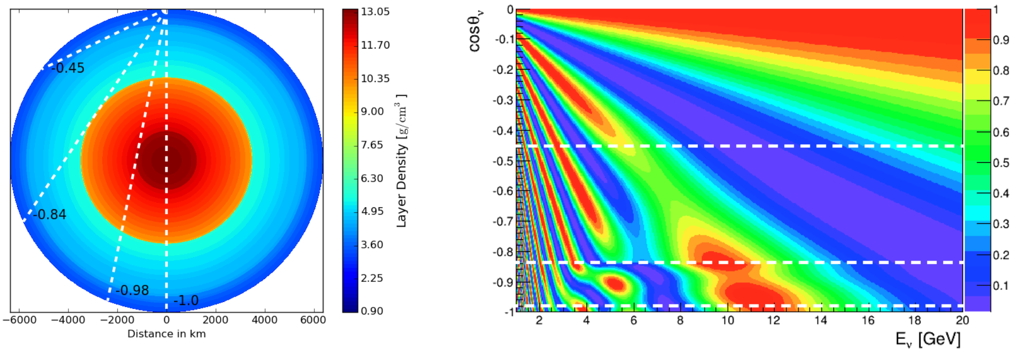


Figure 1.23: $\bar{\nu}_e$ spectrum observed by Borexino. The orange spectrum corresponds to reactor neutrinos from reactors outisde of Italy, the blue histograms correspond to the geo-neutrinos emitted by the ^{235}U and ^{232}Th chains[47].

Atmospheric neutrinos also allow to probe deeper into the Earth, their oscillation patterns are affected by matter effect, i.e. the electron density of each of the Earth's layers. Figure 1.24(a) shows a cross section of the earth with the different layers of density. The detector is placed at the top and looking at up-going neutrinos of angle θ_ν. Figure 1.24(b) shows expectations of the amplitude of oscillations observed as a function of the neutrino energy and the neutrino angle. This amplitude takes into account the densities and the path length in each layer. Discontinuities show sudden density change along the neutrino path, i.e. on lines in Figure 1.24(a).



(a)

(b)

Figure 1.24: Earth electron density tomography by ORCA [48]: 1.24(a): Cross section of the Earth with the density of each layer. The lines correspond to angles (here in values of $\cos\theta$) for which the density changes and a discontinuity is expected in the oscillation pattern.
1.24(b): Expected observed oscillation pattern as a function of the neutrino energy and the neutrino angle $\cos\theta_\nu$. Discontinuities show a sudden change in electron density, i.e. an interface between inner structures in the Earth.

1.7 Conclusion

Neutrinos were first discovered 60 years ago, but are still the center of many interrogations. It is not sure yet how to include them in a coherent $\mathcal{SM}$, whether as a Dirac or Majorana particle, and no theory can predict the smallness of their masses. However, huge progresses have been achieved in the past decade, with the oscillation paradigm finally becoming well understood and quantified. Neutrinos have also a lot of potential in long distance probes, at our own scale, to probe the deep Earth layers, or at cosmic scales, as privileged messenger of astrophysical events.

Chapter 2

Double Chooz

In this chapter will be presented the Double Chooz experiment, the detection of the $\bar{\nu}_e$ and prediction of the neutrino flux. Then, the detectors and the data acquisition chain, and finally, the event reconstruction will be described.

2.1 Description of the Experiment

Double Chooz is a neutrino experiment located at Chooz in the Ardennes region in France, close to the Belgian border (Figure 2.1(a)). Its goal is to measure the mixing angle θ_{13} by measuring the disappearance of $\bar{\nu}_e$ emitted by the two nuclear reactors of CHOOZ-B nuclear power plant, referred to as B1 and B2, and operated by the French company Électricité De France (Figure 2.1(b))

The fluxes and spectra of $\bar{\nu}_e$ are compared between two identical detectors, located respectively at $\sim 400\,\mathrm{m}$ (for the Near Detector — ND) and $\sim 1050\,\mathrm{m}$ (for the Far Detector — FD) from the middle point of the two reactors (see table 2.1).

Double Chooz is a two-phase experiment, the first phase of the experiment started in 2011 with the completion of the Far Detector. During this phase, a Monte Carlo simulation (MC) of the reactor cores was used to normalize the $\bar{\nu}_e$ flux and spectrum shape emitted by the cores. The results were limited by our understanding of the $\bar{\nu}_e$ emitted by the cores. The second phase began in 2015 with the completion of the Near Detector. The two-detector phase allows to greatly reduce the systematic uncertainty because of the correlation of the reactor flux between the two detectors and because of the cancellation of the detection systematics when comparing the Near data to the Far data.

Table 2.1: Experiment configuration, with distances of the detectors to each reactor and the respective overburden in meters of water equivalent.

Detector	Distance to B1 [m]	Distance to B2 [m]	Overburden [m.w.e.]
ND	468.761	355.388	120
FD	1114.656	997.839	300

(a) (b)

Figure 2.1: 2.1(a): The Chooz nuclear power plant is located in the French Ardennes region in the north of France, at the border with Belgium.
2.1(b): The two detector sites (red rectangles) are located at $\sim 400\,\mathrm{m}$ and $\sim 1050\,\mathrm{m}$ from the cores B1 and B2 (yellow circles).

Double Chooz was designed to improve the limit on θ_{13} set by the CHOOZ experiment, and make a measurement of θ_{13}. Most of the systematic uncertainties in CHOOZ came from the $\bar{\nu}_e$ flux prediction and the detection efficiency [26]. In order to reduce these uncertainties, Double Chooz proposed a two-detector concept.

2.2 Detection Process of $\bar{\nu}_e$ in Double Chooz

Double Chooz uses the same $\bar{\nu}_e$ detection principle as Reines and Cowan in the historic Savannah River experiment (See Section 1.1). The $\bar{\nu}_e$ are detected through Inverse β-Decay (IBD) on free protons in a Gadolinium doped organic liquid scintillator.

As previously discussed in section 1.2, neutrinos can only interact through charged current (CC — $W^{\pm}$ exchange) or neutral current (NC — Z^0 exchange).
Elastic scattering for $\bar{\nu}_e$ on free protons, from NC interaction, would imply detecting a recoil of a few keV, well below the sensitivity of the detectors. Elastic scattering on an electron arises from both NC and CC and would create an electron recoil of a few MeV. This signal could be detected, but offer no particular pattern to trigger on, and a single signal of a few MeV would be indistinguishable from ambient radioactivity.
On the other hand, CC of $\bar{\nu}_e$ on free protons, also called Inverse β-Decay, emits a positron and a neutron:

$$\bar{\nu}_e + p \to e^+ + n \tag{2.1}$$

$$n + \mathrm{Gd} \to \mathrm{Gd}^* \to \mathrm{Gd} + \gamma\ (\Sigma\gamma = 8\,\mathrm{MeV}) \tag{2.2}$$

$$n + \mathrm{H} \to \mathrm{H}^* \to \mathrm{H} + \gamma\ (2.2\,\mathrm{MeV}) \tag{2.3}$$

The positron scintillates and annihilates with an electron in the medium, providing an instantaneous signal (prompt signal). The neutron thermalizes through multiple scatterings and captures on a Gadolinium nucleus (Gd). The neutron capture on Gd emits multiple γ for an average total energy of $\sim 8\,\text{MeV}$, well above natural radioactivity energy range (Figure 2.2(a)). The neutron can also capture on Hydrogen (H) and a emit a single γ of $2.2\,\text{MeV}$. The neutron capture, with a characteristic capture time τ_{cpt}, is the delayed signal. The coincidence of a prompt and a delayed signal defines an IBD candidate.
The energy of neutron capture is independent from the $\bar{\nu}_e$'s energy, therefore, it is possible to select n-H or n-Gd captures without impact on the observed $\bar{\nu}_e$ energy.

The n-Gd capture was originally the only one to be used, but the n-H capture was investigated in [49, 50] and provides an independent event sample from the n-Gd capture.

The $\bar{\nu}_e$ energy threshold of the IBD on a free proton is:

$$E_{\text{threshold}} = \frac{(m_e + m_n)^2 - m_p^2}{2m_p} \simeq 1.806\,\text{MeV} \tag{2.4}$$

where m_p, m_n and m_e are respectively the masses of the proton, neutron and positron. The neutron carries only a small kinetic energy, then the positron energy is related to the $\bar{\nu}_e$ energy by:

$$E_{\bar{\nu}_e} = \frac{1}{2}\,\frac{2m_p\,E_{e^+} + m_n^2 - m_p^2 - m_e^2}{m_p - E_{e^+} + \cos\theta\sqrt{E_{e^+}^2 - m_e^2}} \tag{2.5}$$

$$E_{\bar{\nu}_e} \simeq E_{e^+} + \Delta \tag{2.6}$$

where E_{e^+} is the positron energy, θ is the angle of the positron with respect to the direction of the incident $\bar{\nu}_e$ and $\Delta = m_n - m_p = 1.293\,\text{MeV}$.

The visible signal in the detector that carries the neutrino information is the scintillation of the positron and its annihilation with an electron in the medium, so the visible energy is the energy of the positron plus the mass of an electron at rest:

$$E_{vis} = E_{e^+} + m_e = E_{\bar{\nu}_e} - \Delta + m_e = E_{\bar{\nu}_e} - 0.782\,\text{MeV} \tag{2.7}$$

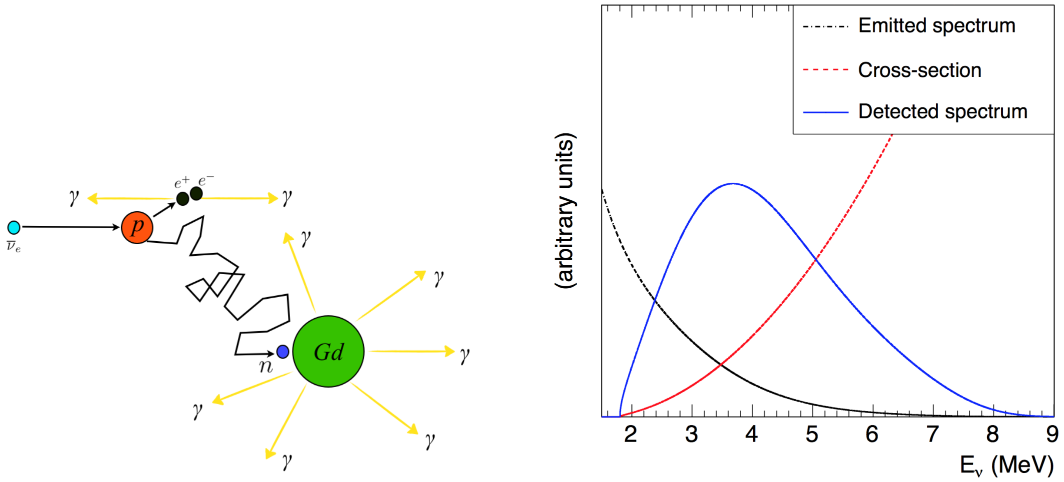
The cross section of the IBD can be described at low energies ($E_{\bar{\nu}_e} \lesssim 60\,\text{MeV}$) by Vogel et al. [51]:

$$\sigma(E_{\bar{\nu}_e}) = K \times (E_{\bar{\nu}_e} - \Delta)\sqrt{(E_{\bar{\nu}_e} - \Delta)^2 - m_e^2} \tag{2.8}$$

$$\text{where } K = \frac{2\pi^2}{m_e^5 f \tau_n} = (9.61 \pm 0.02) \times 10^{-44}\text{cm}^2\text{MeV}^{-2} \tag{2.9}$$

with τ_n is the neutron lifetime [52] and f is a correction factor taking into account the Coulomb term, weak magnetism and external radiative recoil corrections.

Figure 2.2(b) shows a sketch of the typical expected $\bar{\nu}_e$ spectrum in absence of oscillations (solid blue line) as a product of the emitted spectrum from the reactor cores (dotted black line — see Section 2.3) and the IBD cross section (dotted red line).



(a) (b)

Figure 2.2: 2.2(a): $\bar{\nu}_e$ interaction through IBD with neutron capture on Gadolinium.
2.2(b): Sketch of the expected observed anti-neutrino spectrum in the no-oscillation hypothesis (solid blue line), taking into account the emitted spectrum from the reactor cores (dotted black line) and the IBD cross-section (dotted red line) [53].

Successive Analyses Double Chooz released various analyses, using samples with neutron captures on both Gadolinium and Hydrogen and developing new techniques always increase the sensitivity to θ_{13}. Table 2.2 gives an overview of the successive progress stages of the collaboration.

2.3 Neutrino production by the power plant

The two 2nd generation reactors of CHOOZ-B are the most powerful commercial reactors in the world, each with a nominal thermal power of 4.25 $\mathrm{GW_{th}}$ and an electrical efficiency of $\sim 34\%$. The two reactors are Pressurized Water Reactors (PWR) of type N4 and each emits a pure $\bar{\nu}_e$ flux of $\sim 8 \times 10^{20}\ \bar{\nu}_e \cdot \mathrm{s}^{-1}$ at nominal power.

Nuclear reactors rely on the control and continuity of chained fissions in a critical state. Neutrons produced during a fission interact with other nuclei and in turn produce a fission and other neutrons. The core is kept in a critical state, such that the neutron production from fissions balances exactly the neutron loss by absorption or leakage[1].

Choozreactors are fuelled with Uranium oxide ($\mathrm{UO_2}$) enriched at the levels of 1.8 to 4% in ^{235}U. Fresh fuel is composed primarily of ^{235}U and ^{238}U. As the fuel gets depleted, other fissile isotopes such as ^{239}Pu and ^{241}Pu appear by successive neutron captures on fertile isotope ^{238}U:

$$\mathbf{{}^{238}_{92}U} \xrightarrow{(n,\gamma)} {}^{239}_{92}\mathrm{U} \xrightarrow[23.45\mathrm{min}]{\beta^-} {}^{239}_{93}\mathrm{Np} \xrightarrow[2.35\mathrm{d}]{\beta^-} \mathbf{{}^{239}_{94}Pu} \xrightarrow{(n,\gamma)} {}^{240}_{94}\mathrm{Pu} \xrightarrow{(n,\gamma)} \mathbf{{}^{241}_{94}Pu} \tag{2.10}$$

---

[1] A non critical reactor would see its neutron production increase or die out exponentially, whether they are operated respectively above or below the critical threshold.

Table 2.2: Time line of Double Chooz releases with the major specificities of each new analysis.

Name	Year	Live-time (days)	Specificity
DC-I n-Gd	2012	96.8	World first indication of a non-zero θ_{13} by a reactor-based experiment [7]
DC-II n-Gd	2012	227.93	Increased statistics, improved background constraints [54]
DC-II n-H	2012	240.1	Proof of concept for n-H analysis [49]
DC-II RRM	2014	233.93 (n-Gd) 246.4 (n-H)	Background independent measurement of θ_{13} [55]
DC-III n-Gd	2014	467.90	First report of a distortion on the spectral shape [56]
DC-III n-H	2015	462.72	n-H achieves n-Gd equivalent sensitivity [57, 50]
DC-IV n-Gd	2016	468.17 (FD-I) 212.21 (FD-II) 150.76 (ND)	First Double Chooz results with both detectors [58]
DC-IV Gd++	2016	455.21 (FD-I) 362.97 (FD-II) 257.96 (ND)	Combined use of n-Gd and n-H samples

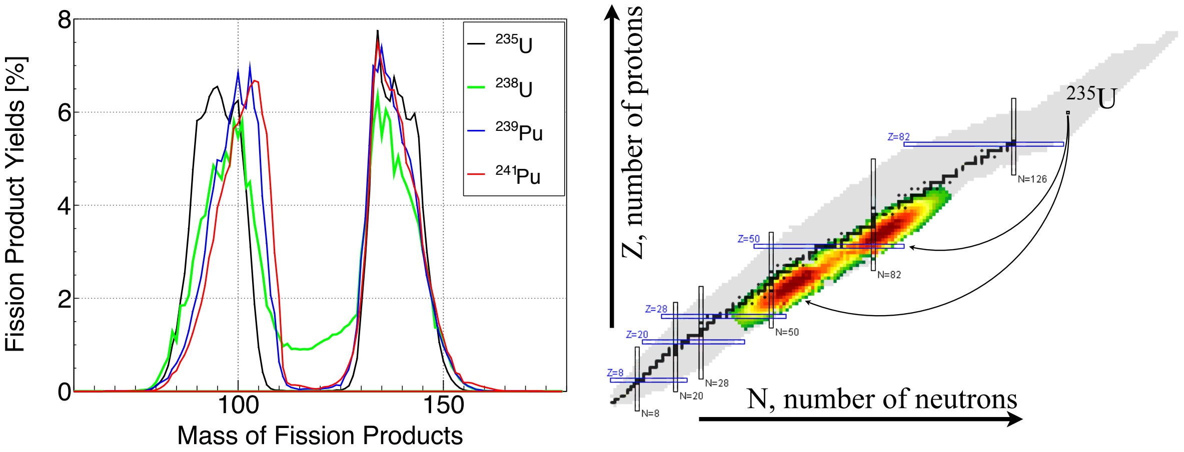
Every year, a third of the fuel is being replaced, leading to reactor off time, which is very useful in Double Chooz's θ_{13} analyses.

The fission cross section of Uranium 235 (the dominant fissile isotope in the cores) is high for neutrons in the thermal range (a few meV) whereas the fission neutrons have an energy ranging from 1 to 10 MeV, with a mean energy of 2 MeV (fast neutrons). A means to thermalize the fast neutrons is then required to keep the chain reaction, this process is called moderation. The fuel is plunged in water, kept in liquid state by a pressure of 155 bar, hence the PWR denomination. Thermalization is achieved by scattering neutrons on the Hydrogen nuclei.

The fission rate, and thus thermal power can be coarsely modulated by regulating the concentration of boron in the primary water loop. The capture of neutron on boron has a high cross section at thermal energies, this process competes with fission and effectively lowers the fission rate. The fission rate can be more finely tuned by inserting control rods of neutron absorbers in the core.

The four dominant nuclei contributing to the thermal power and thus the $\bar{\nu}_e$ flux are ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu. They make up for more than 99.7% of the thermal power of the reactor [54]. A fission releases energy and generates multiple fast neutrons and daughter nuclei called fission products (Figure 2.3(a)). Most of these fission products are rich in neutrons, and will undergo β^--decays to migrate to the valley of stability (Figure 2.3(b)).

The several β-decays following each fission each emit one $\bar{\nu}_e$. Each fissile isotope can thus



(a)

(b)

Figure 2.3: 2.3(a): Fission product yields versus the fission product mass, in atomic units (i.e. number of nucleons), for thermal neutron reactors [59].
2.3(b): Fission product distribution of ^{235}U represented in the isotope table (original figure from [60]).

be associated to an average number of emitted $\bar{\nu}_e$. In average ~ 6 $\bar{\nu}_e$ are emitted per fission. However, $\sim 2/3$ of those $\bar{\nu}_e$ have an energy below the IBD threshold. Table 2.3 gives the mean number of $\bar{\nu}_e$ per fission with an energy above this threshold. However, the fuel is in constant evolution, and so is the $\bar{\nu}_e$ flux and shape. A simulation of the cores is needed to provide a prediction of the un-oscillated $\bar{\nu}_e$ flux and spectrum from the reactors (Figure 2.4(b)) [53].

The expected instantaneous $\bar{\nu}_e$ rate in each detector, with no oscillation ($\theta_{13} = 0$) can be expressed as:

$$N_\nu^{exp}(s^{-1}) = \frac{1}{4\pi} N_p \epsilon \sum_{r=B_1,B_2} \frac{1}{L_r^2} \frac{P_{th,r}}{\langle E_f \rangle_r} \langle \sigma_f \rangle_r \tag{2.11}$$

where N_p is the number of available protons in the detector, ϵ is the detection efficiency, r refers to the reactors B_1 and B_2, L_r is the distance between the considered detector and each reactor r, $P_{th,r}$ is the instantaneous thermal power of the reactor r, $\langle E_f \rangle_r$ is the mean energy released per fission in the reactor r ($\sim 200\,\text{MeV}$), and $\langle \sigma_f \rangle_r$ is the mean cross section per fission in the reactor r.

The mean energy released per fission $\langle E_f \rangle_r$ in a reactor is expressed as a linear combination of the mean energies per fission of the various isotopes contributing to the fission:

$$\langle E_f \rangle_r = \sum_k \alpha_k^r E_{f,k} \tag{2.12}$$

where α_k is the instantaneous fission fraction of the isotope k (with $k = {}^{235}\text{U}, {}^{239}\text{Pu}, {}^{238}\text{U}$, and ${}^{241}\text{Pu}$).

The mean cross section per fission can be computed as:

$$\langle \sigma_f \rangle = \sum_k \alpha_k \langle \sigma_f \rangle_k = \sum_k \alpha_k \int_0^\infty S_k(E) \sigma_{\mathrm{IBD}}(E) dE \tag{2.13}$$

where $S_k(E)$ is the $\bar{\nu}_e$ spectrum per fission of the isotope k, and $\sigma_{\mathrm{IBD}}(E)$ is the IBD cross section in the detector (see Eq. 2.8).

In order to reduce the uncertainties from the $\bar{\nu}_e$ reference spectra $S_k(E)$, currently estimated to be of the order of 2.5% (Table 2.4), Double Chooz chose to normalize the mean cross section per fission using the Bugey4 measurement [40]. The Bugey4 experiment performed the most precise measurement of the mean cross section per fission for a PWR reactor. This measurement was conducted at 15 m from a PWR, found a value of $5.752 \pm 0.08 \cdot 10^{-19}$ barn/fission with a relative uncertainty of 1.4% (1σ C.L.). The normalization is done with:

$$\langle \sigma_f \rangle_r = \langle \sigma_f \rangle^{Bugey} + \sum_k (\alpha_k^r - \alpha_k^{Bugey}) \langle \sigma_f \rangle_k \tag{2.14}$$

The second term takes into account the differences in the compositions of Bugey and Chooz reactor fuels. As previously stated, the fuel composition evolves in time, so does the fission fraction of each isotope. It is then crucial to have a good understanding of the evolution of the composition of the cores. Figure 2.4(a) shows the evolution of the fission fractions of the four dominant isotopes over the data taking period of DC-III [61]. The parts without data correspond to reactor-off time during fuel reprocessing. The mean α_k, averaged over the same period, can be found in Table 2.3, along with the mean energy per fission per isotope. α_k in Double Chooz are estimated using the reactor simulation MURE [62, 63].An other code , called DRAGON [64], was used to cross check MURE results on dedicated sensitivity studies.

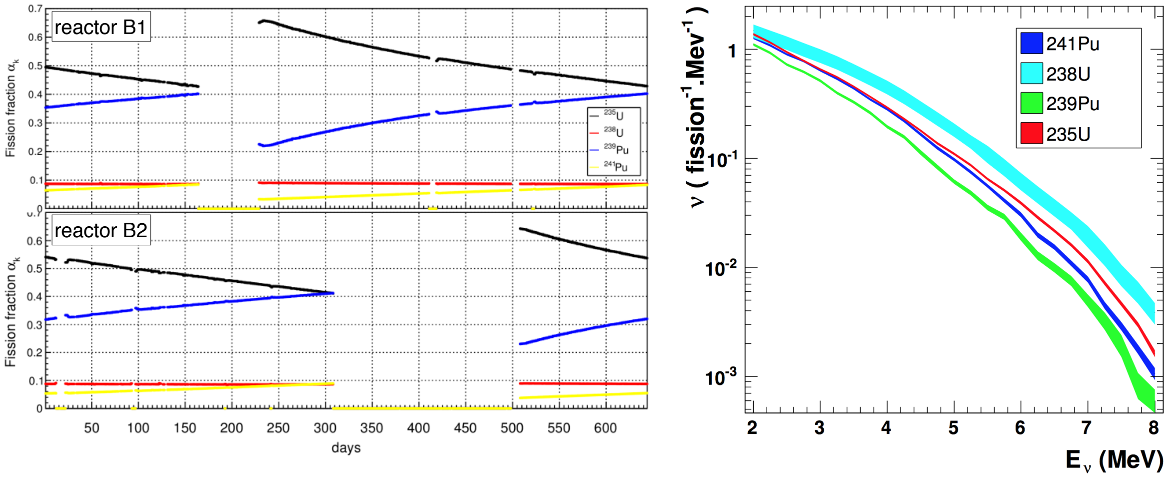
Spectra in Double Chooz reactor cores simulations originate from studies of the β spectrum of the fission products of the main isotopes (^{235}U, ^{239}Pu, and ^{241}Pu) at the ILL research reactor [65]. Figure 2.4(b) shows the $\bar{\nu}_e$ spectra per fission for the four major isotopes from the ILL measurements. The sum of these spectra, weighed by their corresponding α_k, corresponds to the black line in Figure 2.2(b) — this is the $\bar{\nu}_e$ spectrum emitted by the reactors.

The measurement of the instantaneous thermal power $P_{th,r}$ is derived from the in-core instrumentation, providing the temperature, the pressure and the mass flow of the water in the primary and secondary loops. The in-core instrumentation calibration is tested weekly using the heat balance in the secondary loop, which is heated by the primary loop with the water heated by fissions. In the secondary loop, steam is generated to drive turbines. By using measurements of the heat flow in the secondary loop, the thermal power can be measured. This test is performed with the reactor running at full power. The in-core instrumentation is recalibrated if it deviates by more than the uncertainty in the heat balance measurement. At nominal power the final uncertainty is 0.5% (1σ C.L.). The instantaneous thermal power is provided by EDF with time step < 1 min.

Table 2.4 addresses the contribution of the various components of the $\bar{\nu}_e$ predictions to the overall uncertainty on this prediction, in the Far Detector only configuration. Using the measurement of the mean cross section per fission from the Bugey4 experiment helps reduce

Table 2.3: Fission fraction averaged over the data taking period of DC-III [61], mean energy released per fission [66], mean neutrino energy and mean number of neutrinos per fission, with the 1.8 MeV threshold of neutrino detection through IBD for isotopes with the dominant contributions to the core's thermal power [67, 68].

Isotope	$\langle \boldsymbol{\alpha} \rangle$	$\langle \mathbf{E}_f \rangle$ **(MeV)**	$\langle \mathbf{E}_{\bar{\nu}_e} \rangle$ **(MeV)**	$\langle \bar{\boldsymbol{\nu}}_{e,f} \rangle$
^{235}U	0.496 ± 0.016	201.92 ± 0.46	5.226 ± 0.051	$1.92(1 \pm 0.019)$
^{238}U	0.087 ± 0.006	209.99 ± 0.60	7.040 ± 0.326	$2.38(1 \pm 0.020)$
^{239}Pu	0.351 ± 0.013	205.52 ± 0.96	3.799 ± 0.033	$1.45(1 \pm 0.021)$
^{241}Pu	0.066 ± 0.007	213.60 ± 0.65	4.956 ± 0.049	$1.83(1 \pm 0.019)$



(a) (b)

Figure 2.4: 2.4(a): Evolution of the fission fractions of the dominant isotopes over time for the two reactors since the first day of data taking for the DC-III analysis. The parts with no data correspond to reactor-off time, before fuel reprocessing.
2.4(b): Spectrum of anti-neutrinos emitted by isotope.

the overall uncertainty by ∼ 45%.
With the commissioning of the Near Detector, a new measurement of the cross section per fission is under way and is expected to have an uncertainty lower than the measurement performed by the Bugey4 experiment within 3 years of data taking.

2.4 Detector Design

Figure 2.5 shows the schematic organization of Double Chooz detectors.

Each detector is formed from a series of nested cylinders with each volume filled with different

Table 2.4: Uncertainties in $\bar{\nu}_e$ prediction in the Far Detector only configuration. The dominant uncertainty is the mean cross section per fission and can be greatly improved by the use of Bugey4 cross section measurement [69].

Source	**Uncertainty without Bugey4 [%]**	**Uncertainty with Bugey4 [%]**
P_{th}	0.5	0.5
$\langle\sigma_f\rangle^{\text{Bugey4}}$	—	1.4
$S_k(E)\sigma_{\text{IBD}}(E_\nu^{\text{true}})$	2.5	0.2
$\langle E_f\rangle$	0.2	0.2
L_R	< 0.1	< 0.1
α_k^R	0.8	0.9
Total	2.8	1.8

liquids; insensitive Buffer oil for shielding, Gd-doped scintillator as the Target and undoped scintillators for gamma rays, fast neutrons and muons.

The two inner vessels are acrylic and transparent to photons above 400 nm. The innermost vessel is the **Target**, with a diameter of 2.3 m, which contains 10 m^3 of Gadolinium-doped scintillator ($\sim$ 8.3 t).The scintillator contains 1 g/l of Gadolinium. The capture on Gadolinium represents more than 80% of neutron captures in the Target.

The scintillator of the Target is composed of 80% dodecane ($C_{12}H_{26}$), 20% PXE (C_6H_{18}), 7 g/l of PPO ($C_{15}H_{11}NO$), 20 g/l of bis-MSB ($C_{24}H_{22}$), and of 1 g/l of Gadolinium. Time stability for the Gadolinium-doped liquid is ensured by Gadolinium atoms encapsulation in a Gd(thd)$_3$[2] molecule. It ensures an efficient dissolving in the scintillator and a durable transparency [70].

The volume of the Target is accessible from the top of the detector by a Glove Box and a Chimney. A nitrogen atmosphere is maintained in the glove box to prevent a degradation of the scintillator by oxygen.

Enclosing the Target is the **γ-Catcher** volume, with a diameter of 3.4 m, which contains 22 m^3 of undoped scintillator. The purpose of this volume is to detect the gammas emitted in both the neutron-capture process and positron annihilation in the Target, such that gammas emitted from neutrino events occurring in the outer volume of the Target are detected. This results in a well-defined Target volume. The γ-Catcher liquid scintillator is composed of 30% dodecane, 66% mineral oil (Shell Ondina909), 4% PXE, 2 g/l of PPO and 20 mg/l of bis-MSB. This scintillator composition was tuned to match the density and light yield of the Target liquid, in order to assure the safety of acrylic vessels and increase the uniformity of the detector response.

Since the photomultipliers are the most radioactive component of the detector, the inner volumes are shielded by a **Buffer** volume, with a diameter of 5.5 m, filled with non-scintillating paraffin oil. Events occurring in the acrylic volumes are detected by 390 10-inch low background

---

[2] Gd(III)-tris-(2,2,6,6-tetramethyl-heptane-3,5-dionate)

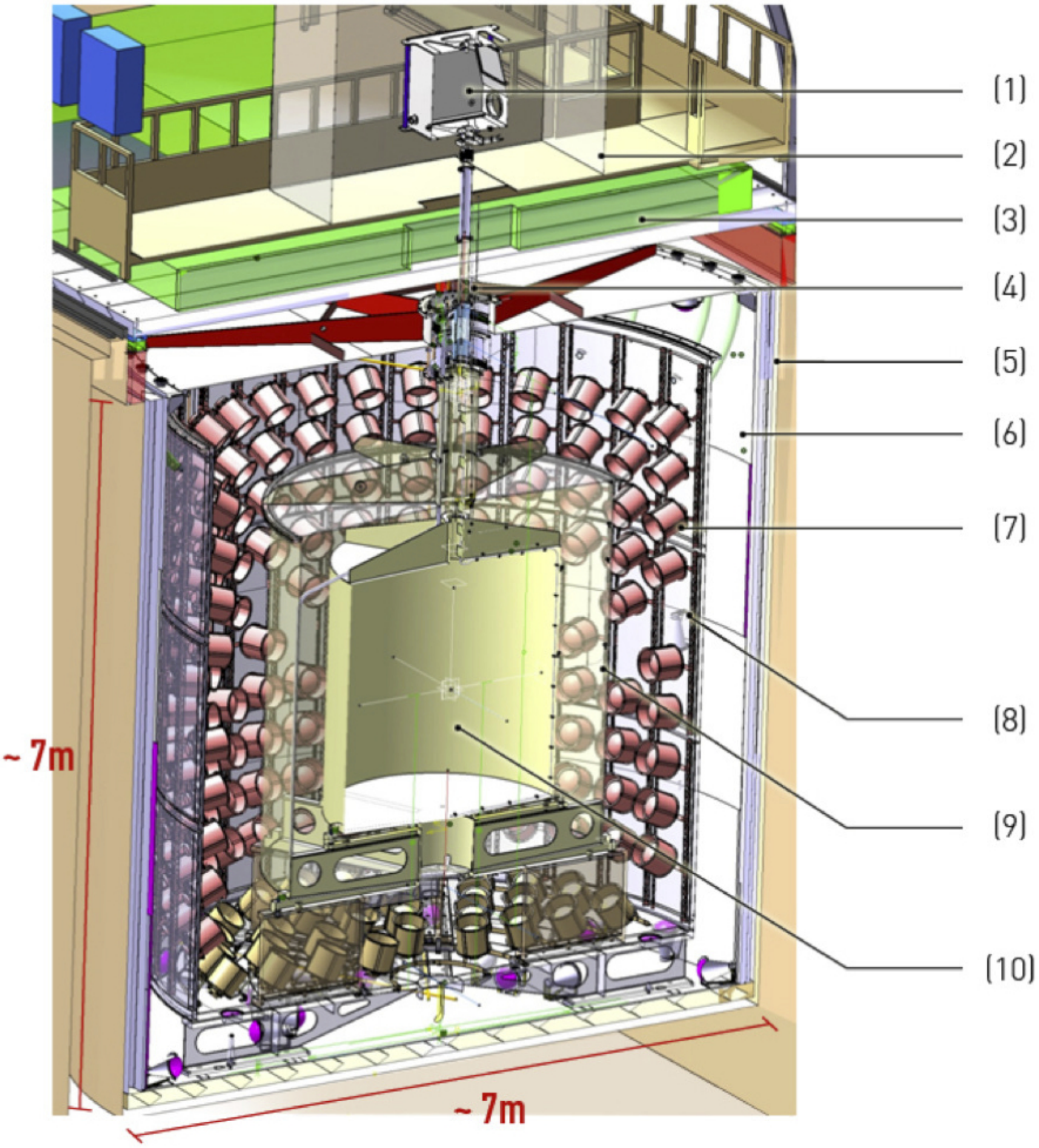


Figure 2.5: Double Chooz detector architecture.
(1) Glove Box (GB),
(2) Clean Tent,
(3) Outer Veto (OV),
(4) Chimney,
(5) Shielding,
(6) Inner Veto (IV),
(7) Inner Detector Photomultiplier (PMTs),
(8) Buffer,
(9) γ-Catcher (GC),
(10) Neutrino Target (NT).

photomultiplier tubes (Hamamatsu R7081, ID-PMT) fixed to the inside of the steel Buffer tank. The photomultiplier tubes are angled to improve the uniformity of light collection efficiency in the inner-most volumes.

The outer detector volume is steel walled, with a diameter of 6.6 m, and filled with scintillator. 78 8-inch photomultipliers (Hamamatsu R1408, IV-PMT) line the outermost wall which is painted with a reflective white coating. This volume is the **Inner Veto** with the purpose of detecting and tracking muons and fast neutrons.

On top of the detector sits the **Outer Veto** (OV). This comprises strips of plastic scintillator and wavelength-shifting fibres. The veto extends further than the detector diameter with the purpose of detecting and tracking muons. The precision of the entry point of a muon, X-Y position, is far more precise than that achieved by the Inner Veto and detector. One of the main objectives is to tag near-miss muons, which interact in the surrounding rock (and not in the detector) but produce fast neutrons. Another important goal is to determine whether a muon entered the inner detector. Muons that do so can produce cosmogenic isotopes (i.e. via a photonuclear interaction on ^{12}C), some of which will produce backgrounds for the experiment (discussed in chapter 6).

Figures 2.6 shows the two Inner Detectors for the far and Near Detectors. The acrylic vessels of the Target and the γ-Catcher are visible as well as the PMTs on the wall of the Buffer. Both detectors are identical from the Buffer tank (inner-most stainless steel vessel) inwards which is a physics requirement. This (Target, γ-Catcher,Buffer) volume is called the **Inner Detector**. Shielding against the radioactivity of the rock is provided by 15 cm of demagnetised steel for the Far Detector and a for the Near Detector, where the detector size is less of an issue and the fast neutron is higher, a 1 m-thick water shielding is used.

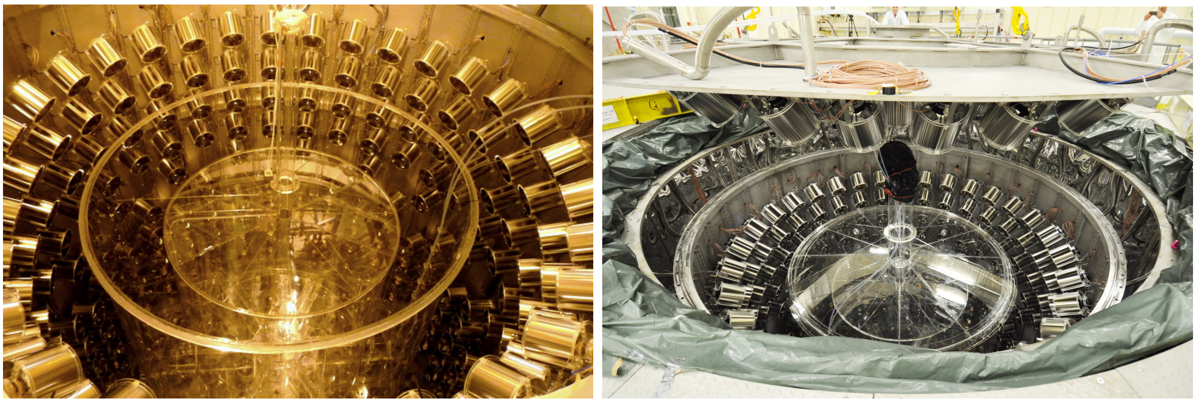


Figure 2.6: Inside of the Inner Detectors of the Far Detector (left) and the Near Detector(right). In both detectors, the Target's acrylic vessel can be seen in the center, encapsulated in the γ-Catcher's acrylic vessel, and surrounded by the PMTs in the Buffer.

2.5 Calibration

Double Chooz relies on the energy deposition of events across the detector in order to build the $\bar{\nu}_e$ energy spectra which will then allow to access θ_{13}. It is essential to master the response of the detector to physical processes considered. Liquid and PMTs properties can be better characterized, even if tests were already performed during the detector installation.

Calibrating the detector is therefore crucial in order to access the energy response and to carefully reconstruct the events vertices. To do so, radioactive sources as well as light injection system are used.

Light Injection systems

Light injection system emitting LED light guided through optical fibers were installed in both the inner detector (IDLI) and the inner veto (IVLI). 46 injection points fixed on the edge of some PMTs are present in the inner detector and 96 in the inner veto. One can set the wavelength of the light to 385 nm, 425 nm or 475 nm for the IDLI and 365 nm or 475 nm for the IVLI. The system allows either a diffuse or pencil beam mode.

Whereas the 385 nm wavelength should be totally absorbed and re-emitted by the liquid scintillator, the 425 nm wavelength is partially absorbed, constraining its detection to the PMTs present on the opposite side of the detector. It allows to understand the absorption of the Target and γ-Catcher as well as the PMTs characteristics. The 475 nm wavelength should not excite the liquid scintillator, acting as a direct light source for the PMTs receiving the light.

Radioactive sources

Up to now, four radioactive sources with different particles emission and different energies are used to calibrate the ID. The deployment of ^{137}Cs, ^{68}Ge, ^{60}Co or ^{252}Cf can be performed either using the Z-Axis Deployment System (ZADS), which brings the sources directly into the Target volume, or the Guide Tube, which brings the sources into the γ-Catcher volume within a rigid loop tube using a motor driven cable. The deployment of the sources is performed through the Glove Box in a clean tent above the detector. When using the ZADS, operators use the Glove Box. ZADS allows to move the sources on a range spanning from 1 mm above the Target bottom up to the Target chimney, with 1 mm precision, and the Guide Tube allows a precision of 2 mm. With the emission of a 667 keV γ, the ^{137}Cs source is the less energetic source. This energy is not of interest for the $\bar{\nu}_e$ selection but remains between the trigger threshold and the IBD reaction threshold of 1.022 MeV according to equation 2.7. The electron capture on ^{68}Ge provides a positron which annihilates with an electron of the medium, leading to the emission of two 511 keV γ. This energy corresponds to the energy threshold of the IBD and helps determining the trigger efficiency. The ^{60}Co source emits two γ of 1.17 MeV and 1.33 MeV in the prompt energy range of interest for the $\bar{\nu}_e$ selection whereas the delayed energy range can be studied thanks to the neutrons release from the spontaneous fission of the ^{252}Cf source.

2.6 Data Acquisition

2.6.a Inner Detector and Inner Veto read-out chains

The acquisition chain can be summarized in Figure 2.7 for the Inner Detector and the Inner Veto. The Outer Veto is an independent system. Far Detector electronic systems were upgraded during the first semester of 2014. The two configurations are respectively called FD-I and FD-II for pre- and post-upgrade.

The light emission from an event in the scintillator is collected by a PMT, operated at a gain of $\sim 10^7$. The electric signal from the PMT, called PMT waveform, is carried out of the

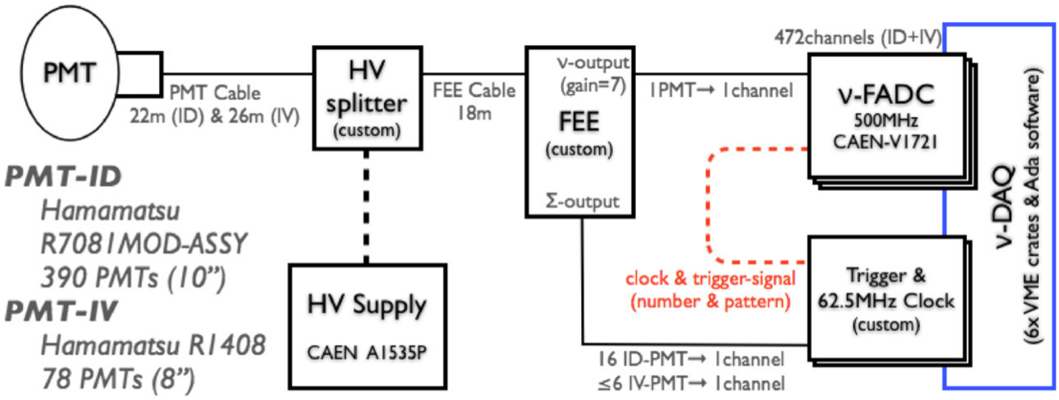


Figure 2.7: Read-out and acquisition chain [54].

detector by the same high voltage cable that powers the PMT.
The separation between low frequency DC high voltage and high frequency PMT waveform is done in the splitter. The waveform is then sent to a custom Front End Electronics that amplifies this signal with a gain of ~ 7 for FD-I and ~ 20 for FD-II.
The amplified analog signal is afterwards digitized by a 8-channel, 8-bit, 500MS/s Flash-ADC (CAEN Vx1721) described later on in Chapter 3.
The FEE also sums up the non amplified analog signals by groups of 16 ID-PMTs, or 5 IV-PMTs, and sends it to the trigger system. Upon discrimination, a trigger signal is simultaneously sent to all Flash-ADCs and the digitized waveforms are stored by the data acquisition system (DAQ).

Later on, a *channel* will refer to the full acquisition chain, from one PMT to the digitized waveform.

2.6.b Outer Veto read-out chain

The Outer Veto read-out system is independent from the Inner Detector and the Inner Veto read-out. The scintillation light coming from the Outer Veto strips is collected by optical fibers. These fibers are connected at one end to a Hamamatsu M64 multi-pixel-PMT. A mirror is positioned at the other end in order to reflect the light.

The OV PMT is connected to a custom FEE board allowing adjustment for each channel gain. The signals are digitized by a 12-bit ADC. If the signals exceed a common threshold, the trigger is generated, and the signals are stored and later read-out via USB.

2.6.c Trigger System

The trigger system is separated in one Trigger Master Board (TMB) and three sub-trigger boards, one for the Inner Veto (TB-IV), and two for the Inner Detector (TB-A and TB-B).

The ID PMTs are divided in 12 geometric sectors of 32 PMTs each. For a given sector, half of the PMTs are connected to the TB-A and the other half to the TB-B. Each PMT connected to the TB-A is surrounded by PMTs connected to the TB-B and vice versa. Therefore, each trigger board looks at the same volume [71].

The IV PMTs are also divided in groups of 5 PMTs. A precise measurement of the deposited energy inside the IV is difficult to obtain, given the low PMTs coverage of the Inner Veto. Thus the Inner Veto trigger condition is rather based on the hit pattern of the PMTs groups. The lowest threshold (neutron) uses multiplicity, requiring 10 groups to exceed the group threshold. The second threshold (muon) is a total sum threshold, is slightly slower than than the multiplicity as more processing is required.

The three trigger boards convert the summed signal of the PMTs (ID) or the PMT hit pattern (IV) into a series of 8-bit sent to the TMB.

Each board has different trigger levels to better characterize the event. TB-A and TB-B have a low energy threshold to characterize a neutrino interaction, an intermediate energy threshold for fast neutrons and a high energy threshold for muons crossing the detector. TB-IV has two levels, one for neutrons and one for muons.
When a muon threshold is crossed in the IV, only low level reconstructed information, computed at the Event Builder's level, are stored, not the complete waveforms of the PMTs.
In turn, the TMB sends out a 8-bit trigger word to the Flash-ADC. This trigger word contains information on the trigger level of the event.

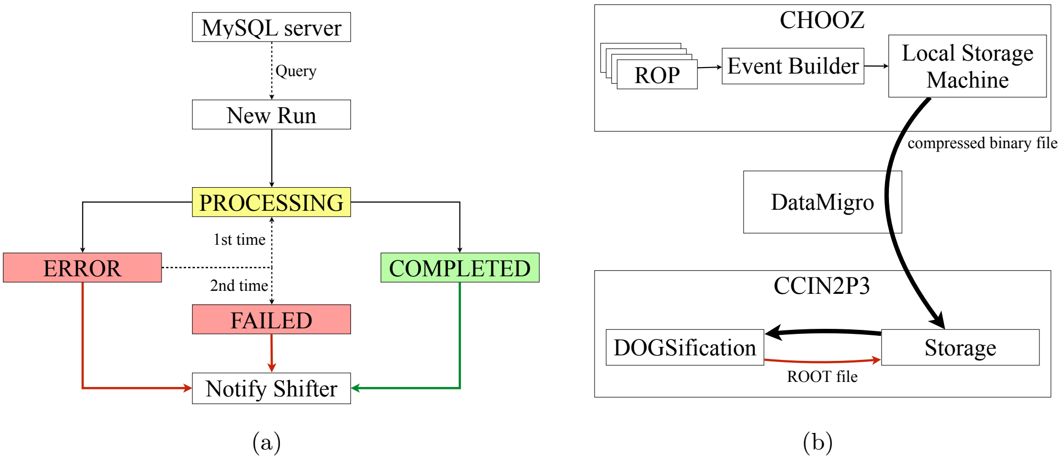
Other systems communicate with the TMB: a fixed rate trigger is sent every second to determine the detector live time, the light injection systems (IDLI and IVLI) provide the LED firing time to the TMB and the Outer Veto provides a flag every time it is triggered.
The TMB also broadcasts a synchronous 62.5 MHz (16 ns period) clock to all subsystems.

2.6.d Data Acquisition

The Double Chooz data acquisition is performed by five read-out processors (ROP) connected to the Flash-ADC cards. When the trigger signal is received, the ROPs transfer the Flash-ADC memory Buffer to the Event Builder Process via a 1 Gb/s ethernet connection using a TCP/IP protocol. The Event Builder is a multi-thread process which collects the data from the five ROPs and writes them on consecutive gzip binary files. The data acquisition is controlled by a Run Control and its Graphical User Interface (GUI). This GUI is the interface between the shifter and the data acquisition system. Shifters can configure the run parameters and start or stop the data acquisition. For each run, the timing information such as the start and stop time, the run length and the run-type are written in a database.

2.7 Datafile handling with DataMigro

Binary data files from the Event Builder are written compressed in gzip format and can be temporarily stored on local machines before being transferred to IN2P3 computing center (CCIN2P3) located at Lyon for processing, back-up and storage. Once at CCIN2P3, binary files are processed in a dedicated file format of the Double Chooz collaboration, based on ROOT (DOGS files) by a process called "DOGSifier". Onsite data handling and data migration to CCIN2P3 are performed by a set of processes called “DataMigro”. The maintenance of this process was one of my responsibilities during my thesis.



(a)

(b)

Figure 2.8: Schematic view of DataMigro process structure.
(a) DataMigro process structure.
(b) File handling from the acquisition on ROPs to their processing and storage at CCIN2P3.

DataMigro is a daemon running on the background of Far Lab and Near Lab machines, it was originally developed by Alberto Remoto [72] and then taken over by Guillaume Pronost [73] and Romain Roncin [74] in the beginning of 2013. I took over responsibility in early 2015.

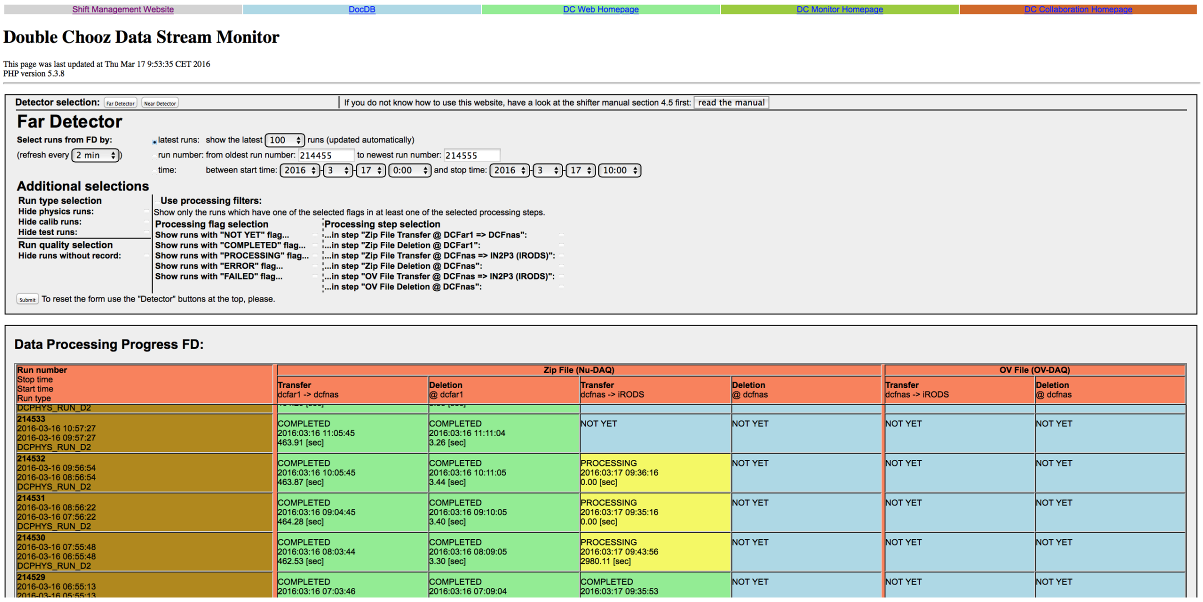
All DataMigro processes have the same structure, described in Figure 2.8. Each process runs on the background of the machines, using a cron system or a simpler bash scripts system. Every minute the process queries a local database to check if a new run can be processed. The status flag of the process ("COMPLETED", "ERROR", or "FAILED") is registered in the database in order to monitor DataMigro and help the experts troubleshooting. This status flag is used to trigger the next DataMigro process. In case of an issue during the processing, an automatic recovery procedure is implemented using a “ERROR” flag in the database. If the recovery procedure failed, the status flag is updated to “FAILED”. DataMigro experts can manually relaunch a failed process for a given run if necessary. DataMigro status is reported to the shifter via a dedicated messaging system. In case of failure, the shifter has to inform DataMigro experts.

The transfer to CCIN2P3 uses the Integrated Rule-Oriented Data System (iRODS), with a $\sim$10 MB/s bandwidth. iRODS is a high performance generic data manager, used in CCIN2P3 with an interface to the permanent CCIN2P3 High Performance Storage System (HPSS). HPSS is a large scale tape storage server with a capacity of 1 to 100 PB.

In order to monitor the DataMigro processes a php-based framework has been developed to allow the monitoring of DataMigro status for each detector in a configurable and user friendly web page, as shown in Figure 2.9.

2.8 Event Reconstruction

Once transferred to CCIN2P3, the binary run files are processed into DOGS files with some basic variable reconstructions. Later on, a second reconstruction layer is also applied to create



Double Chooz Data Stream Monitor

This page was last updated at Thu Mar 17 9:53:35 CET 2016
PHP version 5.3.8

Detector selection: Far Detector Near Detector | If you do not know how to use this website, have a look at the shifter manual section 4.5 first: read the manual

Far Detector

Select runs from FD by: ● latest runs: show the latest 100 runs (updated automatically)
(refresh every 2 min) ○ run number: from oldest run number: 214455 to newest run number: 214555
○ time: between start time: 2016 3 17 0:00 and stop time: 2016 3 17 10:00

Additional selections

Run type selection
Hide physics runs:
Hide calib runs:
Hide test runs:
Run quality selection
Hide runs without record:

Use processing filters:
Show only the runs which have one of the selected flags in at least one of the selected processing steps.

Processing flag selection
Show runs with "NOT YET" flag...
Show runs with "COMPLETED" flag...
Show runs with "PROCESSING" flag...
Show runs with "ERROR" flag...
Show runs with "FAILED" flag...

Processing step selection
...in step "Zip File Transfer @ DCFar1 => DCFnas":
...in step "Zip File Deletion @ DCFar1":
...in step "Zip File Transfer @ DCFnas => IN2P3 (IRODS)":
...in step "Zip File Deletion @ DCFnas":
...in step "OV File Transfer @ DCFnas => IN2P3 (IRODS)":
...in step "OV File Deletion @ DCFnas":

Submit To reset the form use the "Detector" buttons at the top, please.

Data Processing Progress FD:

Run number Stop time Start time Run type	Zip File (Nu-DAQ)				OV File (OV-DAQ)	
	Transfer dcfar1 -> dcfnas	**Deletion** @ dcfar1	**Transfer** dcfnas -> iRODS	**Deletion** @ dcfnas	**Transfer** dcfnas -> iRODS	**Deletion** @ dcfnas
DCPHYS_RUN_D2						
214533 2016-03-16 10:57:27 2016-03-16 09:57:27 DCPHYS_RUN_D2	COMPLETED 2016:03:16 11:05:45 463.91 [sec]	COMPLETED 2016:03:16 11:11:04 3.26 [sec]	NOT YET	NOT YET	NOT YET	NOT YET
214532 2016-03-16 09:56:54 2016-03-16 08:56:54 DCPHYS_RUN_D2	COMPLETED 2016:03:16 10:05:45 463.87 [sec]	COMPLETED 2016:03:16 10:11:05 3.44 [sec]	PROCESSING 2016:03:17 09:36:16 0.00 [sec]	NOT YET	NOT YET	NOT YET
214531 2016-03-16 08:56:22 2016-03-16 07:56:22 DCPHYS_RUN_D2	COMPLETED 2016:03:16 09:04:45 464.28 [sec]	COMPLETED 2016:03:16 09:10:05 3.40 [sec]	PROCESSING 2016:03:17 09:35:16 0.00 [sec]	NOT YET	NOT YET	NOT YET
214530 2016-03-16 07:55:48 2016-03-16 06:55:48 DCPHYS_RUN_D2	COMPLETED 2016:03:16 08:03:44 462.53 [sec]	COMPLETED 2016:03:16 08:09:05 3.30 [sec]	PROCESSING 2016:03:17 09:43:56 2980.11 [sec]	NOT YET	NOT YET	NOT YET
214529 2016-03-16 06:55:13	COMPLETED 2016:03:16 07:03:46	COMPLETED 2016:03:16 07:09:04	COMPLETED 2016:03:17 09:35:53	NOT YET	NOT YET	NOT YET

Figure 2.9: Web interface for the shifters to control the good operation of the transfers.

the Common Trunk files (CT files). The MC files are produced as DOGS format, and then $CT - ed$ like real data files.

2.8.a Energy Reconstruction

An event in Double Chooz consists in a set of waveforms from all the PMTs. For each waveform i, the charge q_i collected by the PMT is obtained by integrating the signal over a time window containing the pulse. A software called RecoPulse is used to extract first level waveform information such as the collected charge, the start time of the pulse, etc. The relevant variables for event reconstruction are defined in Figure 2.10.

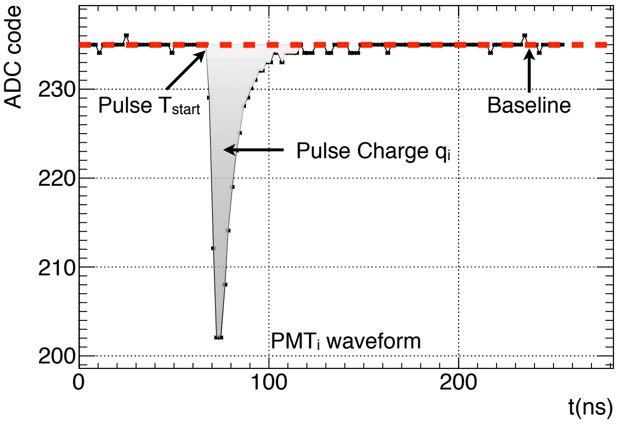


Figure 2.10: Waveform from each PMT i are characterized by their baseline (materialized by the red dashed line), the charge of the pulse q_i, and the start time of the pulse $T_{start,i}$.

The energy reconstruction method is based on the conversion of the total number of photoelectrons, N_{PE}, to the visible energy, E_{vis}. Correction factors for uniformity, f_{u}, energy scale, f_{MeV}, stability for data only, f_{s}, and light non-linearity for MC only, f_{lnl}, are also applied, as follows:

$$E_{\mathrm{vis}} = N_{\mathrm{PE}}(\rho, z, t) \times f_{\mathrm{u}}^{*}(\rho, z) \times f_{\mathrm{MeV}}^{*} \left(\times f_{\mathrm{s}}^{\mathrm{data}}(E_{\mathrm{vis}}^{0}, t)\right) \left(\times f_{\mathrm{lnl}}^{\mathrm{MC}}(E_{\mathrm{vis}}^{0})\right) \tag{2.15}$$

where (ρ, z) stands for the detector coordinates, t for the time, E_{vis}^{0} for the energy once the uniformity correction has been applied and $*$ refers to MC or data sets (FD-I, FD-II, or ND).

The total number of photoelectrons can be extracted from the charge collected by each PMT as follow:

$$N_{PE} = \sum_{i} n_{PE,i} = \sum_{i} \frac{q_i}{gain_i(q_i, t)} \tag{2.16}$$

where $n_{PE,i}$ and q_i represent respectively the number of PE and the charge collected by the PMT i. In Double Chooz, the charge, collected by an individual PMT or summed over the full detector will be expressed in DUQ. This unit corresponds to 1 ADC code integrated over a time sample of 2 ns, knowing the equivalent resistance of a PMT. An ADC-code is the digitization step of the Flash-ADC. $gain_i(q_i, t)$ is a function that translates the charge into a number of PE and corrects for the charge non linearity mostly due to digitization effects (this will be discussed in more details in Section 3.6). This gain function is extracted by calibration taking into account the variation in the course of the data taking and charge dependence, called gain non-linearity, especially sensible at low number of PE.

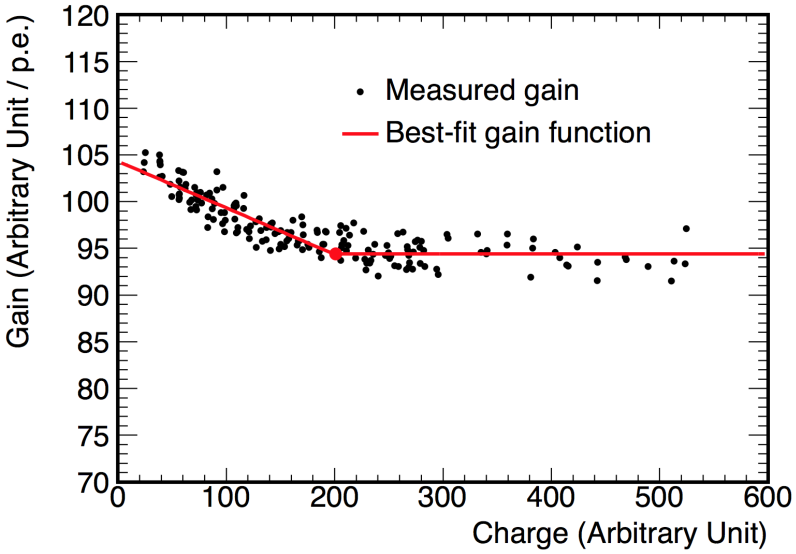


Figure 2.11: Gain as a function of the integrated charge for a typical readout channel. The points correspond to the measurements and the line shows the gain function obtained from a fit [56].

As shown in Figure 2.11 the gain function for each channel i can be described as a constant gain at high charges, and a sloped gain at low charge, due to the gain non-linearity:

$$\begin{aligned} \text{gain}_i(q_i, t) &= g_i(t) + l_i \times (q_i - c_i(t)) & (q_i < c_i(t)) \\ &= g_i(t) & (q_i \geq c_i(t)) \end{aligned} \tag{2.17}$$

where g_i, l_i and c_i are determined for each channel i from a fit of the measured gain for different charge at different time t. The PE calibration, aims to estimate, for each channel, the channel's gains. These gains are measured using the calibration data taken with a constant light yield provided by the Inner Detector light injection calibration system (IDLI). This IDLI calibration, and its Inner Veto equivalent, the IVLI, are performed every week, and after every power cycle, as power cycling the electronics have proven to be the dominant source of the gain non-linearity variations. The average charge to PE conversion factor g_i measured are summarized in Table 2.5 for each of the three detector configurations.

Table 2.5: Charge (DUQ) per PE conversion factor for the various detector settings

Detector	Gain Data (DUQ /PE)	MC (DUQ /PE)
FD-I	88.27	111.44
FD-II	198.50	250.60
ND	185.6	234.30

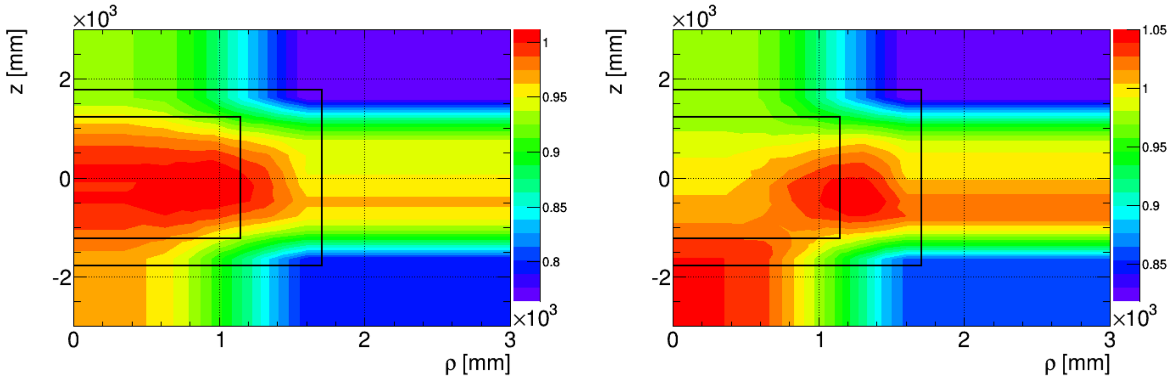
The number of PE collected is then corrected for the effect of the vertex position on the light collection efficiency by the function $f_u(\rho, z)$. This effect is estimated by fitting the neutron capture peak on Hydrogen at a position (ρ, z) in the detector and comparing it to the neutron capture at the center of the Target $(0, 0)$.

$$f_u(\rho, z) = \frac{\text{Peak Mean}(\rho, z)}{\text{Peak Mean}(0, 0)} \tag{2.18}$$

Figure 2.12 shows the uniformity correction maps for the various data sets and MC. Maps for the data were generated using spallation neutron that capture on Hydrogen across both the Target and γ-Catcher volumes. The correction ensures that the visible energy measured for an event does not depend on the event position. The Hydrogen peak is used as it is narrower than the Gadolinium one, and thus provides a more precise measurement of the possible shift in its position.

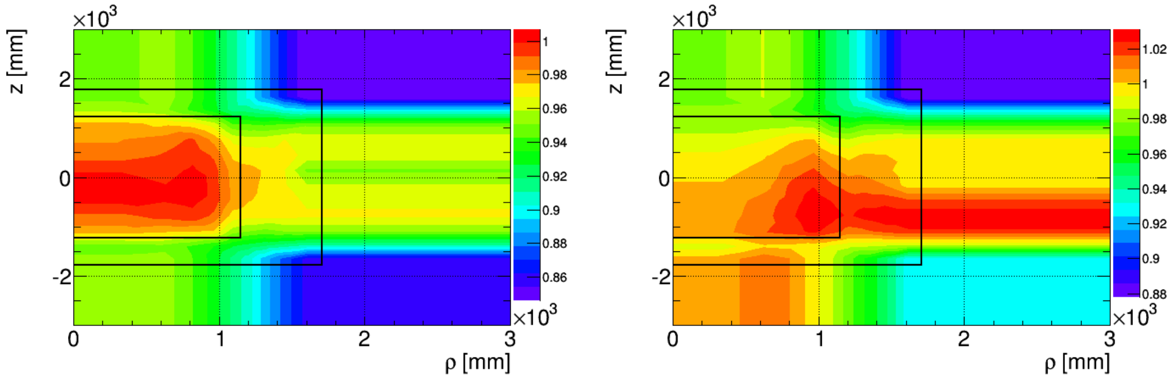
The absolute energy scale f_{MeV} has been studied using the capture on Hydrogen of neutrons from a ^{252}Cf source deployed at the center of the Target.

The energy stability for data sets $f_s(E^0_{vis}, t)$ has been introduced to account for the variation of the electronics response over time, i.e. the time dependence of the reconstructed number of PE. The stability is studied using spallation neutron captures on Hydrogen and Gadolinium as well as the α from the ^{212}Po decay. The energy of the α is 8.8 MeV but results in a visible energy of about 1 MeV due to quenching effects, and therefore allows stability monitoring in the lower energy region below neutron capture peaks.



(a) Uniformity map for MC FD

(b) Uniformity map for MC ND



(c) Uniformity map for Data FD

(d) Uniformity map for Data ND

Figure 2.12: Uniformity maps for MC (top) and data (bottom) in FD (left) and ND (right). The maps are obtained by fitting the mean energy of the neutron capture peak on Hydrogen at different positions in the detector and comparing it to the value at the center. The black lines represent the limits of the Target at the center and the γ-Catcher surrounding it.

Figure 2.13 shows the stability-corrected evolution of the n-Gd, n-H and ^{212}Po α energies. Their evolutions have been stabilized down to the percent level.

The MC visible energy is corrected for the remaining energy non-linearity in order to improve the agreement between data and MC. This energy non-linearity arises from two sources: the charge non-linearity (QNL) and the light non-linearity (LNL). LNL is particle dependent, and therefore, it is not applied in the visible energy equation, only QNL is applied as follow:

$$\begin{aligned} f^{MC}_{NL}(E^0_{vis}) = f^{MC}_{QNL}(E^0_{vis}) &= 0.0023 \times E^0_{vis}[\mathrm{MeV}] + 0.9949 \qquad && (E^0_{vis} < 10\,\mathrm{MeV}) \qquad (2.19)\\ &= 1.0179 && (E^0_{vis} \geq 10\,\mathrm{MeV}) \end{aligned}$$

The LNL is included in the energy model used in the θ_{13} fit. Figure 2.14 shows the energy resolution for FD-I data and for MC. Horizontal axis shows the peak energy obtained by a

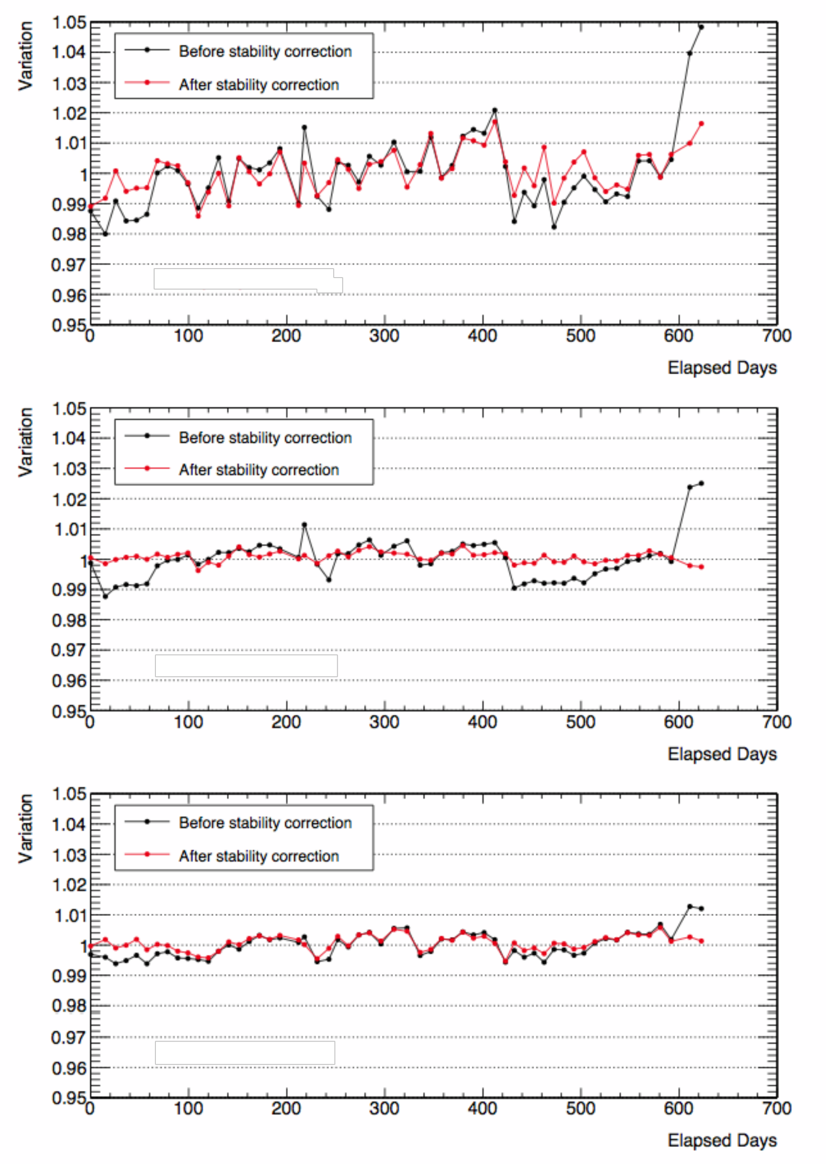


Figure 2.13: Ratio of the peak energy of ^{212}Po α decay (top) and neutron captures on H (center) and Gd (bottom), before and after application of the stability correction, to their nominal value as a function of time for FD-I data. Gain variations over time are dominated by the electronics power-cycling and are especially sensible at low energy (for the ^{212}Po α).

fit and vertical axis shows the energy resolution. Black circles correspond to the data taken with calibration sources at the center of the Target, and black squares correspond to neutron captures on C and Gd from data (IBD or muon-induced spallation neutron). The red points correspond to the equivalent for MC.

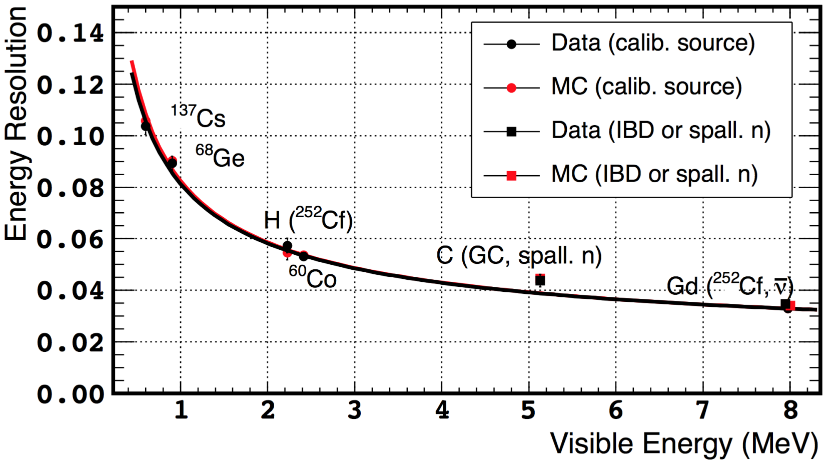


Figure 2.14: Energy resolution for FD-I data. The fits take 3 parameters (a,b,c) where a is the stochastic term of the energy resolution, b a constant term and c is linked to the electronics noise.

2.8.b Vertex Reconstruction

The vertex reconstruction in the Inner Detector is performed by an algorithm called RecoBAMA based on a charge and time likelihood maximization. RecoBAMA assumes an instantaneous, point-like energy deposition in the detector, and an isotropic light propagation. Thanks to these assumptions, a vertex can be described by $\vec{X} = (x_0, y_0, z_0, t_0, \Phi)$ with a single position (x_0, y_0, z_0), time (t_0), and the amount of light per steradians (Φ).

The time of arrival of the light on each PMT i can be predicted as:

$$t_i = t_0 + \frac{r_i}{c_n} \tag{2.20}$$

where r_i is the distance of the PMT i to the vertex position and c_n is the effective light velocity along the light path.

Assuming the light is only affected by pure attenuation, the amount of light μ_i received by each PMT i is expressed as:

$$\mu_i = \Phi \times \epsilon_i \times \Omega_i \times e^{\left(-\frac{r_i}{\lambda}\right)} \tag{2.21}$$

where ϵ_i is the quantum efficiency of the PMT i, Ω_i is the solid angle subtended by the PMT photocathode and λ is the effective light attenuation length along the light path.

The likelihood is defined as:

$$\mathcal{L}(\vec{X}) = \prod_{q_i=0} f_q(0; \mu_i) \prod_{q_i>0} f_q(q_i; \mu_i) f_t(t_i; t_i^{pred}, \mu_i) \tag{2.22}$$

The first product is performed over all PMT with zero charge, which did not collect the expected amount of light μ_i because of quantum efficiency and photon statistics, this is a crucial information for the vertex reconstruction. The second product is performed over the PMT that have been hit, with non-zero charge. This product takes both the charge and time information. f_q and f_t are respectively the charge and time probability density functions (PDF), computed from MC simulations and validated with physics and calibration data.

The reconstructed vertex is the vector $\vec{X}$ that minimizes the negative log-likelihood function, later on called Functional Value:

$$FV(\mathcal{L}) = -\log \mathcal{L}(\vec{X}) = -\sum_{i} \ln f_q(q_i; \vec{X}) - \sum_{q_i>0} \ln f_t(t_i; \vec{X}) = FV_q(\vec{X}) + FV_t(\vec{X}) \tag{2.23}$$

RecoBAMA uses both charge and time informations to reconstruct the vertices, and returns a global FV value, but also the individual FV$_t$ and FV$_q$ values. The accuracy can be measured using deployed sources. Figure 2.15 shows a $\sim$ 13 cm resolution of the reconstruction of the position of a ^{68}Ge source deployed at several positions along the Z axis of the Target.

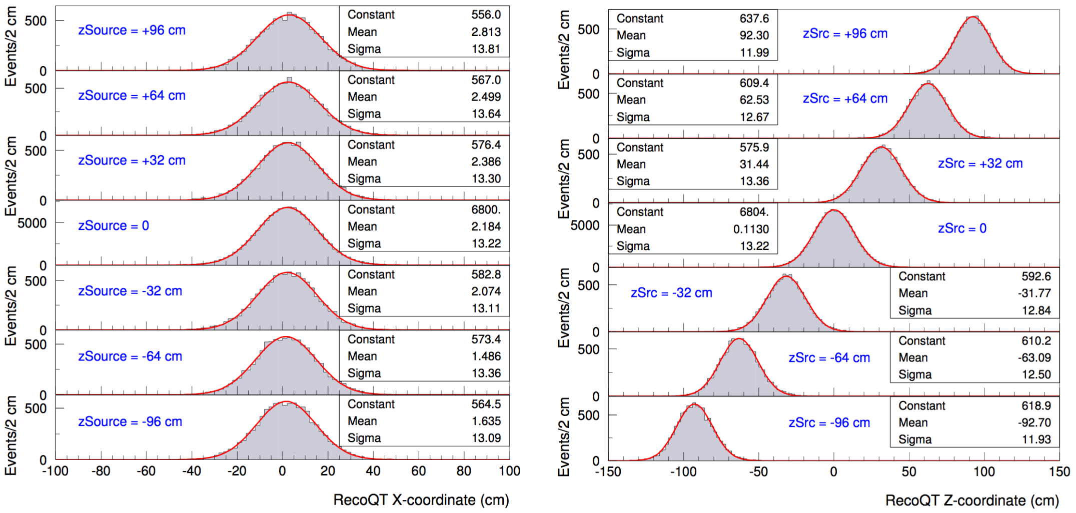


Figure 2.15: Reconstructed vertex from a ^{68}Ge source deployed at several positions along the Z axis of the Target.

2.9 Conclusion

Double Chooz proposed the now widely used multi-detector design. The multi-detector design helps reducing the correlated detection uncertainties and the reactor flux prediction uncertainties. In the single-detector phase, flux uncertainties are reduced by using the normalization measurement performed by Bugey4 as a virtual Near Detector.

The physical containment of the Gd-loaded scintillator allows for a better control of the detection volume, and the presence of the γ-Catcher allows to ensure an efficient energy collection over the full Target volume. The addition of the buffer helps reduce the radioactivity from the PMT, and the Inner Veto allows an active rejection of external contamination such as Muons, Fast Neutrons and penetrating high-energy γ from the surrounding rock radioactivity.

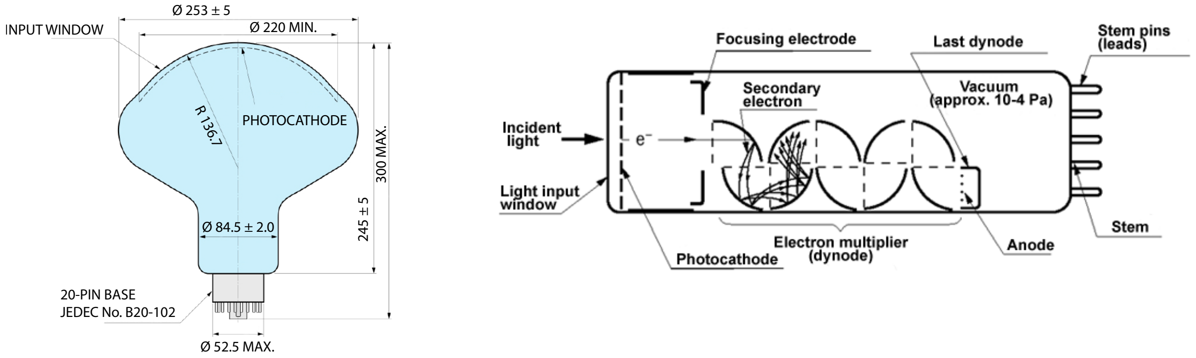
The energy reconstruction is based on the light collection by the 390 PMTs in the Inner Detector, and multiple calibration methods have been developed to ensure a control of the energy non linearity at low energy.

Chapter 3

Calorimetry with Flash-ADC Digitizers

3.1 PhotoMultiplier Tube Principle

Light collection in Double Chooz is achieved by the use of PMTs (Hamamatsu R7081). The working principle of a PMT is described in Figure 3.1: A Photo-Electron (PE) is emitted by photoelectric effect when a photon in the visible range hits the PMT's photocathode. This PE is then accelerated and focused on the first dynode by a focusing electrode. The dynodes produce secondary electrons at each stage, amplifying the electron flux. At the end of the dynode chain, the electrons are collected on an anode, producing a measurable electric current.



(a) (b)

Figure 3.1: 3.1(a): Shape of Double Chooz's ID-PMT Hamamatsu R7081.
3.1(b): Working principle of a PMT: an incident photon knocks an electron (Photo-Electron) out of the photocathode through photoelectric effect. This PE is then accelerated onto the first dynode and is amplified through an electron avalanche at each dynode.

Double Chooz's scintillator yields an average of $800\gamma/\mathrm{MeV}$ with a maximum of emission at a wavelength of 420 nm, which, with a 25% quantum efficiency of the PMT photocathode at this wavelength [75] results in 200 PE/MeV for the full 390 ID-PMT, which translates into 0.5

PE per MeV and per PMT. Signals from IBD in the target range from an average of 1 to 5 PE per PMT, thus, in order to digitize these signals well enough while not saturating, PMTs are operated at a gain of 10^7.

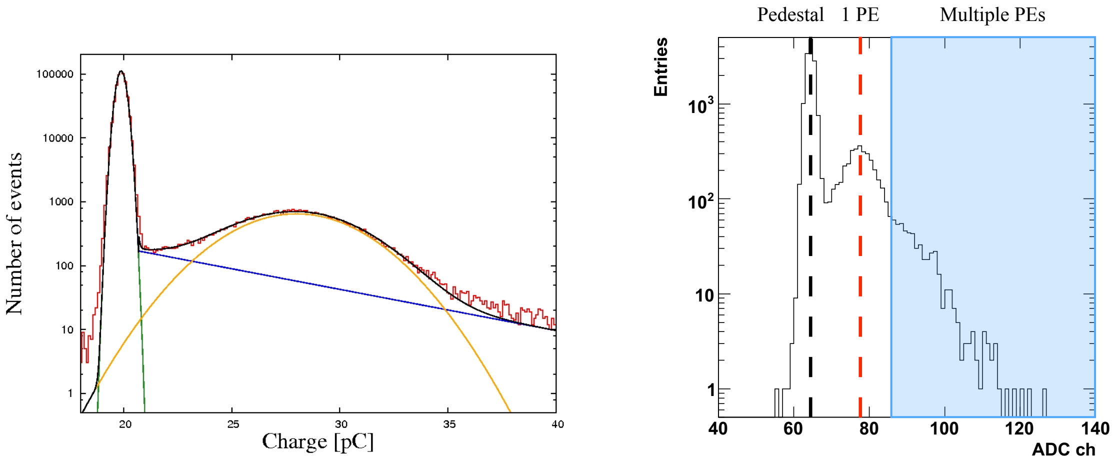
In the following, a PE will refer equally to either the primary electron or the corresponding pulse on the waveform output by the PMT at the end of the amplification.

Amplification of a PE is a series of stochastic processes and is largely dependent on the energy of the electron hitting the first dynode. This energy depends on the position where the photon hits the photocathode, and the homogeneity of the accelerating and focusing field. In general, compared to PMTs with small photocathodes, those with larger photocathodes are less homogeneous and the center of their photocathode is more efficient than the edges. This leads to a widening of the final charge per PE distribution.

Imperfections of a PMT may come from different sources:

- Non-uniformity of the photocathode quantum efficiency (i.e. the probability to emit a PE for every photon hitting the photocathode).
- Non-uniformity of the focusing electric field.
- Probability of first emission at the first dynode level, if a visible photon crosses the photocathode and hits the first dynode, producing a PE, and thus missing an amplification step.
- Probability of after pulse, when an ion is knocked out of the photocathode as well as an electron. The ion will slowly drift back to the photocathode and may re-emit an electron, thus producing a second avalanche. This process is very slow, however, the second peak comes after a few microsecond, and will not be on the same waveform in our case.
- Probability of late pulse, when the electron is in-elastically scattered at the dynode, reflected and then re-accelerated. This process is much faster than the previous one, and only take $\sim 100\,\text{ns}$ in Double Chooz.
- Rate of thermal emission: The energy of electrons in the dynode follows the Boltzmann distribution in $e^{-E/k_B T}$. Because the PMTs are at room temperature, some electrons may have enough energy to leave the dynodes, which produces a measurable avalanche.

Figure 3.2(a) shows the charge distribution of a PMT in Double Chooz that was acquired during PMT tests performed during the commissioning of the Far Detector and is reported in [75]. The PMTs were illuminated with a LASER LED with an intensity of 0.1 PE per pulse. The distribution can be described by a series of Gaussians and an underlying exponential:
The first Gaussian, centred on $\sim 20\,\text{pC}$, is the pedestal and corresponds to the integration of empty waveforms, with no PE.
The second Gaussian corresponds to waveforms with PE and is centred on the value of the pedestal plus the gain of the channel (gain of the PMT and possible later amplification).
The underlying exponential corresponds to thermal electron emission from the first dynode.



(a)

(b)

Figure 3.2: 3.2(a): Charge distribution (in pico Coulomb) for Double Chooz ID-PMT illuminated by a LASER LED with an intensity of 0.1 PE per waveform acquired during the PMT tests for the commissioning of the Far Detector reported in [75]. The distribution can be described by a two of Gaussians (respectively 0 and 1 PE) and an underlying exponential.
3.2(b): Same distribution with a LED intensity of 0.25 PE per waveform obtained in [76]. The pedestal and 1 PE are visible as well as the multiple PE contributions.

The number of PE is the physical observable since it is directly linked to the energy collected on the PMT and thus to the energy of the event. A PMT waveform is a superposition of multiple single PE contributions. Given the wide charge distribution for a single PE, it is often impossible to distinguish between one PE with high charge or two PE with each a low charge as shown in Figure 3.2(b) where 1-PE and 2-PE contributions are piling-up. It is however possible to give a statistical estimation of the number of PE by dividing the charge by the gain, as described in Section 2.8.a.

3.2 Double Chooz Digitization System

Flash-ADC are a means to perform an Analog to Digital Conversion (ADC) of a voltage or current with fast sampling. Waveform digitization allows to measure a variety of parameters on the signal, such as pulse charge, start time, rise time and quality estimators such as the baseline estimation. This also allows to store the waveforms and to re-analyse them later on if necessary, e.g. if the reconstruction algorithm changes, which is impossible when such measurements are done with analog electronics.

The read-out chain has been presented in Section 2.6.a. The Front End Electronics (FEE) amplifies the PMT waveforms for future digitization, and performs a summation of groups of waveforms for trigger decision taking (groups of 16 channels for ID-PMT and 6 channels for IV-PMT).

After amplification, the mean single PE amplitude is of the order of 70 mV for FD-II and

ND configurations, and $\sim 35\,\mathrm{mV}$ for FD-I. The analog RMS noise level was measured to be $\sim 1.2\,\mathrm{mV}$ for FD-I configuration and $\sim 1.9\,\mathrm{mV}$ for FD-II and ND.

The amplified waveforms are then digitized by a 8-channel, 8-bit, 500 MS/s Flash-ADC (CAEN Vx1721). All PMT signals are sampled synchronously every 2 ns. The Least Significant Bit (LSB) is the voltage step of the digitizer. It corresponds to the analog voltage equivalent to one ADC code. The waveforms are digitized over 8 bits, so with $2^8 = 256$ADC codes, for a dynamic range of 1 V. Therefore the LSB is 3.9 mV. The analog noise RMS is ~ 0.3 LSB for FD-I [77], and ~ 0.37 LSB for FD-II and ND and originates dominantly from FEE amplification. The average amplitude of a single PE waveform is digitized over ~ 18ADC codes for FD-I and ~ 50ADC codes for FD-II and ND.

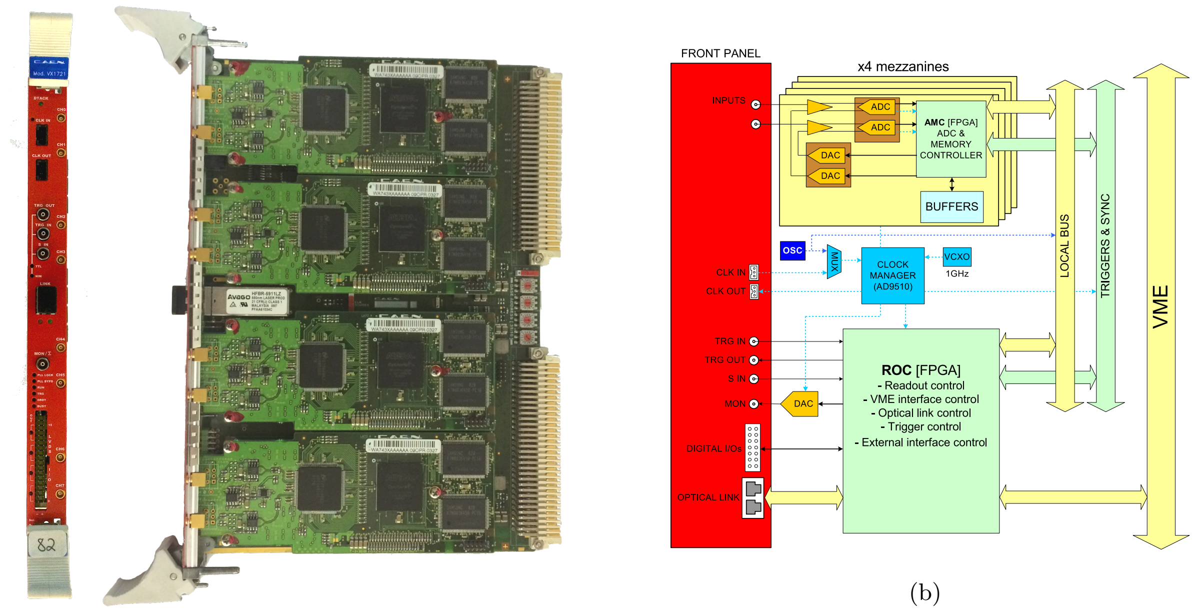
Flash-ADC architecture The Flash-ADC boards are hosted in a VME crate that provides power and communication with the data acquisition system controlled by a single-board computer (PowerPC) with software written in Ada-2005. Digitized waveforms are sent to the computer through the VME64x transfer protocol 2eSST in which the data is transferred by blocks of 64 bytes using the double edged Source Synchronized Transfer (transfer on the rising and falling edges of the clock signal), giving a maximum transfer speed of 320 Mbytes/s.

Figure 3.3(a) shows a front and side view of Double Chooz Flash-ADC cards. The 8 digitization channels are divided in 4 groups of 2 channels. Each group of 2 channels is mounted on a separate mezzanine and the 4 mezzanines are then mounted on a motherboard.
Each mezzanine contains a 2-channel ADC chip, a 2-channel DAC chip, a FPGA chip (Field Programmable Gate Array) and two RAM buffers. The FPGA on each mezzanine controls the digitization of the two channels, the reference voltage, the addressing and read-out in the buffers. On the motherboard, a FPGA controls the trigger signal and clocks distribution, the data read-out from the mezzanines, and all input/output processes with the VME crate and the front panel connectors. A block diagram of the Flash-ADC board organization is shown in Figure 3.3(b).

Event meta-data The Trigger system generates information about the triggering conditions and calculates the event number. Synchronous to the trigger, this information is passed to the Flash-ADC boards through the 16-bit LVDS input connector on the front-panel. This information, together with the number of clock ticks since the previous trigger and an internal trigger counter, are stored by the Flash-ADC. This constitutes the event meta-data; it can be read from the VME, and used to categorise the event types and differentiate data handling.

Clock distribution within the Flash-ADC The Flash-ADC cards are fed a 62.5 MHz clock from the trigger system, synchronously to all other Flash-ADC cards. This external clock is sent to a Clock Manager. The Clock Manager has a PLL that synchronizes an internal 1 GHz oscillator to the external clock, and outputs several clock signals.

Four 500 MHz sampling clocks are sent to the ADC on each of the four mezzanines, a 62.5 MHz is used as the clock output, and another is sent to the motherboard FPGA, with a



(a)

(b)

Figure 3.3: 3.3(a): View of the front panel and the four mezzanines of a Flash-ADC card. The motherboard is hidden below the mezzanines.
3.3(b): Block diagram of a Double Chooz Flash-ADC card.

software-tunable delay. The motherboard FPGA then distributes the synchronous 62.5 MHz to each mezzanine FPGA to serve as synchronization between the mezzanine FPGA and the ADC. When synchronizing the Flash-ADCs, this tunable delay is adjusted so that all mezzanines FPGA on all Flash-ADC cards see the same clock phase, i.e. are synchronized within 2 ns.

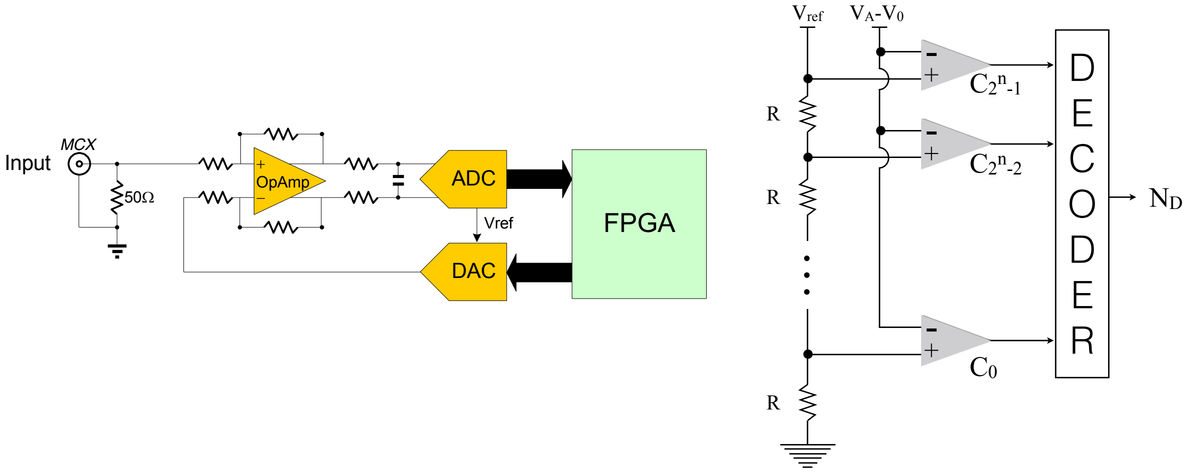
The communication between the ADC and the FPGA on the mezzanines is achieved by writing blocks of 16 bits per channel at 125 MHz. This transfer is done using the DDR protocol, by writing on both fronts of the clock signal. Data is transferred from the mezzanine to the motherboard through a FIFO by sending groups of 64 bits at 62.5 MHz.

3.3 Signal Digitization

The digitization principle of the signal from one acquisition channel is described in Figure 3.4. Figure 3.4(a) shows the diagram of the input signal digitization by Double Chooz's Flash-ADC. Every 2 ns, the input analog signal (V_A) and a reference voltage (V_{ref}) are fed to a 8-bit Analog to Digital Converter (ADC) that outputs a digital value N_D to a FPGA. N_D ranges from 0 to $2^8 - 1 = 255$.

A constant reference voltage V_0 from a 12-bit DAC is subtracted from the analog signal. The value of V_0 can be written dynamically through the VME bus. This value sets the baseline of the PMT waveform and allows us to optimize the dynamic range to digitize the down-going PMT signal while being able to digitize the possible overshooting of the signal following a high energy event such as a muon crossing the detector. V_0 is set so that the baseline is at 210.5 ADC codes in the FD-I configuration. We saw no benefit from such a large positive dynamic range in FD-I data, and decided to reduce it and set the baseline to 230.5 ADC codes in the FD-II and ND configurations.

The stability of V_0 is crucial in the reconstruction of the events by RecoPulse, as it is the level zero from which the digitized waveform is integrated to compute the charge collected by the PMTs. The effects of the baseline on the energy linearity will be discussed in more details in Section 3.6.



(a) (b)

Figure 3.4: 3.4(a): Double Chooz Flash-ADC digitization block diagram. This diagram represents the signal path from input to digitization by the ADC and storage on the FPGA.
3.4(b): Digitization principle of an ADC. The input voltage $V_A + V_0$ is compared to each of the $2^n - 1$ step of a voltage ladder of total voltage V_{ref}, where n is the number of bits over which the signal is digitized.

Figure 3.4(b) shows the working principle of the ADC. The analog signal $(V_A - V_0)$ is compared to each step of a voltage ladder by $2^n - 1$ comparators (C_i). The output of each comparator is a high or low (1 or 0) saturated signal. The binary output is then coded into a 8-bit ADC count N_D by the decoder and sent to the FPGA. N_D corresponds to the number of the comparator with the voltage step closest, while still being lower, to the analog value, performing an effective flooring of the signal.

The FPGA in turn writes each byte in a circular buffer of 1024 pages of 2048 samples (4096 ns). A page is written continuously up to the 2048^{th} sample. If no trigger signal is received during that time, the first sample of the same page is overwritten by the 2049^{th} sample and so on. When a trigger signal is received, the writing process is moved to the next page and the current page is ready to be read-out.

The address on the page of the trigger point can be calculated knowing the page duration, 4 μs and the trigger time, the time between this event and the precedent. A "slice" of data of 256 ns is then sent to a FIFO in the motherboard FPGA. From this FIFO, the waveforms can be transferred across the VME backplane to the PowerPC.

While a page is read-out, another page is written, so there could be multiple triggers in the time it takes to read a page. Once the waveform has been transferred, the corresponding page is cleared. If reading is too slow compared to the trigger rate, the number of written pages can increase, up to a maximum of 1024 pages. The firmware developed for Double Chooz thus allows a variable size of waveforms to read from 0 µs (no data) to a maximum of 4 µs. In Double Chooz, the mean trigger rate can go up to 1 kHz before a problem arises. When that happens, an exception is raised, data taking stops and an alert is sent to the shifter.

3.4 Digitization Limitations

Several effects intrinsically appear when digitizing the PMT waveform that will later on limit our knowledge of the signal, and thus its reconstruction.

In the digitization of an analog waveform, two discretizations occur:

- sampling: discretization in time,
- quantization: discretization in amplitude.

An obvious limitation arises from the sampling frequency of 500 MS/s (Mega Samples per second). The Nyquist-Shannon sampling theorem holds for signals at frequencies lower than 250 MHz. The speed of the PMT coupled with the bandwidth limitations of the cables and the Front-End amplifiers ($< 200\,\mathrm{MHz}$) ensure that the condition of the sampling theorem is met. In addition, a bandwidth filter inside the Flash-ADC, before the digitizer, limits the frequency of the signal to below the 250 MHz limit.

Similarly to the Nyquist-Shannon theorem, B. Widrow derived two quantization theorems (QT1 and QT2 [78]). Those theorems link the signal pdf to the quantization step, which in our case corresponds to 1 ADC count, or 1 LSB.

The first theorem, QT1, states that if the amplitude of the analog signal is greater than 1 LSB, the statistical description of the analog signal can be exactly reconstructed from the digitized output.
In Double Chooz, the average single PE pulse is well digitized, over $\sim$ 18ADC codes, and as such can be well reconstructed.

The second theorem, QT2, is a looser condition: if the amplitude of the analog signal is greater than 0.5 LSB, the estimated mean and variance of the output digitized signal are equivalent to that of the input analog signal.

The baseline is estimated from waveforms with no PE pulse, thus the signal is essentially flat, with fluctuations only due to the analog noise. In that case, it is the noise RMS that becomes critical to estimate the precise position of the baseline.

If the waveform is noiseless, the baseline determination is impossible. As only one code would be hit, all that could be known would be the position of the baseline within 1 ADC code. On the contrary, if the Gaussian noise RMS is big, the hit code distribution would be a well drawn Gaussian, with a well measurable mean and variance. This is the essence of QT2.

Double Chooz analog RMS noise level is below 0.5 LSB, so there is an intrinsic bias between the true and the measured baselines. Carbone and Pietri derived the biases for constant signals with several noise distributions including the Gaussian case in [79]. The offset between the observed and the true baselines $m_e(e_q)$ can be expressed as a sinusoid, the amplitude of which decreases exponentially with the noise RMS σ_n:

$$m_e(e_q) \sim \frac{1}{\pi} e^{-2\pi^2\sigma_n^2} \sin(2\pi e_q) \ , \sigma_n \geq 0.3\,\mathrm{LSB} \tag{3.1}$$

where e_q is the true offset of the baseline within one ADC code. Figure 3.5(a) shows the simulated evolution of the bias between the true and the reconstructed waveforms with respect to the positions of the true baseline, for different noise levels. The bias evolves as a sinusoid, as described in Equation 3.1, but the flooring performed by the Flash-ADC introduces a constant bias of -0.5 LSB, for any noise level. Therefore, the baseline position is constantly underestimated.

The evolution of the bias amplitude with the noise level proves problematic since every channel does not necessarily have the same level of analog noise. Placing the true baseline at an integer value of a code minimizes this evolution since for any level of noise, the signal has a 50% probability to hit the desired code N, and a 50% probability to hit the code $N-1$. This leads to a constant offset of 0.5 LSB when averaging over the samples, placing the observed baseline at code $N-0.5$ regardless of the noise level.

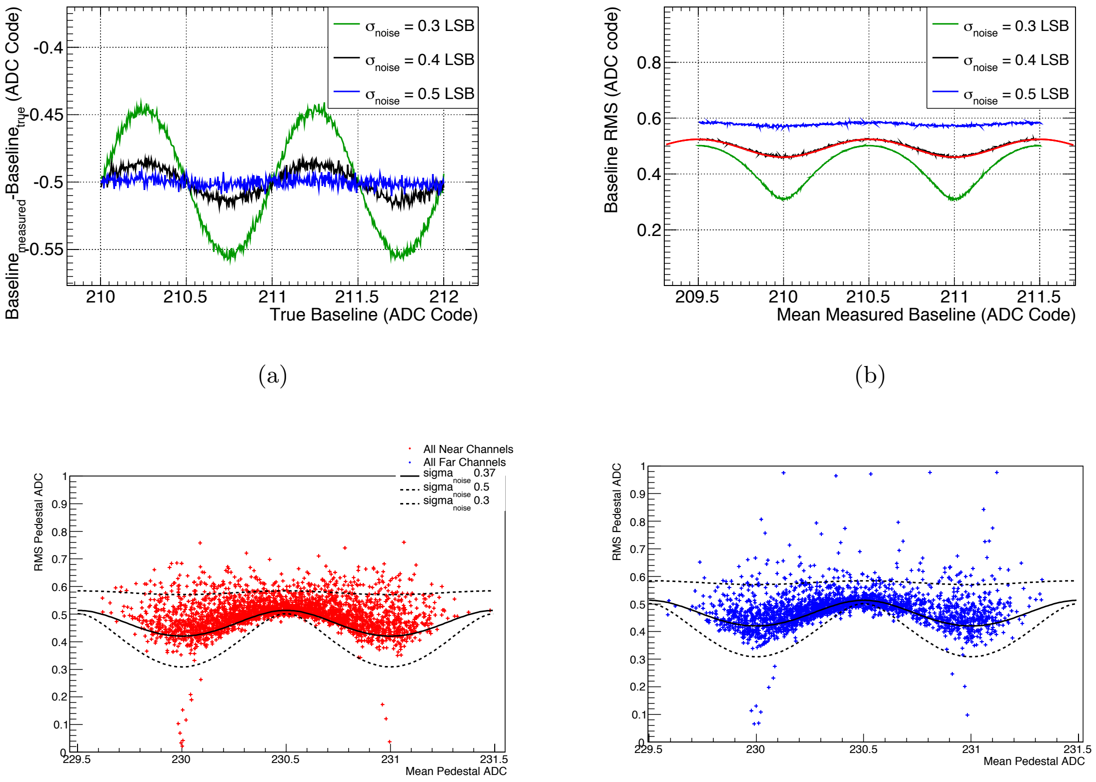
The bias in the baseline determination leads to a non-linearity in energy as discussed in more details in Section 3.6.

The variance of the output signal σ^2, i.e. the variance of the hit code distribution, can be expressed as [79].

$$\sigma^2 = \frac{1}{12} + \sigma_n^2 - e^{-2\pi^2\sigma_n^2}\left[\left(4\sigma_n^2 + \frac{1}{\pi^2}\right)\cos(2\pi e_q) - \frac{1}{\pi^2}e^{-2\pi^2\sigma_n^2}\sin^2(2\pi e_q)\right] \ , \sigma_n \geq 0.3\,\mathrm{LSB} \tag{3.2}$$

The evolution of the RMS of the output signal with the observed baseline has been simulated in Figure 3.5(b) for various noise levels. The red line shows the analytical description of Equation 3.2 for a noise level of 0.4 LSB. The maximum spread of the output digitized values is obtained for an observed baseline at the middle of the code ($N-0.5$), which corresponds to a true baseline at code N.

An evaluation of the analog noise was performed during the FD-II electronics update and the Near Detector commissioning by estimating the RMS of the output digitized signal while varying the baseline position and is presented in Figure 3.5(c). The red and blue dots correspond to



(a)

(b)

(c)

Figure 3.5: 3.5(a): Simulated evolution of the bias between the measured baseline and the true baseline.
3.5(b): Evolution of the RMS of the hit code distribution for simulated empty waveforms with different analog noise levels with respect to the measured baseline value. The red line is the analytical description of this behaviour, discussed in [79].
3.5(c): Measurement of the RMS of the hit code distribution for channels of the Near Detector (red dots) and the Far Detector (blue dots). The lines correspond to the analytical evolution for analog noises of (from bottom to top) 0.3 ADC code, 0.37 ADC code and 0.5 ADC code.

respectively the Near Detector and the Far Detector, superimposed are the shapes of Equation 3.2 for noise levels of 0.3 ADC code, 0.37 ADC code and 0.5 ADC code. The mean noise level over all channels is consistent with 0.37 ADC code. A few points do not fit with the rest of the distribution, they correspond to empty channel, with no signal. The RMS of the noise on those channels is much lower than 0.3 ADC code, so Equation 3.2 is no longer valid.

3.5 Evaluation of the Near Detector's Flash-ADCs

3.5.a Linearity Tests

The digitization limitations discussed previously are intrinsic to the digitization process. Other sources of bias may arise, from the Flash-ADC being not ideal. The linearity of the

Flash-ADC is the linearity of the voltage-to-ADC function.

The two relevant values characterizing the linearity of an ADC are the Differential Non-Linearity (DNL) which measures the deviation from the ideal size of 1 LSB and the Integral Non-Linearity (INL) which measures the cumulative effect of the DNL. It is the difference between the Flash-ADC and the ideal voltage-to-ADC transfer function. Those parameters determine the real voltage-to-ADC transfer function of the considered Flash-ADC, how an analog voltage input is converted to an ADC code.

The baseline is estimated by averaging the first few samples of the digitized waveform. The relative size of the ADC codes that correspond to the baseline (i.e. the DNL of those codes) will impact the probability of each of those codes being hit, and thus impact the baseline determination. The INL is the overall bias of the digitizer's output compared to the analog waveform. A high INL can affect the shape and amplitude of the PMT signal, but not the baseline since the bias is constant in time, as long as the Flash-ADC is not power-cycled.

During the Near Detector commissioning, the Flash-ADC's linearity and working state needed to be asserted before we could install them on-site and begin data taking. An evaluation of the Far Detector's Flash-ADC was performed as well in [77] during its commissioning. I was charged to provide a test to evaluate the Flash-ADCs and to screen the defective cards.

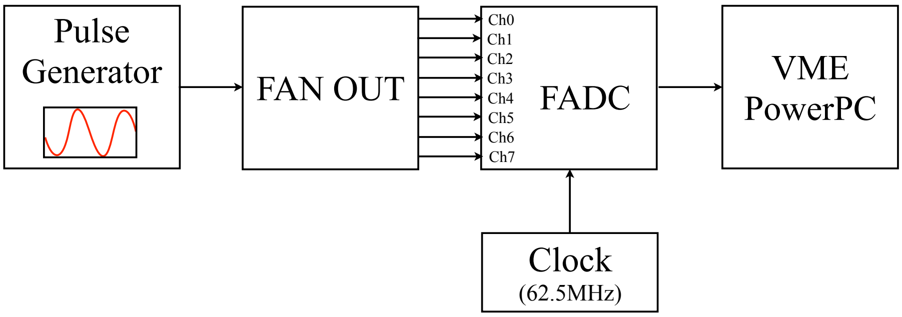
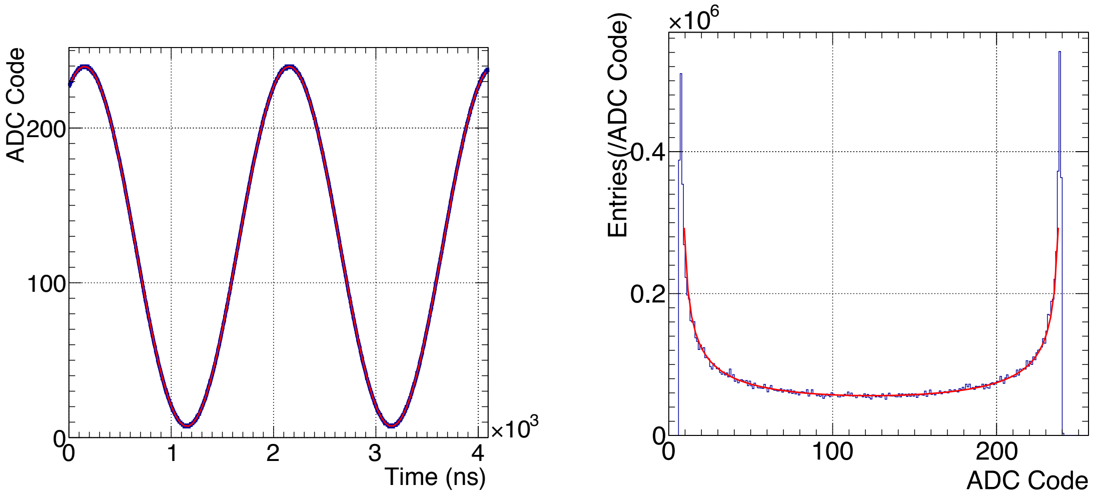


Figure 3.6: Schematics of the set-up used to test the linearity of the Flash-ADCs.

In order to measure the DNL of each channel, the set-up described in Figure 3.6 has been used: a 1 V, 12-bit pulse generator outputs a 3 MHz sinusoid signal of amplitude 4.5 V by incremental steps of 244 µV. The signal from the pulse generator is then fed to each of the 8 channels of the 8-bit Flash-ADC card via a FAN OUT. Figure 3.7(a) shows the Flash-ADC digitized waveform (blue), and the input sinusoid from the pulse generator. The trigger was realized internally to the Flash-ADC, asynchronously from the pulse-generator signal.

For each channel, 100 waveforms of 2048 samples were recorded and a histogram of the sampled codes was made. A perfect Flash-ADC would show a hit code distribution following the pdf of a sinusoid. Figure 3.7(b) shows the hit code distribution for a typical Flash-ADC channel (blue), and the fitted sinusoid pdf.

If a code size deviates from 1 LSB, the code's hit probability will change, and that code's



(a)

(b)

Figure 3.7: Waveform and pdf of a test of a single channel on a Flash-ADC.

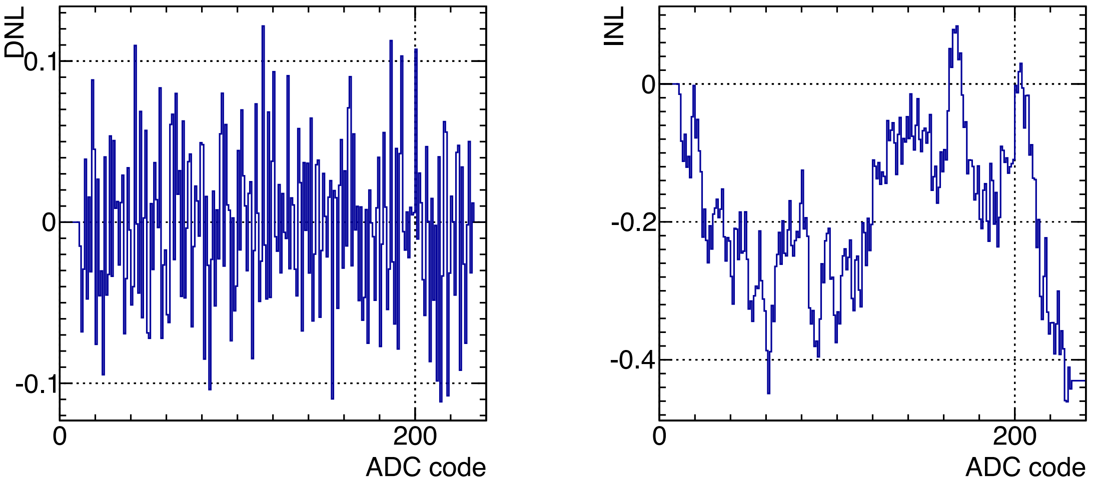
value in the hit code distribution will be further away from the fitted pdf. Statistical fluctuations from analog noise are expected to be negligible when averaging over numerous waveforms as the effects of the DNL are constant but not the noise fluctuations. Comparing the number of hits for each code to the pdf of a sinusoid function allows to estimate the DNL per code. The DNL for each tested code is shown in Figure 3.8(a). The INL per channel is computed as the cumulative of all the DNLs of the channel and is shown in Figure 3.8(b).

Each Flash-ADC channel is characterized by computing the standard deviation of the DNL and INL distributions, as shown in Figure 3.9. In our sample of cards, the average DNL is $\sim 4.1 \cdot 10^{-2}$LSB with a maximum of $6 \cdot 10^{-2}$LSB and the average INL is ~ 0.17 LSB with a maximum of 0.38 LSB. All values obtained are well within the specifications of the manufacturer, where the typical maximum DNL value is 0.16 LSB with limits of ±0.6 LSB and typical INL value is 0.3 LSB with limits of ±0.9 LSB [80].

It is most important that the INL stays within the manufacturer's limits as the INL is the non-linearity between the output code with respect to the corresponding input voltage. Too big an INL would mean an important bias in the waveform amplitude, and thus on the measured charge of the event.

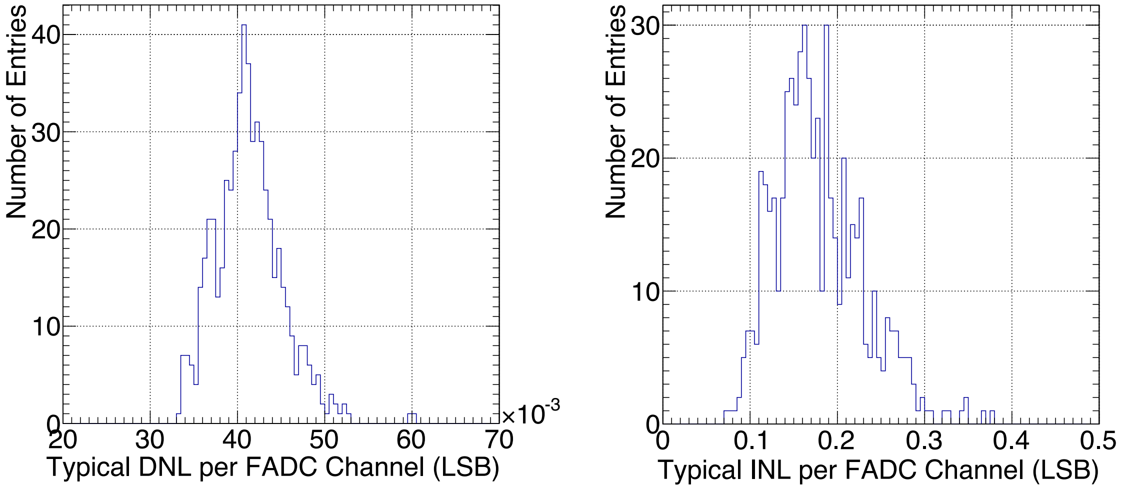
3.5.b High Frequency Correlated Noise

Figure 3.10(a) shows the mean waveform, obtained by averaging over 100 sinusoidal waveforms with random trigger phases in dark blue. When doing so, the analog noises average out, as does most of the sinusoid contribution. Unfortunately, the finite number of phases averaged on does not enable to completely cancel the sinusoid. The residual waveform has a very important component of high frequency noise. This noise does not cancel out as the analog noise or



(a) (b)

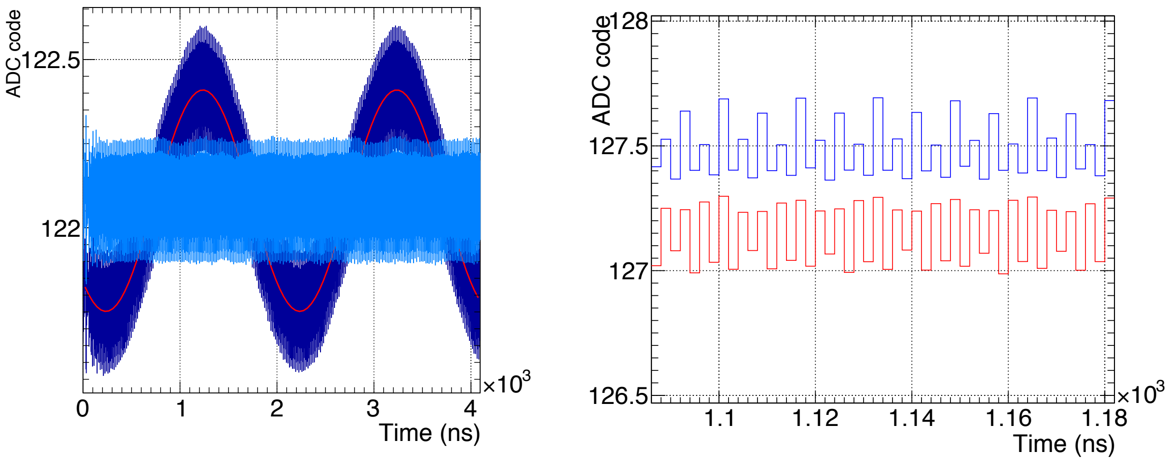
Figure 3.8: Non linearity measurements on a single channel:
3.8(a): DNL per code for a single channel. 3.8(b): INL evolution per code on the same channel.



(a) (b)

Figure 3.9: Non linearity measurements of all channels.
3.9(a): The typical DNL (i.e. the RMS of the DNL) per channel for all tested cards.
3.9(b): The typical INL (i.e. the RMS of the INL) per channel for all tested cards.

the random phased signal since it is not random. Subtracting the residual sinusoid component leaves only the high frequency noise (light blue line in Figure 3.10(a)).



(a)

(b)

Figure 3.10: 3.10(a): Dark blue: signal averaged over 100 waveforms with random trigger phase. Light blue: after subtraction of the sinusoid component, only the clock-induced noise remains. 3.10(b): Clock-induced noise: close-up. The 16 ns period is clearly observable.

The red and blue line in Figure 3.10(b) correspond to the noise acquired at several weeks of interval. This noise is clearly not random, and shows a repetitive pattern every 16 ns. Moreover, the phase of the noise remains the same. The 16 ns period corresponds to the harmonic of the various clocks in the Flash-ADC card. This high-frequency noise is in fact a cross-talk between all the clocks and the analog input.

The slowest 16 ns oscillation corresponds to the main 62.5 MHz clock, and the faster frequencies are harmonics of the 500 MHz sampling clock.

3.5.c Test Results

The method previously used to estimate the linearity in [77] consisted in acquiring constant waveforms for each code. Linearity results with this method were biased by the digitization error, whereas the sinusoid method is not, since the waveform is sampled over many ADC codes, the true sinusoid distribution can be well reconstructed when fitting the full digitized waveform. The DNL measurement used to contain the digitization error and the actual DNL inherent to the Flash-ADC. Thus, the DNL and INL found were twice bigger than the ones I measured.

Most of the tested Flash-ADC complied with the manufacturer's specifications. The cards with the highest INL were kept as spares in case of a Flash-ADC failure during data taking. One card was rejected, due to an octet inversion in the data propagation within the card, and could be repaired by our laboratory's engineer.

The high frequency clock-induced noise does not have an impact on the PE peak in the signal as it gets integrated over many periods, but the baseline estimation has to be done over a multiple of 8 samples.

After testing the Flash-ADC, I went to the Near Hall to take part in the installation of the electronics, I installed the NIM and VME crates for the FEE and the Flash-ADC, as well as the Flash-ADC cards.

3.6 Charge Reconstruction with RecoPulse

As explained in Section 2.8.a, the energy reconstruction in Double Chooz relies heavily on the reconstruction of the charge collected by each PMT for a given trigger. Double Chooz's ability to reconstruct charge and convert it into a number of PE using the gain calibration drives the final energy resolution.

Calorimetry with PMT can be achieved in different manners:

- Counting the number of hit PMT. In a low luminosity experiment, where an average of less than 1 photon per PMT is expected, the number of hit PMT gives a direct measurement of the energy of the event.
- Integrating the full PMT signal when a trigger is sent to the Flash-ADC.
- Integrating the signal over a reduced time window, corresponding to the time of the signal. This method allows to reduce the non-linearity in charge introduced by electronic noise, by providing a better constraint on the baseline using the no-signal period in a waveform.

The latter is the method widely used in Flash-ADC based experiments. In Double Chooz, this charge reconstruction method is called *RecoPulse*.

3.6.a Baseline estimation

When storing the waveforms of a trigger, a pre-window of $\sim 50\,\text{ns}$ is stored before the beginning of the first pulse on the first PMT (grey stripped area in Figure 3.11). Every waveform thus has the first 50 ns free of signal, plus an additional time corresponding to the difference of time of flight that the photons have to cover to reach the various PMTs. On an event-by-event basis, the baseline is estimated by averaging the 8 first samples (green rectangle). The number of samples has to be a multiple of 8 because of the high frequency noise with a period of 16 ns induced by the cross-talk of the clocks inside a Flash-ADC (see Section 3.5).

Every second, a fixed rate trigger is sent to the Flash-ADC by the Trigger Master Board (see Section 2.6.c). Given the trigger rates of $\sim 500\,\text{Hz}$ in the FD and $\sim 300\,\text{Hz}$ in the ND, an average of $\sim 10^{-4}$ events are expected during a 256 ns integration window. Therefore, in the vast majority of case, the waveforms of those triggers are empty. The baseline of each channel

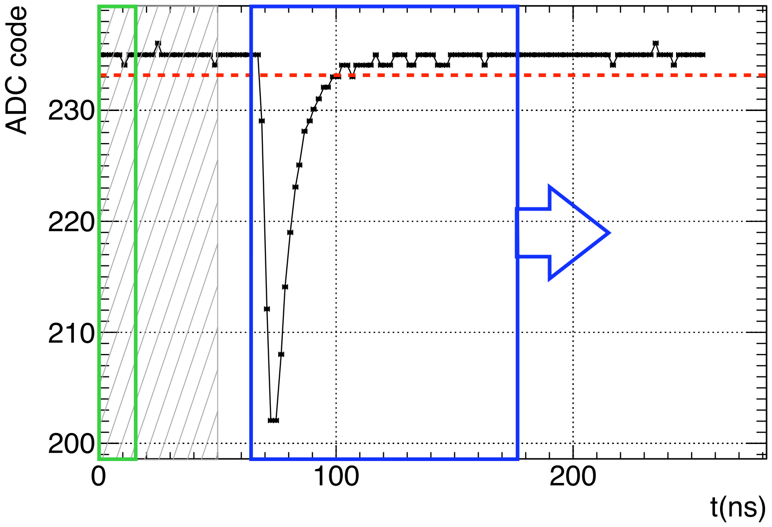


Figure 3.11: Charge reconstruction by RecoPulse. When a trigger is sent by the TMB, the waveform is registered with a 50 ns pre-window before the trigger time (grey striped area), providing a signal-free time during which the baseline is estimated over the first 8 time samples (green rectangle). The waveform is scanned by an integration window of 56 samples (blue rectangle) until the maximum integrated charge in the window found. Once the peak in this window is reconstructed, the process is iterated to the un-analysed part of the waveform until the average amplitude in the window is lower than the threshold (red dotted line).

is then estimated by averaging the full waveform, over the 128 samples.

In case the value of the baseline measured from the first 8 samples of a waveform is consistent with the latest fixed rate trigger estimation within 0.5 ADC code, the fixed rate value is kept to estimate the charge in the following procedure. Else, the event-by-event value is used.

As the fixed rate estimation of the baseline is done over the full waveform, its estimation is more precise. However, the event-by-event estimation is very useful as it provides a way to estimate the fluctuation of the mean value and the RMS of the baseline over time, and even allows for a real-time quality check of the waveforms, for which those values should remain within acceptable boundaries.

After a high energy event such as a muon crossing the detector, the PMTs take some time to recovers, and the baseline of triggers directly following a muon can be outside of the accepted range, or not flat, leading to a too high RMS of the first 8 samples. The estimation event-by-event of the baseline allows to flag those events as symptomatic.

The estimation of the waveform baseline is intrinsically biased by the Flash-ADC measurement. Since the amplitude of the analog noise is lower than 0.5 LSB, the digitization theorem in Section 3.4 states that the baseline cannot be perfectly reconstructed. The Flash-ADC produces a flooring of the analog value of the waveform, hence the measured baseline is consistently under-estimated with respect to the true analog value.

3.6.b Pulse Charge Reconstruction

Pulse reconstruction is achieved using a sliding window algorithm. The waveform is scanned by a fixed size window of 56 samples.
The first peak is found when the integral of the signal in the window is maximal. Then the process is iterated over the un-analysed signal until the integrated charge in the window is lower than a threshold value $Q_{\text{threshold}}$ with

$$Q_{\text{threshold}} = 5\ \sigma_{\text{baseline}} \cdot \sqrt{\text{WS}} \tag{3.3}$$

where WS is the integration window size in number of samples.

The integrated charges in each time windows are summed to obtain the total charge of the waveform.

As shown in Figure 3.12, because of the constant underestimation of the baseline, the reconstructed charge is misreconstructed by $\Delta Q = \Delta b \times T_{pulse}$ where Δb is the bias between the true baseline (red dashed line) and the reconstructed one (blue dashed line) and T_{pulse} is the pulse time $(T_{rise} + T_{fall})$. The charge non-linearity is the relative bias on the charge $\Delta Q/Q \propto T_{pulse}/Q$. This bias depends on the shape of the PMT pulse, e.g. for two PE pulses well separated (left), the charge non-linearity will be twice as big as the one in the case where the same two pulses arrive at the same time (right). Moreover, a pulse with a big charge has a smaller charge non-linearity than a pulse with a smaller charge, for roughly the same pulse duration.

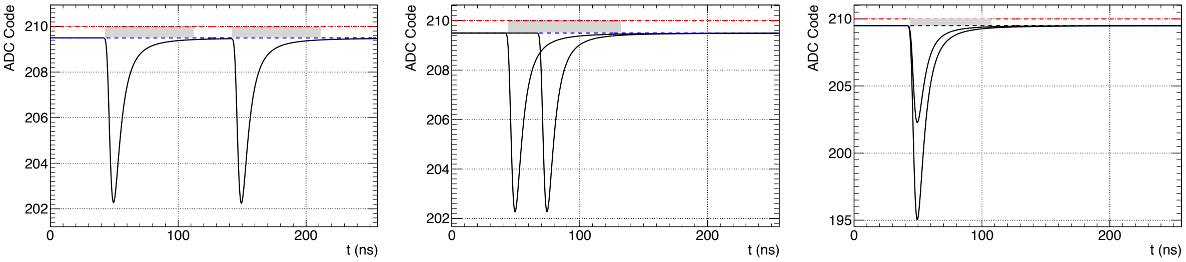


Figure 3.12: RecoPulse non-linearity source: The mismatch of the reconstructed baseline (dashed blue line) and the true baseline (red dashed line) leads to a bias in the charge reconstruction proportional to the bias on the baseline and the duration of the pulse (grey area). This non linearity depends on the waveform topology, i.e. the time and charge of each PE pulses.

Figure 2.11 shows the evolution of the gain with the charge of the waveform. The increase at low charge corresponds to the charge non-linearity becoming non-negligible at low charge.

Every time the Flash-ADCs are power-cycled, the reference voltage changes slightly, as does the true baseline position. Thus, the mean charge per PE (gain) and the charge non-linearity changes as well. This effect was presented earlier on in Figure 2.13 where the evolution of the energy prior to gain corrections presented steps correlated in time with Flash-ADC power cycles. The effect is more visible at low energies, with more waveforms at low number of PE, where the ratio $\Delta Q/Q$ is bigger.

3.6.c Pulse Time Reconstruction

A set of times is estimated for each pulse found:

- T_0 is the start time at which the integration window starts,
- T_1 is the end time of the integration window,
- T_{start} arrival time of the pulse, is computed as a constant fraction of 30% of the peak maximum amplitude,
- T_{end} is the time corresponding to the signal reaching 20% of its maximal amplitude after the maximum has been achieved,
- T_{max} is the time at maximum amplitude,
- T_{rise} is defined as $T_{\text{max}} - T_{\text{start}}$, and T_{fall} is defined as $T_{\text{end}} - T_{\text{max}}$.

Most of those times are defined from the baseline and as such depend on the knowledge of the baseline. Figure 3.13 shows the modification on T_{start} and T_{end} induced by the mis-reconstruction of the baseline. T_{start} is the least sensitive time as the PMT pulse have a steep rising edge. However, a PE with an amplitude smaller than 30% of the maximum peak amplitude could be ignored here, biasing the actual T_{start}.
T_{max}, however, is independent of the baseline but corresponds to the time of the maximum amplitude of the waveform, not necessarily the time of the first PE.

3.7 Charge Reconstruction with RecoZoR

3.7.a From a PMT to RecoZoR

The strength of RecoPulse is that it is shape independent. However the linearity of the charge reconstruction, especially at low number of PEs is limited by the under-sampling of the baseline.

The waveforms are composed of a succession of PEs. The contribution of a single PE has been studied in a dedicated setup at APC described in Figure 3.14.

Waveforms from a PMT set in a black box illuminated by a LED are acquired with a 8-bit, 1 GHz oscilloscope with a sampling frequency of 2.5 GS/s (LeCroy Teledyne HD0 6304).
The dynamic range of the oscilloscope is 80 mV, with an LSB of 0.3125 mV, which allows for a perfect digitization of the noise and baseline reconstruction, which is not possible with a Flash-ADC because of the much bigger dynamic range.

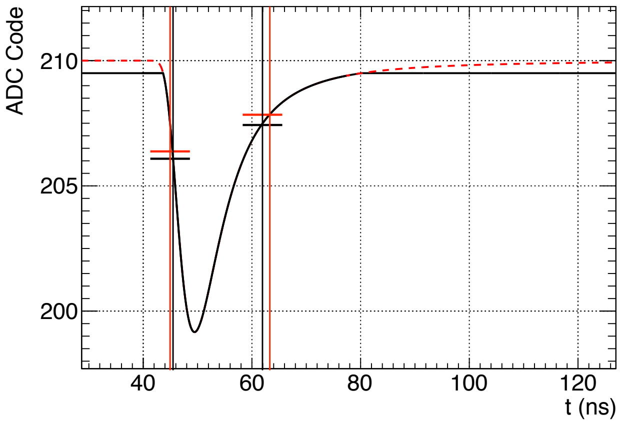


Figure 3.13: Impact of the mis-reconstruction of the baseline on the T_{start} and T_{end} reconstruction. Black lines correspond to the time determination with the mis-reconstructed baseline, red lines correspond to the ideal case where the baseline can be reconstructed exactly.

The signal cables used are spares from Double Chooz, with the same length to approach as much as possible the signal distortion of the *in situ* measurement.
The PMT is powered by a high voltage power supply so as to be operated with a gain of 10^7. The LED is set to low luminosity and delivers a trigger signal synchronous to the LED flashing. A delay is included in the trigger signal in order to take into account the length of the signal cables and have the PMT pulse synchronous to the trigger within 200 ns to minimize the possibility of accidental dark noise coincidence. Waveforms are stored for every trigger from the LED, and an average of 1.24% of the waveforms are not empty. This rate corresponds to an expectation of $1.24 \cdot 10^{-2}$ PE per waveform, and a probability of $7.7 \cdot 10^{-5}$ to have two or more PE per waveform. Thus, charges measured from those non-empty waveforms correspond almost exclusively to single PE waveforms.

Figure 3.15(a) shows the charge integrated from the acquired waveforms from 50 ns to 100 ns, with the PMT pulse time being 60 ns, as :

$$Q = \int_{50}^{100} \frac{V(t)\mathrm{dt}}{R\, G_{\text{FEE}}\, e} \tag{3.4}$$

where $V(t)$ is the value of each sample of the waveform, G_{FEE} is the gain of the FEE, R is the equivalent resistance of the PMT, and e is the elementary charge $1.6 \cdot 10^{-19}\, C$. The distribution can be described by two Gaussians and an exponential. The first Gaussian (black dashed line), centred on zero corresponds to the integral of the pedestal, and is simply the integrated fluctuations of analog noise on empty waveforms. The second Gaussian (blue solid line) corresponds to the charge of the PE signal and is centred on the PMT gain. The exponential (green solid line) models the PMT thermal emission, the effect by which an electron gets emitted from one of the dynode and triggers a amplification avalanche.

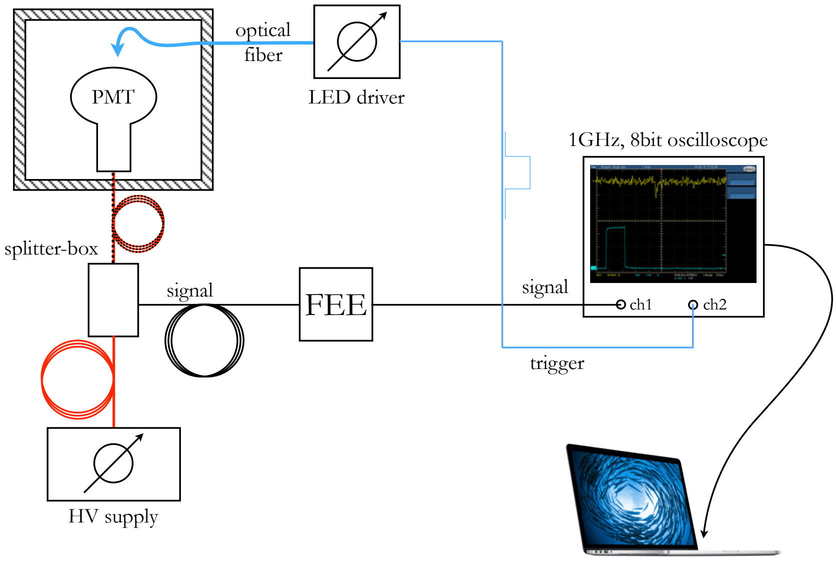


Figure 3.14: Acquisition of waveforms from a low-intensity LED.

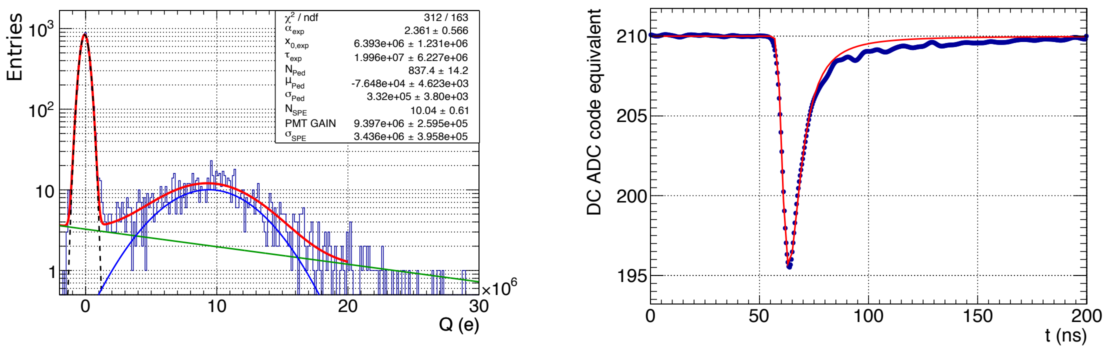
Figure 3.15(b) shows the average 1PE waveform obtained by averaging the waveforms with charge within $\pm 1\sigma$ of the gain measured in Figure 3.15(a). The single PE peak can be described by a Landau function (red line) with a good approximation; the absolute difference between the Landau and the actual waveform does not exceed 1 LSB of Double Chooz's Flash-ADC and is washed out by the analog noise and the quantization effects.

If a single PE peak can be described as a Landau, then a multiple PE waveform $s(\#PE)$ with $\#PE$ PE could be described by the sum of multiple Landaus corresponding to each PE present in the waveform:

$$s(1PE) = B - L(q_1, t_1, \sigma_1) \tag{3.5}$$

$$s(\#PE) = B - \sum_{i=1}^{\#PE} L(q_i, t_i, \sigma_i) \tag{3.6}$$

where $\#PE$ is the integer number of PE in the waveform, by opposition to N_{PE} which is a real number derived from the integrated charge of the waveform in RecoPulse, B is the baseline, q_i, t_i and σ_i correspond to respectively the charge, time of maximum amplitude, and rise-time of each PE i. The waveform can thus be fully described by a set of parameters of baseline, charge and time for each PE : $s(B, \vec{q}, \vec{t}, \vec{\sigma})$. The rise time σ of a PE pulse depends on the geometry of the dynodes, and can vary from one PMT to another, however, it was found to be stable for a given PMT and can therefore be fixed when fitting the waveform. Hence, the function is $s(B, \vec{q}, \vec{t})$.



(a)

(b)

Figure 3.15: 3.15(a): Integrated charge distribution, in units of the elementary charge $e = 1.6 \cdot 10^{-19} C$. The peak centred in zero corresponds to the pedestal, the integrated noise over the waveform. The exponential models dynode thermal electron emission. The second peak corresponds to the PE charges. The PMT is operated at a gain of $(9.39 \pm 0.26) \cdot 10^6$.
3.15(b): Mean 1PE waveform obtained by averaging waveforms acquired with a 8-bit, 1 GHz, 2.5 GS/s oscilloscope, with an illumination of $1.24 \cdot 10^{-2}$ PE per waveform. Averaged waveform must have a charge within $\pm 1\sigma$ of the gain measured in 3.15(a)

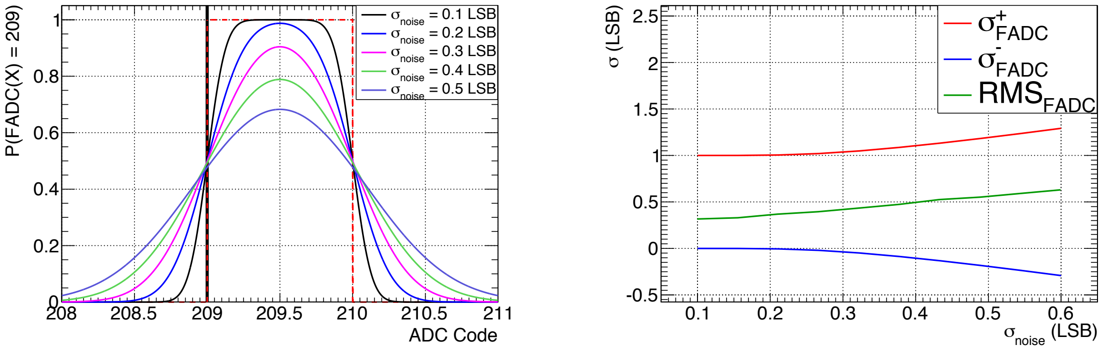
Fitting the waveform with this description adds new information in the waveform reconstruction: the knowledge of the shape. The fit results include the baseline determination, the number of PE, the individual charge and hit time of each PE. In addition, the baseline is estimated over the full waveform, including the PE peaks, which are well digitized. The resolution of the waveform on an event-by-event basis is thus expected to be greatly improved compared to RecoPulse.

This reconstruction method is doing the opposite of Double Chooz's Read Out System Simulation (ROSS) that start with the information of each PE and simulates the behaviour of the PMT and the full acquisition chain to get the simulated digitized waveform. Thus, the reconstruction has been named RecoZoR (ROSS ⇒ SSOR ⇒ ZOR).

3.7.b Flash-ADC Measurement Error

In order to fit properly the waveform, attention must be paid to the errors attached to each sample. Fitting with no error would lead to an impossibility to estimate the goodness of the results. *A priori*, the measurement from the Flash-ADC is absolute and no error is associated to that. However, the Flash-ADC, by construction, returns an integer value, realizing a flooring of the actual analog value. When the Flash-ADC returns an ADC code N, the true analog value of the waveform has a uniform probability of being anywhere within the codes $[N, N+1]$.

Figure 3.16(a) shows the pdf of the true signal position for which the Flash-ADC output is a given ADC code (here 209) for various analog noise values and assuming a perfect Flash-ADC with no DNL (i.e. a code size of exactly 1 LSB). For a noiseless signal, the probability is uniform within the code, no a priori position can be postulated. As the noise level increases, signal with true average value outside of the code can fluctuate inside because of the noise.



(a) (b)

Figure 3.16: 3.16(a): Probability for a true baseline at X that the Flash-ADC returns the code 209 for various analog noise levels. For a noiseless signal, the probability to get a given code N is the same for any baseline between N and N+1 and is zero outside (red dashed line). As a Gaussian noise is introduced, a probability for the signal to fluctuate into a lower or higher code appears.
3.16(b): From the previous pdf we can define the error as the values for which the pdf reaches half of its maximum. The red and blue line respectively show the higher and lower boundaries of the error as a function of the noise level. The green line shows the RMS of the pdf.

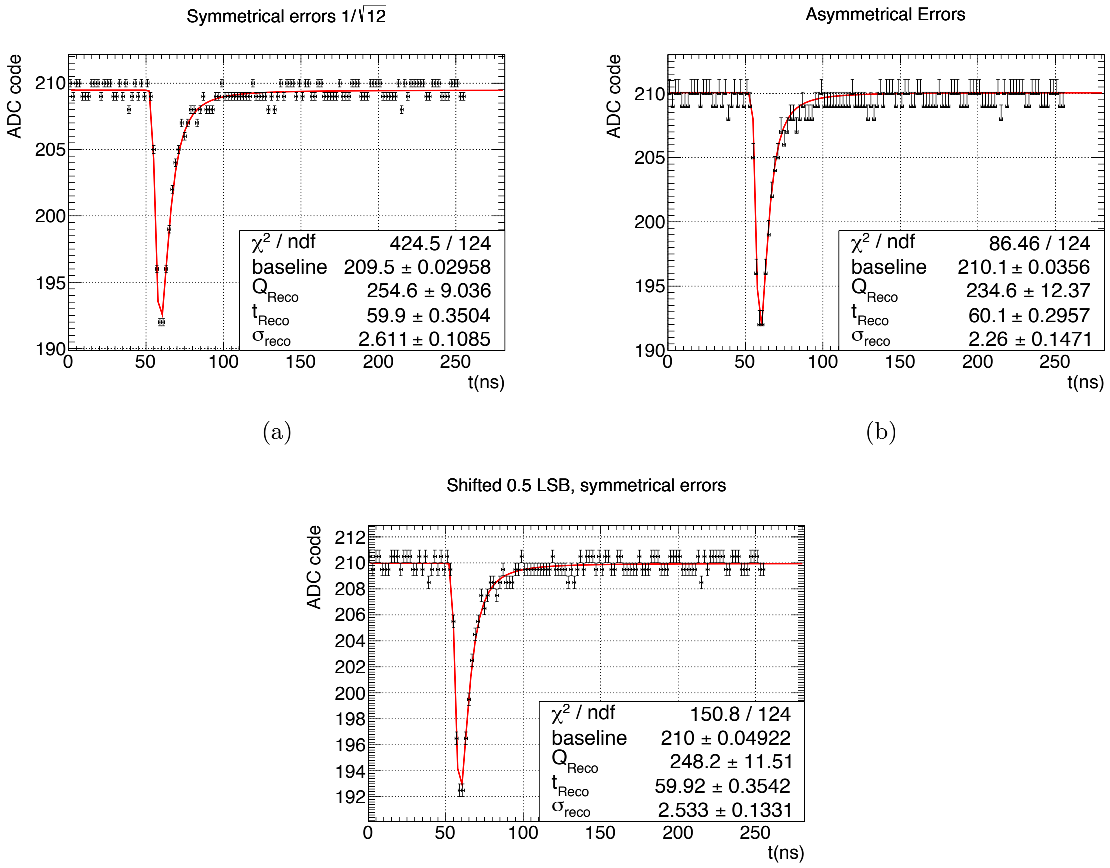
Inversely, signal with true average value inside the code can fluctuate outside. The pdf is a convolution of the gate function of the Flash-ADC and the Gaussian distribution of the noise as described in Equation 3.7. With increasing noise levels, the pdf becomes closer to a Gaussian.

$$\text{gate(LSB)} \otimes \text{gaus}(0, \sigma_e) = \int_0^{LSB} \frac{1}{\sigma_e\sqrt{2\pi}} e^{\frac{-(t-x)^2}{2\sigma_e^2}} \text{dt} \tag{3.7}$$

We already saw this behaviour in the baseline determination on Section 3.4, and it needs to be taken into account when fitting the waveform, as the signal that was digitized at a code N is not necessarily at the true value of N, and does not have symmetrical probability to be higher or lower than N (i.e. error bars should not be centred on N and symmetrical).

A better error description could be to take the limits of the pdf at half maximum to define σ^{+}_{FADC} and σ^{-}_{FADC}. Those values would be respectively -0 and $+1$ at low noise level, where the pdf can still be described mostly by a gate function. As the gain increases however, so must the error bars. This behaviour is illustrated in Figure 3.16(b), where σ^{+}_{FADC} and σ^{-}_{FADC} are respectively the red and blue lines.

Another possibility is to remark that the pdf is symmetrical, with a mean value of $N + 0.5$. The full waveform could be shifted by 0.5 ADC and be attributed symmetrical error bars of the RMS of the pdf. The evolution of the RMS of the Flash-ADC output is represented Figure 3.16(b) as the green line. At low noise level, the RMS tends to $1/\sqrt{12}$ as the pdf tends to a gate function of width 1 LSB.



(a) (b)

(c)

Figure 3.17: Simulated single PE waveform with an analog noise of 0.37 LSB and a charge of 150 DUQ.
3.17(a): Symmetrical error of $1/\sqrt{12}$ LSB on the digitized output to account for the Flash-ADC flooring.
3.17(b): Asymmetrical error of $^{+1.1}_{-0.1}$ LSB on the digitized output.
3.17(c): Symmetrical error of 0.49 LSB (RMS of the pdf for a noise level of 0.37 LSB) on the digitized output shifted by 0.5 LSB.
Each waveform is fitted with a single PE function.

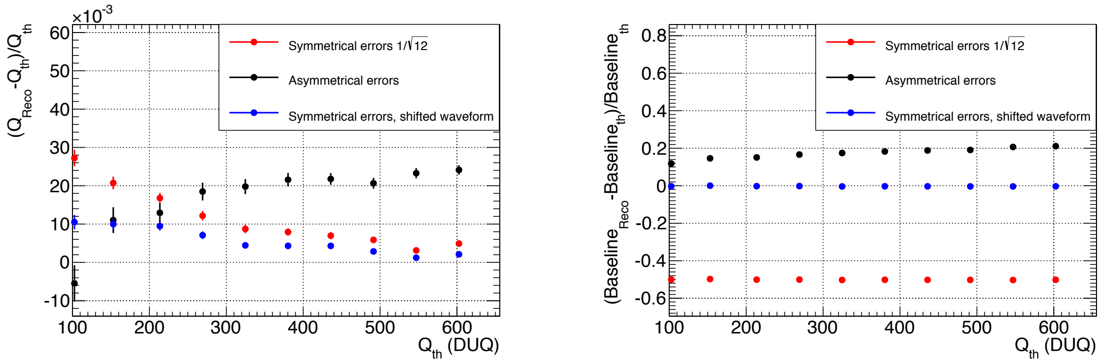
Figure 3.17 shows a simulated single PE waveform with an analog noise of 0.37 LSB, to be in the noise configuration of FD-II and ND, and a charge $Q_{th} = 250$ DUQ. The errors associated to the Flash-ADC measurement have been treated in three different ways, applied on the same waveform:

- 3.17(a): Symmetrical error of $1/\sqrt{12}$ LSB on the output value of the Flash-ADC, to account for the flooring of the Flash-ADC only,

- 3.17(b): Asymmetrical error of $^{+1.1}_{-0.1}$ LSB on the output value of the Flash-ADC,

- 3.17(c): Symmetrical error of 0.49 LSB on the output value of the Flash-ADC shifted by 0.5 LSB to account for the position of maximum probability.

Each configuration is fitted with a single PE function $s(1PE)$ as defined in Equation 3.5. The χ^2/NDF is greater in the symmetrical, $1/\sqrt{12}$ LSB error, description than in the shifted, 0.49 LSB error, description as the errors are under-estimated in the first description because of the noiseless hypothesis. On the contrary, the χ^2/NDF is too low in the asymmetrical description because the errors are overestimated. Moreover, the fit expects a Gaussian error definition, which is not the case here, thus this error definition is not suitable.



(a)

(b)

Figure 3.18: 3.18(a): Non-linearity of reconstruction on the 1000 single PE waveforms for each of the three error handling methods as previously described.
3.18(b): Baseline bias in the reconstruction of waveforms for each of the error handling methods.

Figure 3.18(a) shows the evolution of the non-linearity with the charge of a single pulse for each of the three error handling configurations. Although the charge of a single PE cannot go up to 600 DUQ in Double Chooz, it is probable enough that multiple PE could arrive in a very short time interval, as shown in Section 3.7.d, which would give a similar behaviour.
The linearity of the two symmetrical error configurations evolve identically at high charge, but the error description using the RMS of the pdf and shifted values is more linear at low charge. The linearity of the asymmetrical errors is worse, however, than the symmetrical case at high charge.

Figure 3.18(b) shows the evolution of the baseline with the charge. Here again, the symmetrical description shows a more constant behaviour than the asymmetrical one. The configuration with shifted value gives a baseline close to the true value, which is a direct consequence of the 0.5 LSB shift.

From now on, the error description will be symmetrical, using the RMS of the pdf on figure 3.16(a) and the digitized waveform values will be shifted by 0.5 LSB. The linearity and baseline consistency are not impacted by the shift, but the reconstructed baseline is closer to its true value.

3.7.c Description of the Algorithm

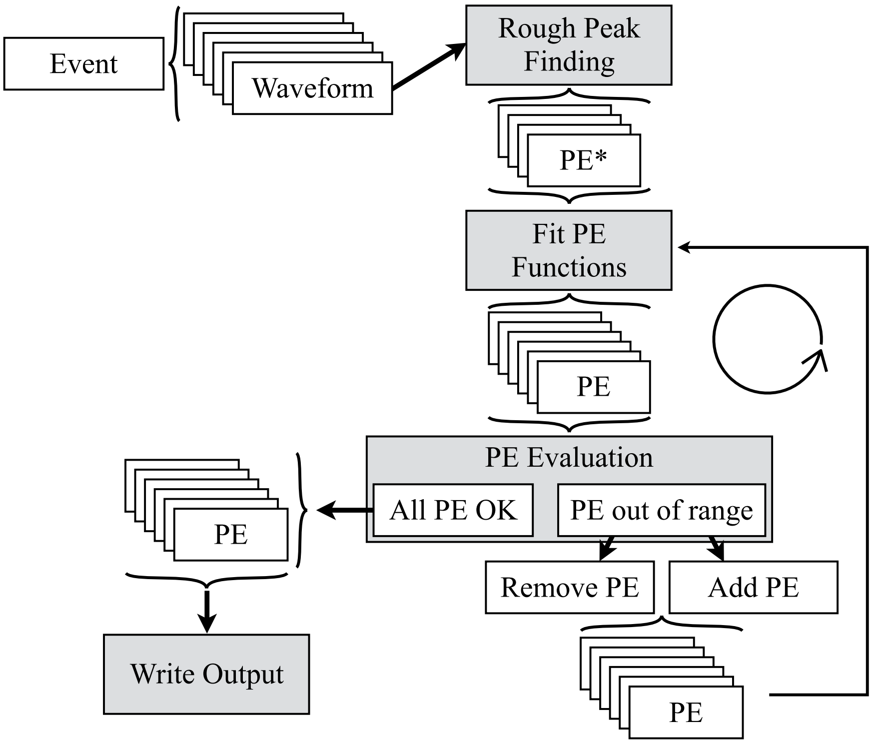


Figure 3.19: RecoZoR organizational sketch.

Figure 3.19 presents an organization diagram of RecoZoR: An event in Double Chooz is a collection of waveform, each waveform is treated separately by the algorithm. The goal is to break down the waveform into a series of PE, each characterized by a time and a charge. The output of RecoZoR consists in a total number of PE ($\#PE$), with the time and charge for each, but also the total charge of the waveform, defined as the integral of the function s($\#PE$) and the time of the first PE. The integrated charge allows to reconstruct a continuous PE estimator similar to the PE reconstruction by RecoPulse called N_{PE}.

The waveform is shifted by 0.5 LSB and is being attributed an error of $\sigma_{FADC} = 0.42\,\mathrm{LSB}$ if in the FD-I configuration, corresponding to a 0.3 LSB analog noise level, and $\sigma_{FADC} = 0.49\,\mathrm{LSB}$ if in the FD-II or ND configuration, corresponding to a 0.37 LSB analog noise level.

First, a peak finding method is applied to the waveform to determine how many peaks are present, i.e. to give the minimal number of PE needed to describe the waveform. Indeed, there can be more than one PE per peak, but no less than one. Each peak is treated as a PE (let's call them PE*). They form the first, rough, collection of PE.

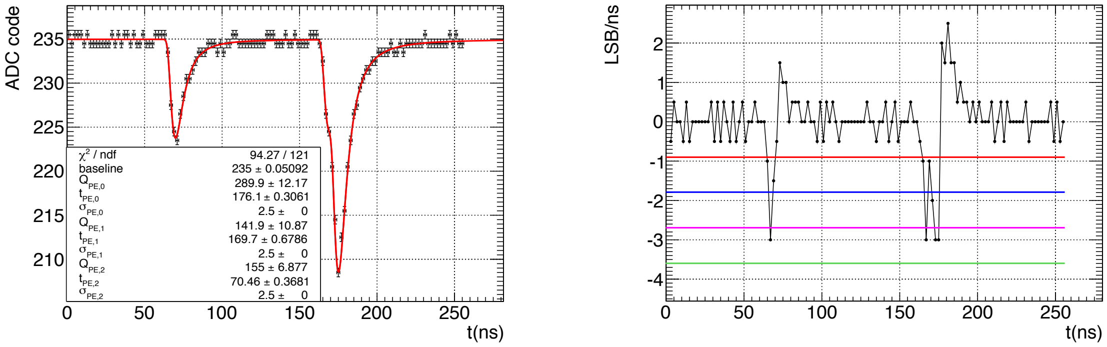
At this point, the algorithm enters a loop which performs three operations:

- fit the waveform with the current list of PE

- evaluate the sanity of the fit result, i.e. are the found PE consistent with the gain of the PMT, and is the waveform described properly by the PE list?
- if necessary, modify the PE list by adding or subtracting PEs.

when no more PE are out of range, or a maximum number of iteration is reached, the loop is exited and the list of PE and waveform-wise reconstructed variables are written as output of the reconstruction.

Peak Finding could be achieved with the same method as RecoPulse's sliding window, with a threshold on the integrated charge on a given window. However, having a threshold would imply missing pulses with lower charges, and thus an inefficiency at low energy.
To address this issue, we use the derivative of the waveform. Indeed, the derivative holds the information on the shape of the pulses. random fluctuation from the noise do not exceed a ±0.5 LSB /ns as they are fluctuations of only one ADC for one sample. The PE pulse however, have a very steep rising edge, and thus their derivative neatly contrasts, even more so when multiple PE pile up in a short time period. A threshold in the derivative space translate in a shape discrimination between the pulse and the noise fluctuation, rather than a charge criterion. More over, when multiple PE pile up, there is a higher probability that the derivative will go back to zero, marking more clearly different pluses than the waveform itself.
The threshold on the derivative is set to be -0.9 LSB /ns and allows to detect peaks that vary by more than 2 ADC codes in one sample.
When two PE are close enough, the derivative can remain negative between two peaks and not go back to zero. To address this behaviour, the threshold is risen by increments of -0.9 LSB /ns until no peak is found as materialized in Figure 3.20(b), where the threshold evolves from the red line to the green line. The number of peak corresponds to the maximum number found when rising the threshold.



(a) (b)

Figure 3.20: 3.20(a): Simulated 3PE waveform Fitted s(3PE) function.
3.20(b): Derivative of the simulated waveform, two peaks are easily identifiable. Allowing the threshold to evolve from the red line to the green line allows to find three peaks, providing a better starting point for RecoZoR.

Figure 3.20(a) shows a simulated waveform with RecoZoR reconstructed shape and Figure

3.20(b) its derivative. Three PE are present in the waveform, with only two pulses. A simple threshold on the waveform would only have found two peaks as would a fixed threshold on the derivative have. Allowing the threshold to be risen allows to find three peaks, thus giving a better starting point for the rest of the reconstruction.

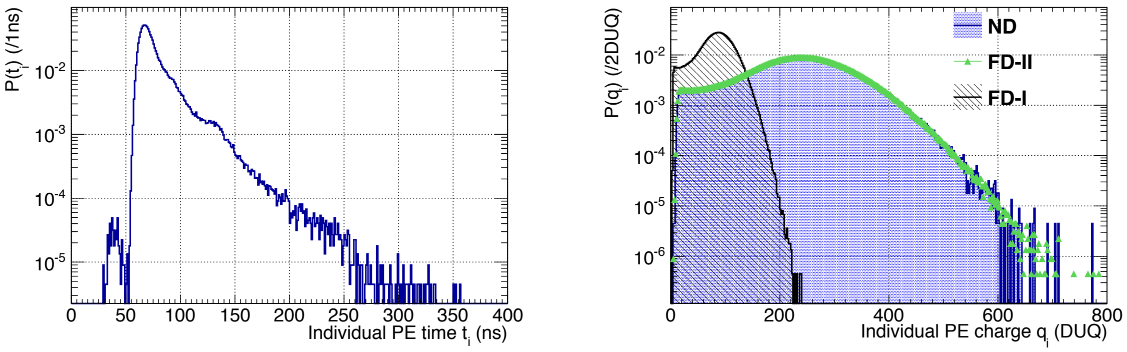
The output of this step consists in a list of pulses defined by their integrated charge, and the start and stop time of each pulse. Those parameters are used to initialize the function used in the first iteration of the fitting loop.

Fitting the waveform is performed using the Minuit minimization package and a χ^2 minimization method. The parameters are allowed to vary within their authorized range with a flat prior.
The charge of each is allowed to vary between $0.1 \cdot G$ and $1.9 \cdot G$ where G is the gain of the PMT (in DUQ/PE).
The time of the maximum amplitude of each PE is allowed to vary between the start and stop times of the pulse found in the first step in the first iteration of the loop, and within $\pm 2\sigma$ for the other iterations.
The rise time σ is fixed *a priori* for each PMT. When reconstructing ROSS generated waveforms, σ is fixed to 2.5 ns as it is the value used in the simulation.

Status Evaluation is achieved with two methods. First, if the charge of a $\mathrm{PE}(q_0, t_0)$ is at the limit of $1.9 \cdot G$, the PE is divided in two, the two new PEs are set to $\mathrm{PE}(q_0/2,\ t_0 - 2\,\mathrm{ns})$ and $\mathrm{PE}(q_0/2,\ t_0 + 2\,\mathrm{ns})$. Else, if the charge of a PE is at the lower limit of $0.1 \cdot G$, the PE is removed for the next iteration.
Second, the residual of the waveform with respect to the fitted function (once the previous modifications have been realised) is obtained. If the value of the residual is lower than a threshold of $2\sigma_{FADC}$, or the value residual is lower than σ_{FADC} for more than 2 consecutive samples, a PE is added at this position with a default charge of $0.2 \cdot G$.

The Time and Charge pdf can also be used to determine the likelihood of the reconstruction. The time pdf, shown in Figure 3.21(a), corresponds to the scintillation response of Double Chooz's liquid scintillator and can be modelled to first order by the convolution of a Gaussian and the sum of three exponentials [81]. In addition to this scintillation behaviour, pre- and late-pulses are added. Pre-pulses correspond to photons that cross the photocathode without producing PE but hit the first dynode, producing a primary PE from the cathode, and effectively missing an amplification stage. These PE usually have a smaller charge and are not triggered on. The late-pulses correspond to an ion being knocked out of the photocathode as well as a PE. The ion then slowly drifts back to the photocathode and may knock an electron from the photocathode, producing a later PE pulse.

The charge pdf, shown in Figure 3.21(b) for FD-I and ND configurations of ROSS, corresponds to the charge fluctuation of a single PE and have been extensively studied in [75, 76] and discussed in Section 3.7.a, Figure 3.15(a). Here only the 1PE peak is simulated, the pedestal is not shown as it corresponds to no PE in the waveform. For each waveform, the number of generated PE is chosen randomly to follow the Poisson distribution centred on the true expected number of PE for the energy of the generated event from MC simulations.



(a) (b)

Figure 3.21: 3.21(a): Time pdf of PE. To the exponential behaviour from the scintillator are added features such as pre-pulses (before 50 ns) and late-pulses (around 135 ns).
3.21(b): Charge pdf of PE for each detector configurations. Here only the 1PE peak is simulated, the pedestal is not shown as it corresponds to no PE in the waveform. PMT are operated at the same gain of $\sim$ 240 DUQ /PE in FD-II and ND and at a gain of 88 DUQ /PE in FD−I.

Unfortunately, due to time constraints, the use of the pdfs has not been implemented in the algorithm.

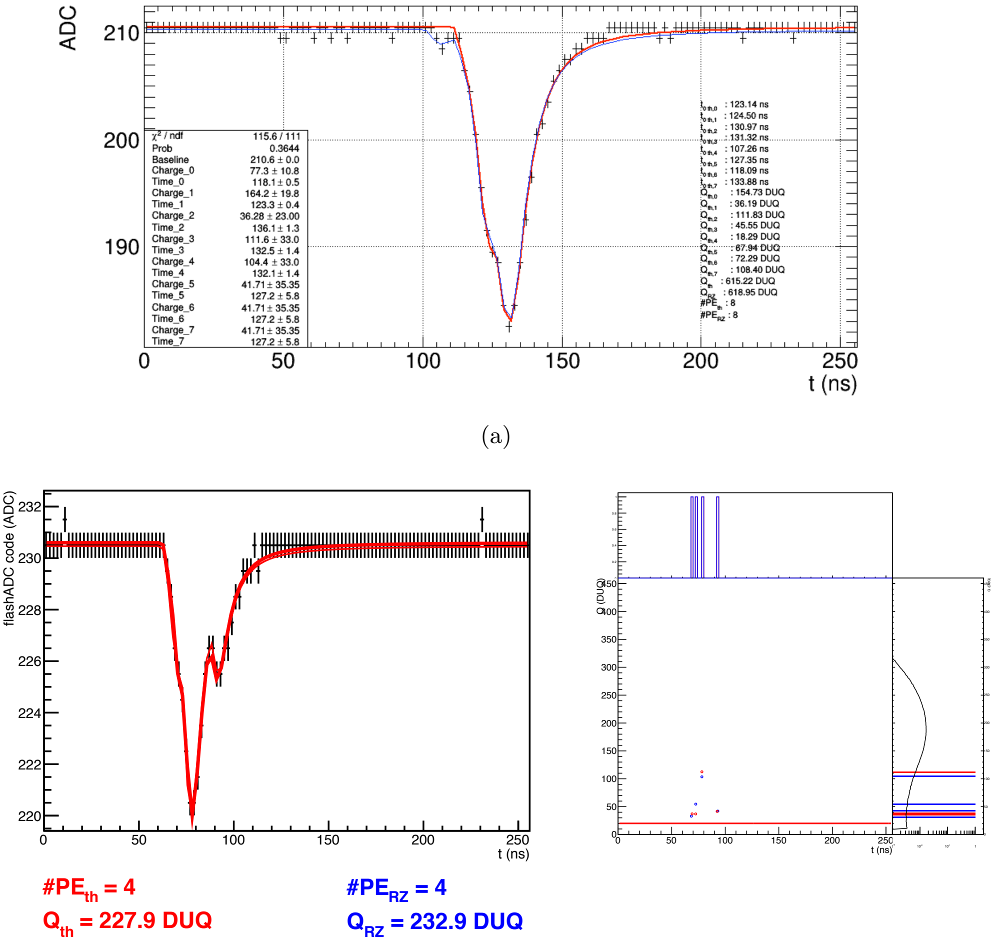
Output The output of RecoZoR for each waveform consists in the value of the baseline, an ordered list of PE with their hit time, charge and rise time, as well as the summed charge of all the PEs in the waveform, the number of PE used to fit ($\#PE$), the charge-estimated number of PE (N_{PE}), and a likelihood value, computed from the time and charge pdf.

3.7.d Reconstruction of Simulated Events

RecoZoR has been developed and evaluated on MC simulated events. Attention was especially paid to charge linearity as it is directly comparable to RecoPulse without having to worry about the definition of linearised PE described in Section 2.8.a.

Figure 3.22(a) shows a 8 PE waveform reconstructed by RecoZoR. The blue line corresponds to the true, noiseless, waveform and the red line corresponds to the reconstructed waveform. RecoZoR misses the first PE at 107 ns because it has a very low charge. However, several virtual PE are reconstructed with the same time and charge, and RecoZoR reconstructs the correct number of PE. Moreover, the reconstructed charge agrees with the true charge within 0.5%.
Figure 3.22(b) shows a 4 PE waveform reconstructed by RecoZoR. The thin red line corresponds to the true, noiseless waveform, and the thick line corresponds to the reconstructed waveform. Here the number of PE is consistent as well with the true number, and the reconstructed charge agrees with the true charge within 2%.
Figure 3.22(c) shows another way to display the information on the same waveform as Figure 3.22(b), that allow to check the consistency of the reconstructed PEs. The horizontal axis represents the time at which the PEs are reconstructed, and the vertical axis represents the charge of each PE. Red dots are the true PEs and blue dots are the reconstructed PEs. Time and

charge informations are projected on the sides of the graph. The projected charge information is compared to the charge pdf for a single PE (black line).



(a)

(b)

(c)

Figure 3.22: 3.22(a): Reconstructed 8 PE waveform. The blue line shows the true analog, noiseless, waveform before digitization, points are the digitized waveform, and the red line shows the reconstructed waveform from RecoZoR.
3.22(b): Reconstructed 4 PE waveform. Here the thin red line corresponds to the true analog, noiseless, waveform, and the thick red line the reconstructed RecoZoR output.
3.22(c): Another way of comparing the true waveform and the reconstructed one is to look at the time and charge distribution of the reconstructed PEs. Here the time is displayed on the horizontal axis, and charge on the vertical axis. The red points correspond to simulated PEs, and the blue ones correspond to the reconstructed ones. The 2D plot is projected on each axis, and the charge distribution is compared to the expected pdf.

Baseline Reconstruction Figure 3.23 shows the difference between the reconstructed and the true baseline. Baselines reconstructed by RecoPulse (green quantized histogram) are here reconstructed by averaging the first eight samples, on an event-by-event basis as described in Section 3.6.a. They are shifted in average by 0.5 as expected from digitization flooring effects. The resolution on the baseline from RecoPulse does not depend on the number of PE as they are estimated on a fixed number of samples, with no signal at this point.

Baselines reconstructed by RecoZoR are estimated over the full fitted waveform and the resolution depends on the number of PE. Contributions of various PE number (0,1,2 and 3) are shown in separate histograms (respectively in black, red, green and blue). The integrated effect from 0 to 10 PE is shown in the pink histogram. Each of RecoZoR histogram is closer to the true baseline than RecoPulse's estimation, as a result of the 0.5 LSB shift applied to the full waveform. Furthermore, the RMS of RecoZoR-reconstructed waveforms is significantly better than RecoPulse-reconstructed baselines.

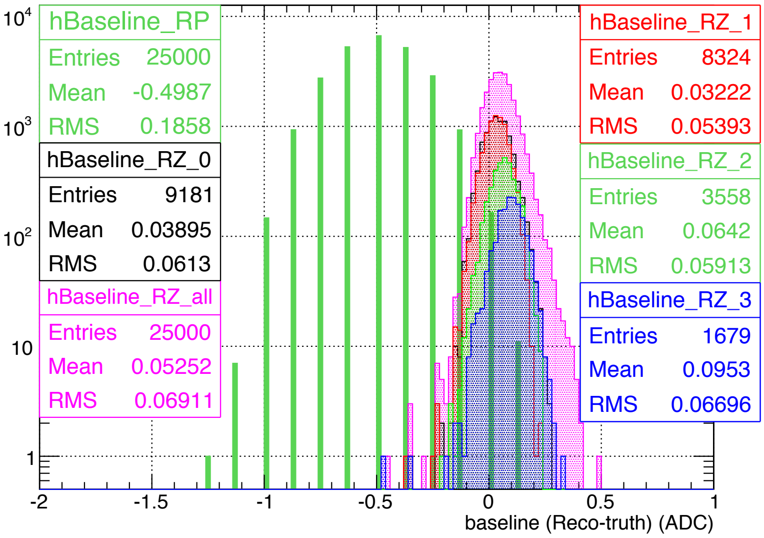


Figure 3.23: Difference between the reconstructed and the true baseline for RecoPulse (in green) and RecoZoR for waveforms with 0, 1, 2, 3 and up to 10 PE (respectively in black, red, green, blue and pink). RecoZoR baseline is consistently closer to the true value, and with a much smaller dispersion of the reconstructed value.

Table 3.1 shows the number of samples on which RecoPulse should evaluate the baseline in order to get the same resolution as RecoZoR.
For empty waveform (0PE), RecoZoR performs a fit by a constant function, that is equivalent to averaging the full waveform. This is equivalent to the baseline reconstruction by RecoPulse on fixed rate triggers, every second. Thus RecoZoR is equivalent to RecoPulse with 128 samples, but this can be done on every empty waveform.
Waveforms with 1 or 2 PE add a new information to the fit: the PE shape, and especially, the "ground level" of the PE. As a consequence, the waveform is reconstructed more accurately (equivalent to RecoPulse on more samples). This information degrades as the number of PE increases since the charge of each PE is able to vary and PE piling up may mimic a different

"ground level".
However, the equivalent number of PE remains well above the eight samples used by RecoPulse, hence providing, on an event-by-event basis, a much better, and more consistent, estimation of the baseline.

Table 3.1: Number of samples on which the baseline must be averaged to obtain the same performance as RecoZoR.

#PE	RecoZoR baseline RMS	N_{samples} RecoPulse equivalent
0	0.061	128
1	0.053	169.56
2	0.059	136.83
3	0.066	109.34
0→10	0.069	100.04

Charge Linearity with RecoZoR Charge non-linearity is estimated as the deviation from zero of the average of the residual charge between the reconstructed charge and the true simulated charge:

$$(Q_{Reco} - Q_{th})/Q_{th} \tag{3.8}$$

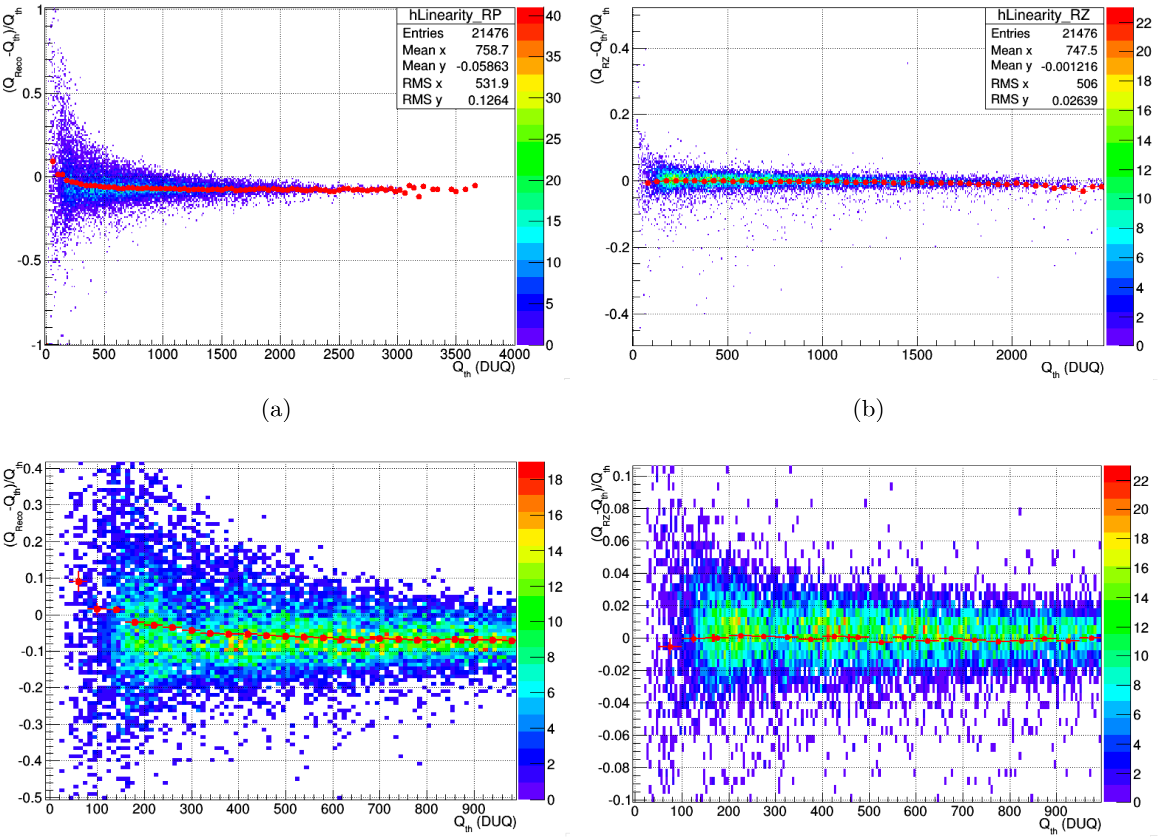
Figure 3.24 shows the evolution of the linearity with the charge of the waveform for RecoPulse (left) and RecoZoR (right). RecoPulse misses a constant fraction of the charge ($\sim$ 10%) at high charge. However, this constant fraction can easily be calibrated out. However, RecoPulse's non-linearity increases at low charge, up to 20% in amplitude from the high charge behaviour.

RecoZoR on the other hand shows a very stable linearity, below the percent level, over the full range of charges. RecoZoR is also more consistent than RecoPulse, as the RMS of the linearity is of the order of 12% for RecoPulse, and 3% for RecoZoR.

Reconstruction of the Number of PE RecoZoR being very linear, the estimation of the number of PE N_{PE} in a waveform can be achieved by simply dividing the obtained charge by the PMT gain. The linearity of such a reconstruction will be the same as the charge, since it only consists in dividing by a constant value. The estimation of N_{PE} for RecoPulse however is more complex, due to its non linearity at low charge. The reconstruction of N_{PE} is described in Section 2.8.a and relies on a non-constant gain function.

Another estimator on the number of PE in RecoZoR is the actual number of PE contribution fitted on the waveform #PE, i.e. the number of Landau functions used to describe the waveform. Figure 3.25(a) shows the evolution of the residual number of PE defined in Equation 3.9 with the true number of PE present in the waveform.

$$\frac{\#\text{PE}_{RZ} - \#\text{PE}_{th}}{\#\text{PE}_{th}} \tag{3.9}$$



(a)

(b)

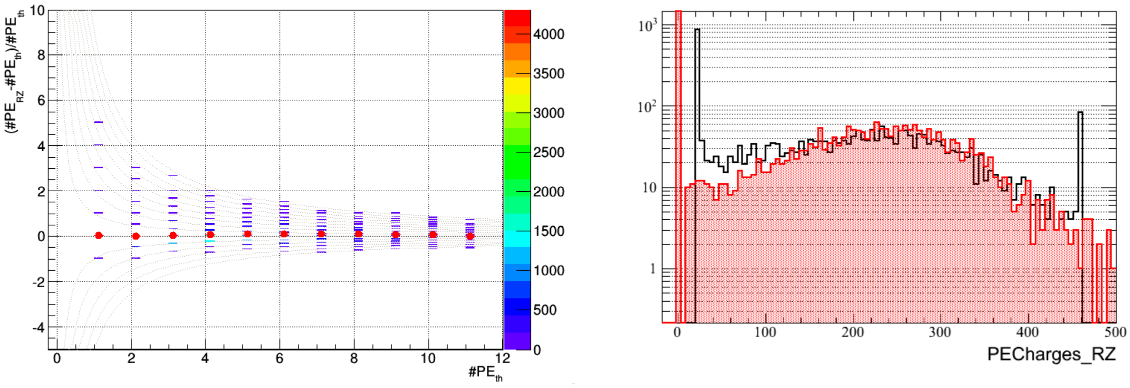
(c)

(d)

Figure 3.24: Charge linearity comparison between RecoPulse and RecoZoR:
3.24(a) and 3.24(b): Evolution of the charge residual between the reconstructed charge and the true simulated charge of the baseline : $(Q_{Reco} - Q_{th})/Q_{th}$ vsQ_{th} for RecoPulse (a) and RecoZoR (b).
3.24(c) and 3.24(d): Above distribution zoomed in on the 0 → 4 PE region for RecoPulse (c) and RecoZoR (d).

The values line-up on a family of N/x functions where N is number of additional, or missing, PE used to fit the waveform. Red points represents the mean value for each number of #PE$_{th}$. In average, up to 7 PE, a mistake of less than one PE is made.

Several effects can lead to a mis-reconstructed number of PE. Due to the time pdf of the light emission by the scintillator, presented in Figure 3.21(a), being very peaked at the beginning of the pulse, there is a high probability of multiple PE piling-up at the beginning of the pulse. When PEs pile up, it becomes more difficult to separate them, and it is possible to only reconstruct one bigger PE, or multiple smaller PEs. As a result of the fitting process, it would be advantageous to fit a high number of small PE, thus describing perfectly the waveform. Abnormally large noise fluctuations can be mis-understood as an additional, low charge, virtual PE. This is especially possible at the falling tail of a pulse, when the signal is slowly crossing the ADC code thresholds.



(a)

(b)

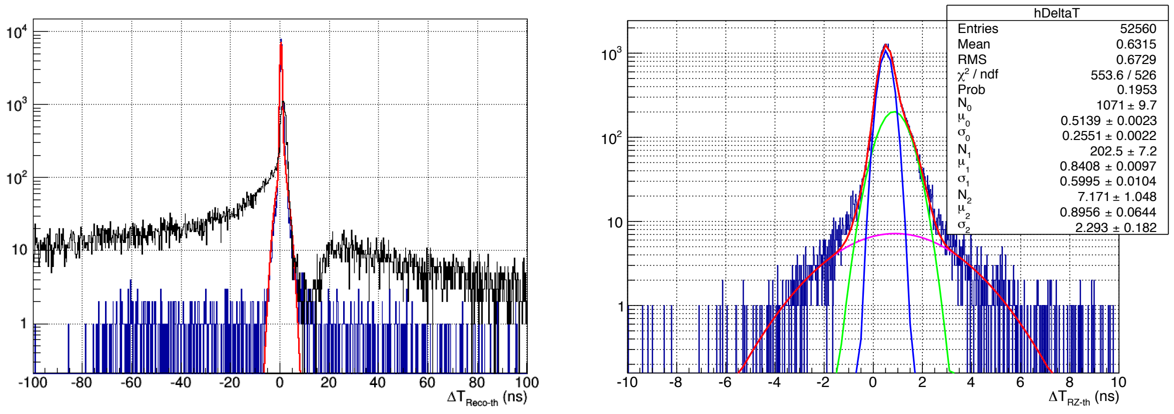
Figure 3.25: 3.25(a): Evolution of the residual number of PE reconstructed by RecoZoR with the true number of PE present within the waveform.
3.25(b): Charge per reconstructed PE (black histogram) compared to the charge per simulated PE (red histogram).

These effects are visible when comparing the distribution of charge per PE for the simulated PE and the reconstructed ones on Figure 3.25(b). The PE resulting of a fusion of two smaller ones will inflate the high-charge part of the distribution, and the one that results from an artificial creation of PE will inflate the lower-charge part of the distribution.

Time Resolution RecoPulse start time is dependent on the baseline determination, and can be biased by low charge PE with an amplitude lower than 30% of the maximal amplitude of the peak. RecoZoR start time however, is the time of the first reconstructed PE. It can be biased by the creation of virtual PE before the actual time of the pulse, or missing the first PE.
The division of a high-charge PE into two low-charge PE is transparent to the start time since the two PE will have the same time, as is the case for the fusion of two small PE.
Missing the first PE of the pulse, given the peaked pdf increasing the PE pile-up at the beginning of the pulse is very unlikely. Thus only PE added by noise fluctuation before the pulse can bias the start time.

Figure 3.26(a) shows the difference between the reconstructed T_{start} and the true hit time for RecoPulse (black line) and RecoZoR (blue line). The RecoZoR distribution shows much less spread than the RecoPulse one. The central value is fitted with three Gaussians as detailed in Figure 3.26(b), the blue Gaussian represents well reconstructed PEs. The pink Gaussian represents mis-reconstructed virtual PEs from noise fluctuation. Finally, the green Gaussian represents two close PEs reconstructed as only one, which can have the effect of slightly shifting the reconstructed time to a higher value.

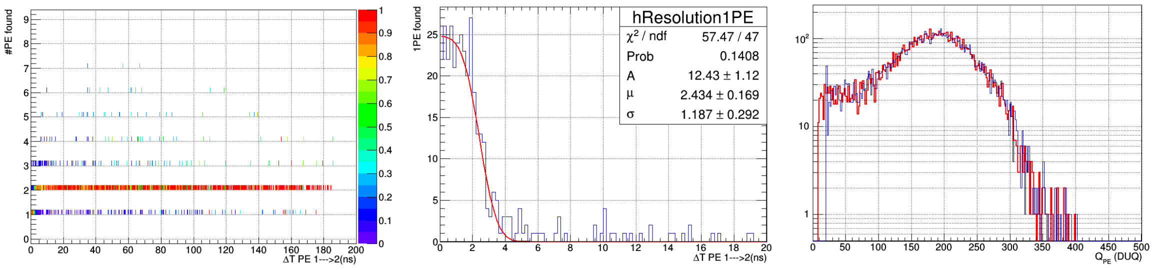
Time accuracy depends on the time difference between PEs. The time resolution of the reconstruction, i.e. how close two PEs can be and still be separated, can be studied by plotting the number of reconstructed PE from 2 PE-waveforms for various time separations between the two PEs $\Delta TPE_{1\rightarrow 2}$, as shown in Figure 3.27(a). The number of reconstructed PEs should be 1



(a) (b)

Figure 3.26: 3.26(a): Difference between the T_{start} reconstructed and the true hit time for RecoPules (black) and RecoZoR (blue). The T_{start} from RecoZoR shows a lot less dispersion than that of RecoPulse.

3.26(b): The central region is fitted by three Gaussians, for the well reconstructed PE and each of the two major mis-reconstruction sources.



(a) (b) (c)

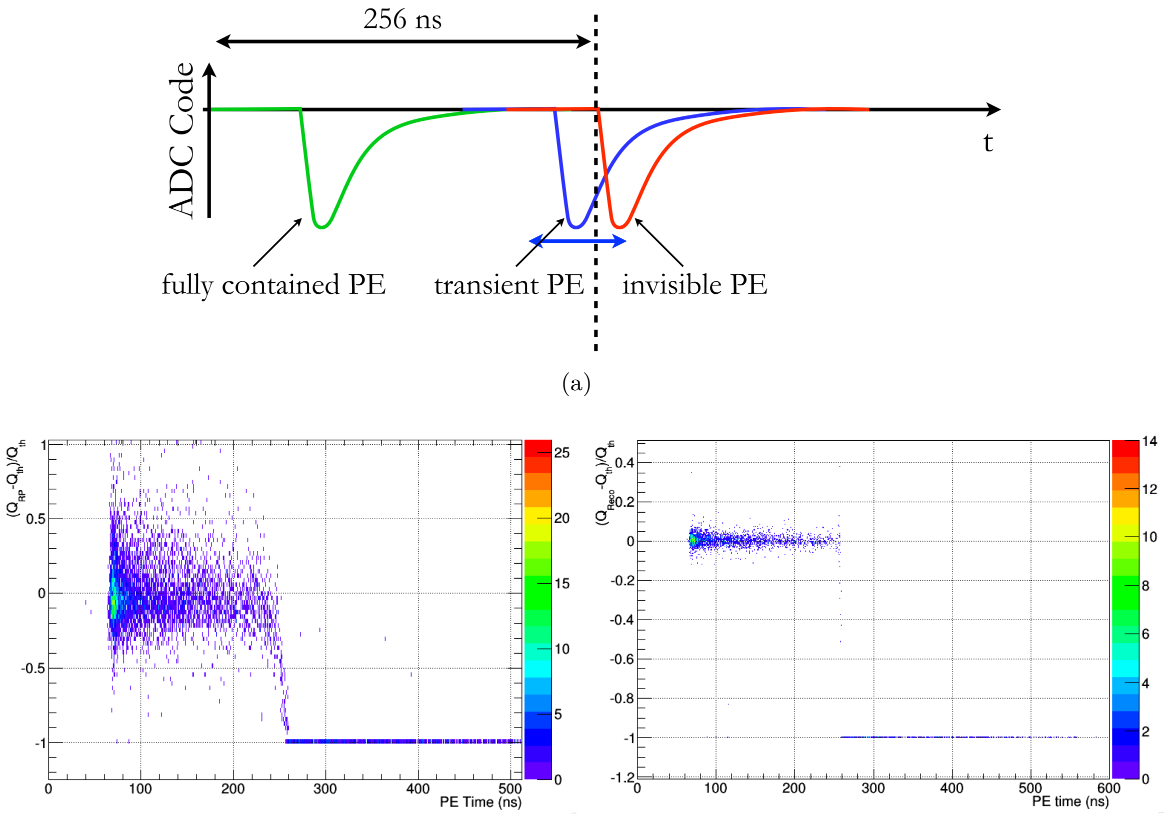
Figure 3.27: 3.27(a): Number of reconstructed PE for a 2 PE-waveform as a function of the time difference between the two PEs.

3.27(b): Number of 2 PE-waveform reconstructed as one PE versus the time difference between the two PEs. The reconstruction time resolution is 2.43 ± 0.17ns and was estimated fitting the histogram with an Error Function.

3.27(c): Charge per PE distribution for simulated waveforms (red line) and RecoZoR reconstructed (blue line).

when the time difference is below the resolution. Figure 3.27(b) shows the number of waveforms reconstructed as 1 PE as a function of $\Delta TPE_{1\rightarrow 2}$. The resolution of 2.43 ± 0.17ns is defined as the mean of the Error Function used to fit the histogram.

Figure 3.27(c) shows the charge per PE distribution for RecoZoR (blue line) and the true distribution (red line) for two PE waveforms with $\Delta TPE_{1\rightarrow 2} > 5$ ns. In this case, the two distributions agree perfectly, confirming the previous hypothesis that the discrepancies between the two distributions came from mis-behaviour due to PE pile up.



(a)

(b)

(c)

Figure 3.28: 3.28(a): Illustration of a possible non-linearity induced by the possibility of a PE to arrive at the end of the waveform and not be fully contained (transient PE).
3.28(b): Non-linearity of RecoPulse as a function of the hit time of the PE. As the PE arrives closer to the end of the waveform, the integrated charge represents a lower fraction of the total true charge, down to not seeing the PE at all.
3.28(c): Non-linearity of the charge reconstructed with RecoZoR. As long as the PE is fitted, the full information of the charge is known. Thus, no non-linearity is visible up to 256 ns when the PE becomes invisible.

Time Window Integration As seen in Figure 3.21(a), PEs have a non-negligible probability of arriving at the end of the waveform. Integrating the waveform, as is the case in RecoPulse, thus leads to possible inefficiency. Figure 3.28(a) presents the cases of a fully contained PE (in green) that does not suffer from no additional non-linearity, an invisible PE (in red) arriving outside the the waveform and therefore that is not reconstructed at all, and a transient PE, only partially contained and that leads to non-linearity.

Figure 3.28(b) shows the evolution of the non-linearity versus the hit time of the PE for RecoPulse. As the PE arrives closer to the end of the waveform, the integrated charge represents a lower fraction of the true charge. Thus, the amplitude of the non-linearity increases at high PE time.

Figure 3.28(c) shows the evolution of the non-linearity for RecoZoR. The non-linearity remains stable whatever the position of the PE since as long as the PE can be fitted, its complete charge is known.

3.8 Conclusion

We have discussed in length the sources of non linearity in RecoPulse's calorimetry. These sources are well understood by Double Chooz's collaboration and are taken into account in the energy reconstruction.
We presented RecoZoR, a new method of calibration, currently under development, that already seems very promising as a complement to RecoPulse.

RecoZoR has shown a more linear response in charge reconstruction, either at low charge and high PMT hit time, compared to RecoPulse. Moreover, RecoZoR charge linearity is much less dependent from the baseline true position than RecoPulse.

The reconstruction of the number of PEs from RecoZoR still requires work, however, the charge and time reconstructions already give RecoZoR all features of RecoPulse, with the same output capabilities, but without any of its non-linearities.

Proof of feasibility for RecoZoR was first introduced by Guillaume Pronost as part of his thesis [73]; I took over the development of his work and went back to the basics to make sure each effect observed was understood ans quantifiable for performance evaluation.
Due to time constraints, I could not implement RecoZoR in the framework of Double Chooz, however, numerous improvements are expected from this new reconstruction:

Energy Definition Stability Every time the electronics is power-cycled, the baseline position is changed, leading to a modification of the charge per PE definition which translates in the steps observed in Figure 2.13. A better baseline reconstruction, and a charge-independent PE definition will allow to reduce, if not cancel, the impact of this behaviour of the Flash-ADCs.

Energy Resolution Cancelling RecoPulse's energy non-linearity at low number of PE is expected to improve the energy resolution, especially at low energies. This would translate in a reduction of the n-H capture peak width.

Start Time Resolution RecoZoR provides a start time that is independent of a threshold, thus the time of the first PE is known with a great precision, and does not lose the first PE however small with respect to the heavier peak in the Waveform.

Spatial Resolution RecoBAMA relies on the hit time and charge seen by each PMT to reconstruct the interaction time and position within the detector. As RecoZoR improves the time and charge resolution, it is reasonable to expect an improvement in the vertex reconstruction as well.

Chapter 4

θ_{13} Measurement

The Double Chooz collaboration released $\sin^2(2\theta_{13})$ measurements based on different candidate selections and analysis methods. These analyses are described briefly in Chapter 2. In this chapter, the selection of IBD candidates and the $\sin^2(2\theta_{13})$ analyses are described. I will focus on DC-III and DC-IV measurements, describing analyses based on the candidates with neutron captures on Gadolinium (n-Gd), Hydrogen (n-H), and both captures together (Gd++) as shown in Figure 4.1. The n-Gd analysis has been presented at the Moriond 2016 conference [58].

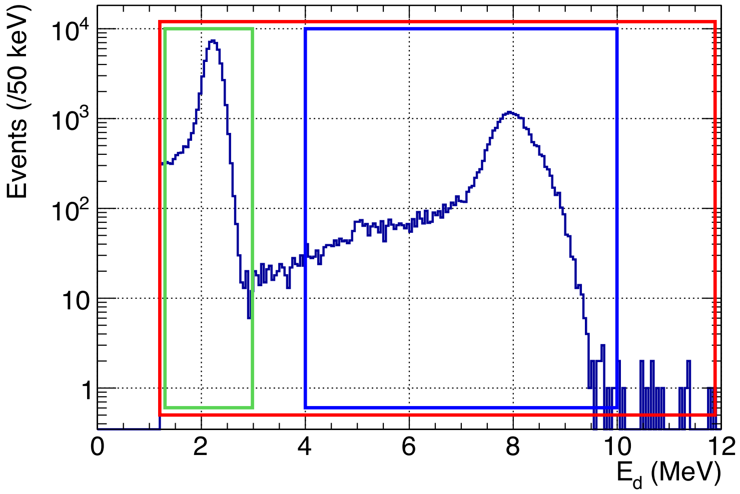


Figure 4.1: Delayed energy of events in the Near Detector. The green rectangle limits the n-H energy range, the blue rectangle limits the n-Gd energy range. The Gd++ analysis includes all neutron captures (red rectangle).

Background reduction is a fundamental part of $\bar{\nu}_e$ analyses. Neutrino detectors are often located deep underground to limit the impact of atmospheric muon contamination, as well as other muon-induced backgrounds.

Natural radioactivity is also a major source of background. During the building of the detectors, efforts were made to minimize the contamination of the volumes, to make Double Chooz as much of a low background experiment as possible, by selecting low radioactivity materials and imposing a ISO 6 clean room standard in the entire lab during the assembly (fewer than 10^6 particles of $0.1\,\mu\mathrm{m}$ per m^{-3}).

4.1 Candidate Selection

The selection of IBD candidates for the neutrino analyses is done through the coincidence in time and space of two valid triggers corresponding to a possible positron (prompt trigger) and subsequent neutron capture on Gd or H (delayed trigger). Most of the uncorrelated background can be rejected at the single trigger level by the definition of a valid trigger.

To be considered valid, a trigger must not be a muon, or a muon-correlated, a Light Noise (PMT-induced light), or a fixed rate trigger (automatic 1 Hz trigger to compute the detector live-time and the waveform baselines), and its Inner Detector energy must be within $[0.4, 100]$ MeV.

Muon veto: A trigger is considered as a muon trigger when one of these three conditions is verified:

- The total charge in the Inner Veto, summed over all the IV-PMT is $Q_{IV} > 50 \cdot 10^3$DUQ in the Far Detector, and $Q_{IV} > 30 \cdot 10^3$ DUQ in the Near Detector.
- The number of IV-PMT groups hit is higher than 10 over the 18 groups, meaning that more than 50% of the IV-PMT are hit.
- The energy in the Inner Detector is $E_{ID} > 100\,\mathrm{MeV}$.

When muons cross the detector, they produce cosmogenic isotopes and fast neutron showers. A muon veto time of $500\,\mu\mathrm{s}$ is applied after a muon trigger.

Light Noise veto: Finally, a condition has been included to reject Light Noise. This background is produced by ID-PMTs sporadically emitting flashes of light from their epoxy bases [82]. The rate of Light Noise increased since the Far Detector first started data taking in 2011 from $\sim 30\,\mathrm{Hz}$ to nearly $400\,\mathrm{Hz}$ in 2016. Light is emitted at the base of a single PMT, and therefore other PMTs see much less light than the one hit, and much later. *A contrario*, an IBD-like trigger happens in the Target or the γ-catcher, in a more central volume of the detector, with an isotropic light propagation. Thus, all PMTs are hit at approximately the same time, with a short time of flight delay. The RMS of the charge distribution and the hit-time of each PMT in a trigger provides a means to discriminate between Light Noise and IBD events, as well as the ratio $Q_{\mathrm{max}}/Q_{\mathrm{tot}}$ of the maximum charge collected by a PMT over the total charge collected in the Inner Detector.

Another variable has been constructed in order to reduce Light Noise. The PMT with highest charge collection Q_{max} is found, and the PMTs in a 1 m sphere radius centred on that PMT are selected to compute the variable Q_{diff}:

$$Q_{\text{diff}} = \frac{1}{N} \sum_{i}^{N} \frac{(Q_{max} - q_i)^2}{q_i} \tag{4.1}$$

The signal difference between the PMT with maximum charge and its neighbours should lead to a low Q_{diff} in the case of a physical event in the Inner Detector, whereas it is expected to be large in the case of a Light Noise trigger.

Light Noise rejection is achieved with the following conditions :

- $\text{RMS}_{T_{\text{start}}} < 36$ ns or $\text{RMS}_Q < 1680 - 28 \times \text{RMS}_{T_{\text{start}}}$
- $Q_{max}/Q_{tot} < 0.20$
- $Q_{\text{diff}} < 100\,000$

Figure 4.2(a) and 4.2(b) show the RMS_Q vs RMS_t distribution in the Far Detector and the Near Detector respectively. In both cases, the cut of the Far Detector are materialized by the black line, any event on the right side of this line is rejected. In order to avoid as much as possible any Light Noise contamination in the Near Detector, the bases of the PMTs have been covered with black plastic. The absence of a strong contamination of Light Noise in the Near Detector is noticeable by the disappearance of the right-hand lobe of the distribution.

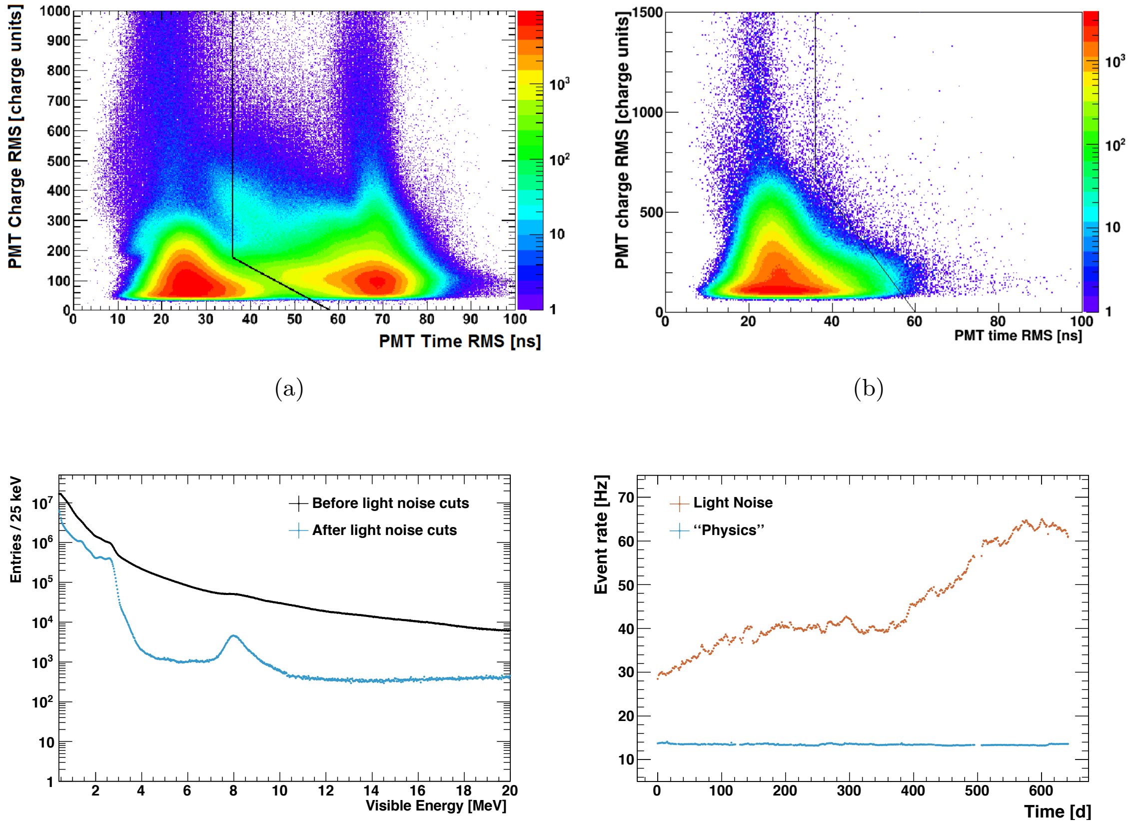
Figures 4.2(c) and 4.2(d) respectively show the energy spectrum and rate of events before and after the application of the Light Noise cuts for the Far Detector. Physical features appear in the energy spectrum after application of the cuts, especially the ^{208}Tl peak at 2.6 MeV and the n-Gd capture peak at 8 MeV. The event rate is very stable after application of the cuts, whereas the Light Noise rate is consistently increasing with time.

Any trigger that does not verify one of the three conditions mentioned above is rejected. At this point valid triggers form what is called the "Single triggers selection": they are not yet paired in prompt/delayed candidates.

Among the valid triggers, additional selection criteria are applied that depend on the studied sample (neutron capture on H, Gd, or the so-called Gd++ sample where all neutron captures are used). Those criteria are an additional Muon veto time (ΔT_μ), an energy range for prompt and delayed candidates, a time and space correlation between the prompt and the delay (ΔR, ΔT) and two isolation time windows ΔT_{before} and ΔT_{after} as summarized in Table 4.1.

There can be no valid trigger in the prompt isolation window $[-\Delta T_{before}, 0\,\mu\text{s}]$ before a prompt candidate, and only one delayed candidate in the delayed isolation window $[0.5\,\mu\text{s}, \Delta T_{after}]$ after the prompt candidate as illustrated in Figure 4.3.

The isolation criterion ΔT_{before} is meant to ensure that a prompt trigger does not directly follow another potential candidate, and thus cannot be a delayed trigger misidentified as a prompt. ΔT_{after} on the other hand, helps to reduce high multiplicity events, that cannot be IBD since IBD events only have two triggers.



(a) (b)

(c) (d)

Figure 4.2: 4.2(a) and 4.2(b): RMS_Q vs RMS_t distribution for the Far Detector (4.2(a)) and the Near Detector (4.2(b)). The right-hand lobe of the distribution corresponds to the Light Noise, that is no longer present in the Near Detector. The lines in both plots correspond to the cut applied in the Far Detector.
4.2(c): Energy spectrum before and after application of Light Noise cuts for the Far Detector. The ^{208}Tl peak from radioactivity becomes visible at 2.6 MeV, as well as the n-Gd peak at 8 MeV.
4.2(d): Evolution of the event rate over time before and after the application of Light Noise cuts for the Far Detector. No more Light Noise remains, leaving a constant rate of physical events.

Table 4.1: Additional vetoes applied to Single triggers to pair them in the definition of a prompt/delayed candidates for the DC-IV analysis.

Selection	ΔT_μ(µs)	E_p(MeV)	E_d(MeV)	ΔT_{before}(µs)	ΔT_{after}(µs)	ΔT(µs)	ΔR(m)
n-Gd	1000	$[0.3, 100]$	$[4.0, 10.0]$	600	600	$[0.5, 150]$	< 1.0
Gd++	1250	$[0.3, 100]$	$[1.4, 100]$	800	900	$[0.5, 800]$	< 1.2

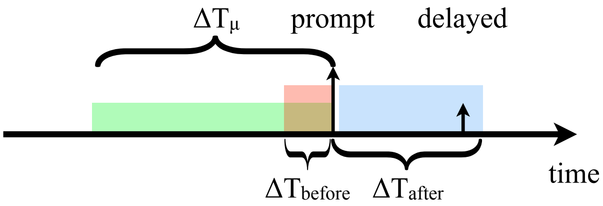


Figure 4.3: Time windows for candidate selection: the green box corresponds to the muon veto, the red box corresponds to the prompt isolation window, and the blue box corresponds to the delayed isolation window.

4.2 Backgrounds in the Candidate Selection

Backgrounds in Double Chooz can be sorted in two main categories: correlated and uncorrelated backgrounds.

Uncorrelated backgrounds: are produced by the random coincidence of two uncorrelated events, e.g. two radioactive decays. These background events will be referred to later on as "Accidentals".

Correlated Backgrounds: are produced by two events from a single physical process, e.g. a muon decaying to a Michel electron.

4.2.a Uncorrelated Background

Accidental coincidences: The coincidence condition between the prompt and delayed trigger allows to greatly reduce the Accidental background.

This background is studied over multiple off-time integration windows as represented in Figure 4.4. In the n-Gd analysis, for each prompt candidate, a series of 2000 windows of 1.2 ms each, starting at 1 s after the prompt trigger and separated by 800 µs are opened. In the Gd++ analysis, 500 windows of 1.7 ms are used.

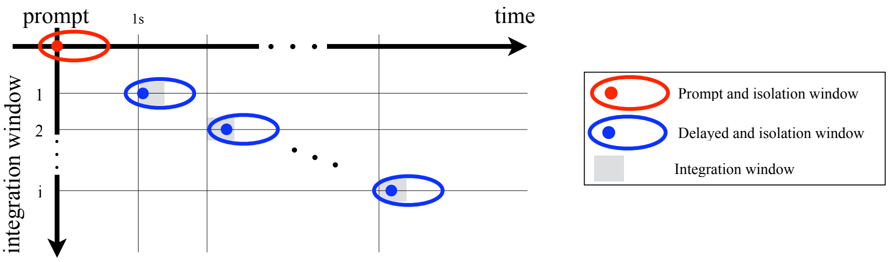
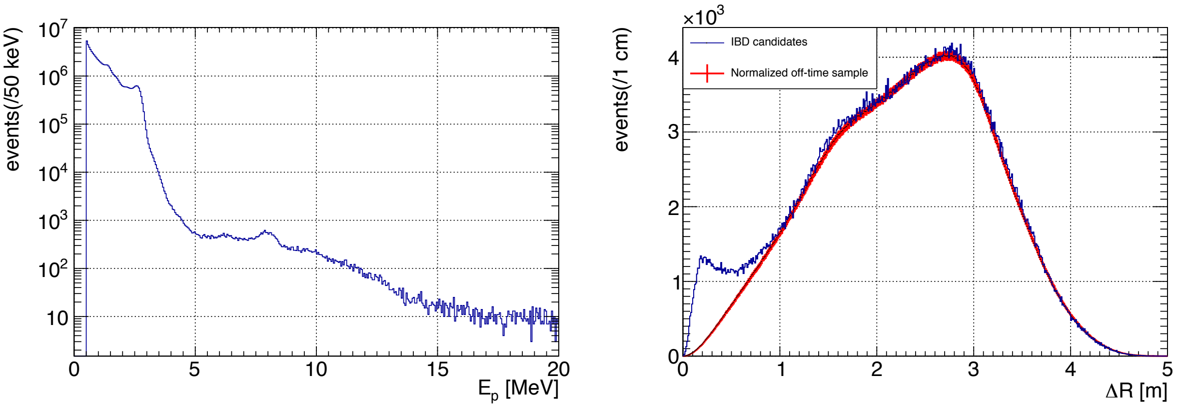


Figure 4.4: Off-time integration windows. For each prompt trigger candidate, multiple integration windows, are opened after a 1 s delay.

Figure 4.5 shows the energy spectrum, the distance correlation ΔR and the time correlation ΔT of candidates (blue) and off-time coincidences (red).

Figure 4.5(a) shows the prompt energy of Accidental events. These events are dominated by natural radioactive decays paired with a later neutron capture. The ^{40}K γ line at 1.460 MeV and



(a)

(b)

Figure 4.5: 4.5(a): Prompt energy of Accidental events. These events are dominated by natural radioactive decays paired with a later neutron capture. The ^{40}K γ line at 1.460 MeV and the ^{208}Tl γ line at 2.614 MeV are visible, as well as the peak of the neutron capture on Gd at 8 MeV. 4.5(b): Correlation distance ΔR between a prompt and the corresponding delayed trigger. The blue line corresponds to the IBD candidates in the n-H data set, the red line is the normalized ΔR distribution of the Accidental sample normalized to the IDB distribution.

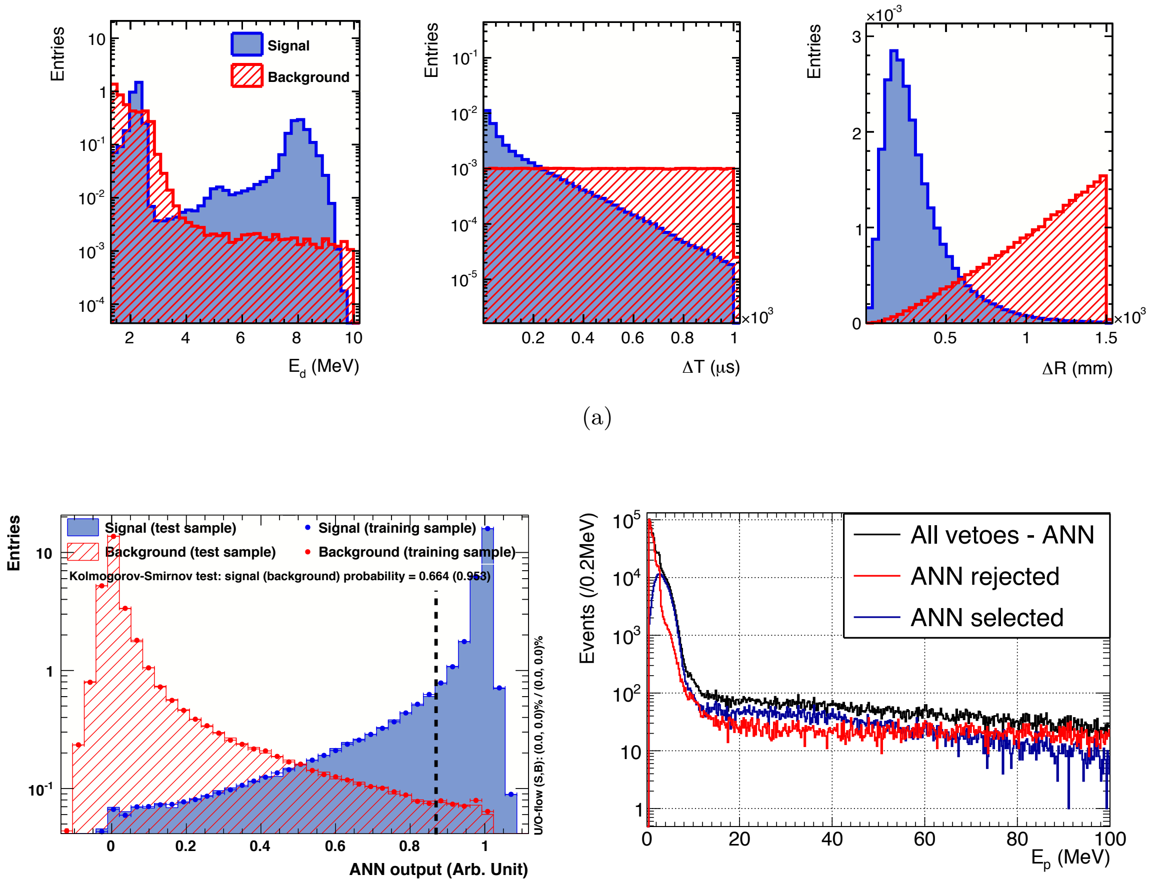
the ^{208}Tl γ line at 2.614 MeV are visible, they are the dominant γ-emitters in the surrounding rock, as well as the peak of the neutron capture on Gd at 8 MeV. The Accidental contamination decreases by more than three orders of magnitude above 4 MeV, and becomes virtually negligible around the n-Gd peak compared to the Accidental contamination of the n-H peak.
Figure 4.5(b) shows the ΔR distribution of events. IBD are correlated in distance as the neutron is emitted with low energy and make up the peak visible below 1 m. However, Accidentals are not and make up the rest of the distribution; their shape is the result of a random sampling of the spatial distribution of the Singles.

In order to take advantage of the time and space correlation, as well as the well-defined delayed energy range of the IBD compared to the Accidental events, an Artificial Neural Network (ANN) has been developed.

The ANN was trained using MC simulated IBD interactions and Accidental coincidences following the distributions studied in the multiple time windows method and presented in Figure 4.5. The training distributions are presented in Figure 4.6(a), and the output distribution for signal and background is shown in Figure 4.6(b). The signal-to-background ratio is optimized by varying the minimum required ANN output for an event to be accepted.

The ANN was originally developed for Accidental rejection in the Hydrogen analysis [50], and has been used in the Gd++ analysis as it includes n-H captures as well as n-Gd.

Events with an ANN output lower than 0.86 in the Near Detector (0.85 in the Far Detector)



(a)

(b)

(c)

Figure 4.6: 4.6(a): Distributions of delayed energy, ΔT and ΔR for signal and backgrounds used to train the ANN.
4.6(b): Distributions of background (red stripped area) and signal (blue area) events as a function of the ANN output, in arbitrary units. This representation allows to optimize the value of the ANN selection to obtain the best signal-to-background ratio while keeping a low signal inefficiency.
4.6(c): Prompt energy distribution of ND Gd++ events, when all vetoes but ANN are applied (black), rejected by the ANN (red), and selected by the ANN (blue).

are rejected (black dotted line in Figure 4.6(b)). The ANN rejects 97.11% of off-time events and 14.54% of the MC training sample of IBD signal events in the Near Detector (respectively 97.21% and 13.60% in the Far Detector). Finally, 36.63% of on-time events are rejected in the Near Detector while 62.04% of the same events are rejected in the Far Detector. This difference can be explained by the fact that Accidentals are no longer the dominant background in the Near Detector, thus they make up for a smaller fraction of the on-time events.

Figure 4.6(c) shows the prompt energy distribution of Gd++ events in the Near Detector. The black line corresponds to events selected when all vetoes are applied except the ANN, the red curve corresponds to events rejected by the ANN, and the blue line corresponds to events accepted by the ANN. The energy spectrum of the Accidentals is clearly visible in the rejected sample, and no longer in the accepted sample.

Cosmogenic Background: Long-lived radio-isotopes ^{9}Li and ^{8}He are created through atmospheric muon-induced spallation on ^{12}C contained in the organic scintillators.

These isotopes are $(\beta\, n)$ emitters, and as such are indistinguishable from the IBD candidates on an event-by-event basis.

The life time of ^{9}Li and ^{8}He are respectively 257 ms and 172 ms respectively, much longer than the 1 ms muon veto time. Given the muon rate in Double Chooz, increasing the muon veto time to reduce these isotopes would lead to an unacceptably large loss in exposure time.

A Li-likelihood, $\mathcal{L}_{^{9}\mathrm{Li}}$, has been developed based on the distance to a muon track and the multiplicity of neutrons following this muon. The muon considered in computing the likelihood of a given event is the one that yields the highest value in a time window of 700 ms preceding the event. ^{9}Li reduction is performed based on the value of this likelihood.

The Cosmogenic background spectral shape presented in Figure 4.7 is obtained by selecting a Li-enriched sample with the likelihood selection $\mathcal{L}_{^{9}\mathrm{Li}} > 0.4$. By fitting the MC spectra of ^{8}He and ^{9}Li to this spectrum, the fractional contribution of ^{8}He was found to be consistent with zero. The Cosmogenic background can therefore be referred to as only ^{9}Li later on.

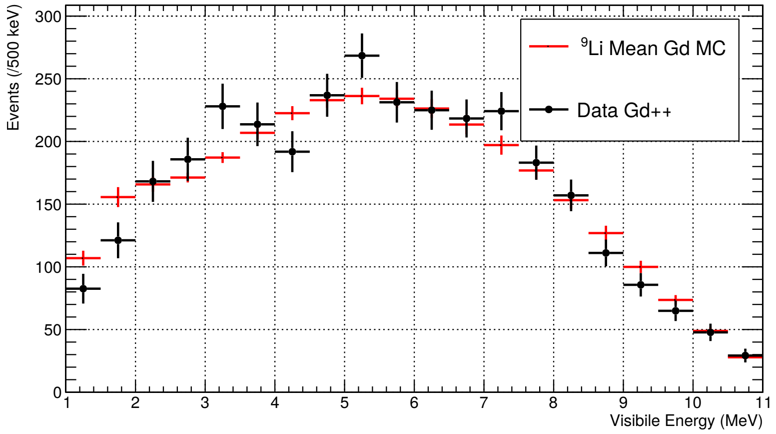


Figure 4.7: ^{9}Li spectrum for the DC-IV Gd++ analysis. The black point correspond to the data with $\mathcal{L}_{^{9}\mathrm{Li}} > 0.4$, the red point correspond to the MC spectrum of the ^{9}Li.

Residual rate of ^{9}Li is computed by fitting the spectral shape in Figure 4.7 to the IBD candidate spectrum along with other background spectral shapes and the oscillated shape of the $\bar{\nu}_e$ as will be presented in Section 4.4.

4.2.b Correlated Background

Correlated Backgrounds in Double Chooz consist of Stopping Muons and Fast Neutrons. Stopping Muons are atmospheric muons stopping and decaying after travelling a short distance in the detector, thus depositing an energy low enough not to be tagged as muons, and thus mis-reconstructed as IBDs. Fast Neutrons shower from a muon crossing the surrounding rocks.

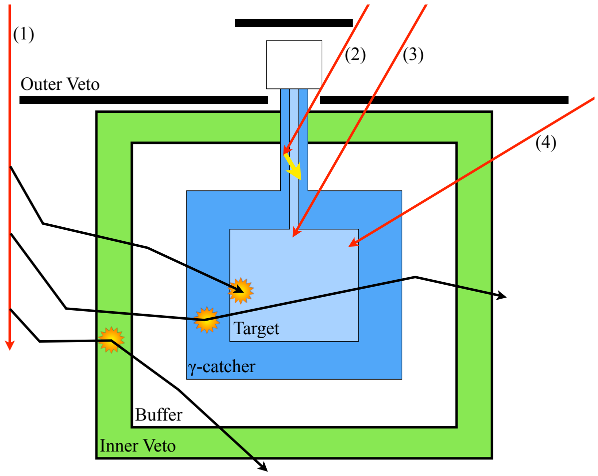


Figure 4.8: Sketch of the Correlated Background processes.
Muon (1) crosses the surrounding rocks and produces Fast Neutron showers through spallation. Some of them will enter the detector and undergo proton recoils or capture, producing the so-called Fast Neutron background.
Muon (2) enters the chimney without being detected by Outer- nor Inner-Veto. If the distance travelled before decaying is short enough, the deposited energy remains below the muon energy threshold. The scintillations from the muon and the subsequent Michel electron mimic a prompt and delayed signal, producing the so-called Stopping Muon background.
Muons (3) and (4) are detected by either the Outer- or Inner-Veto and as such are rejected.

Figure 4.8 shows a sketch summing-up the the processes leading to the creation of Correlated Backgrounds. The red arrows represent atmospheric muons, the black arrows the spallation Fast Neutrons, and the yellow arrow the Michel electron.

Fast Neutrons (FN): Showers of Fast Neutrons are generated by muon spallation in the surrounding rocks. Due to their low interaction cross-section, fast neutrons can cross the Inner Veto unseen and interact in the Inner Detector.

Fast Neutron interactions in the Inner Detector can be prompt-like when a neutron produces a nuclear recoil, or delayed-like when a neutron captures on Gd or H.
Fast Neutrons entering the detector within the same shower are correlated in time, so a nuclear recoil and a later neutron capture are perfectly consistent with an IBD candidate space and time correlation. As such, they cannot be separated from IBD candidates on an event-by-event basis.
Fast Neutron induced proton recoils in the Inner Detector have a wide energy range spanning over the full energy range of Double Chooz. Fast Neutron with visible energy of more than 100 MeV are treated as Muons.

Correlated Background originate outside of the detector, thus the Inner-Veto is a very powerful tool to reject and/or study this background. A sub-sample of Correlated Background can be studied by observing events with a time and space correlation between energy depositions in the Inner Veto and the Inner Detector. However, muons crossing the Inner Veto to enter the

Inner Detector deposit a high energy in the Inner Veto and have already been rejected by the muon veto earlier on. To first order, only Fast Neutrons are left in the sample.

The energy deposition in the Inner Veto by Fast Neutrons is possible when one of the spallation neutrons of the shower scatters in the Inner Veto. The neutrons that deposit energy in the Inner Veto are not necessarily the same as those that interact in the Inner Detector. Thus the energies in both volumes are uncorrelated, and the spectrum of the tagged Fast Neutron sample is expected to be the same as the one of the non-tagged sample.

The selection is based on a correlation in time and space between the two volumes, the number of hit PMT in the Inner Veto, and the charge (or energy) seen in each volume:

- #IV-PMT > 1
- $-40\,\mathrm{ns} < \Delta T_{\text{ID-IV}} < 70\,\mathrm{ns}$
- $\Delta R_{\text{ID-IV}} < 3.7\,\mathrm{m}$

Events that satisfy these conditions are rejected as Correlated Background. This selection is called Inner Veto Veto (IVV). As shown in Figure 4.9(a), a fraction of those IVV-rejected events have anti-correlated energy depositions in the Inner Veto and the Inner Detector. Those events could correspond to a same highly penetrating particle deposits energy in both volumes (e.g. muons or high energy γ).

In order to specifically study Fast Neutrons, those events with anti-correlated energy are rejected, by requiring another condition:

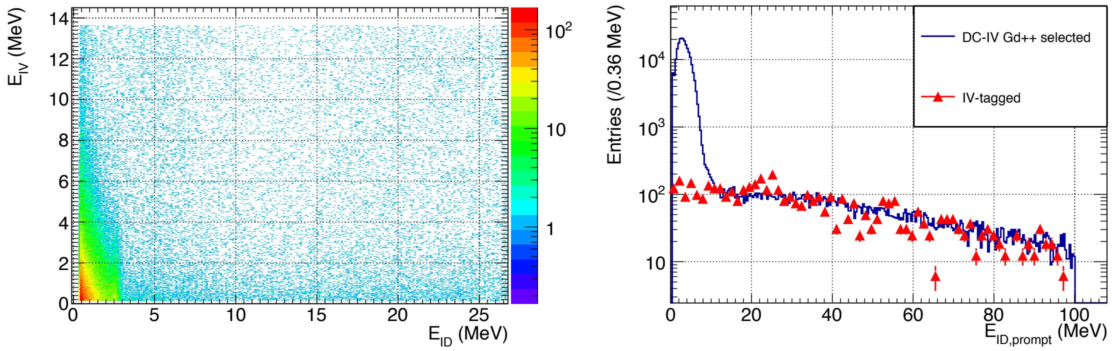
$$Q_{IV} > 25\,000\,\mathrm{DUQ} \mathrel{||} E_{ID} > 10\,\mathrm{MeV}$$

Events that would be rejected by the IVV, but satisfy this condition are called IV-tagged events, and correspond to our tagged Fast Neutron sample.

Figure 4.9(b) shows the prompt energy spectrum of IV-tagged events compared to the distribution of IBD candidates. The Fast Neutron sample is scaled to the IBD candidate sample on the [20,100] MeV sample, allowing an estimation of the remaining Fast Neutron contamination in the IBD sample.

Here the importance of opening the prompt energy range up to 100 MeV when IBD only range up to 10 MeV becomes obvious. The added contamination of Fast Neutrons is outside the neutrino physics energy range, so it does not impact our measurement of θ_{13}, but the handle on the contamination below 10 MeV is greatly improved by being able to estimate the number of Fast Neutrons on a much wider energy range.

In the DC-III n-Gd analysis, the relative uncertainty on the Fast Neutron background was 17%, the contamination was only estimated up to 20 MeV. In the DC-IV Gd analysis, the relative uncertainty is now of the order of 5%, and 3.5% for the DC-IV Gd++ analysis. The increased precision comes from more statistics, and from the increased energy range that allows a better estimation of the contamination.



(a) (b)

Figure 4.9: 4.9(a): Inner Veto deposited prompt energy versus Inner Detector deposited prompt energy for IVV rejected events, zoomed-in on the low energies. The triangular shape at low energies correspond to energy deposition in both volumes by a single particle such as a muon or a high energy γ.
4.9(b): Prompt energy spectrum with IBD selection in the Near Detector. The blue histogram is the selected IBD sample, the red dots correspond to the IV-tagged samples, scaled to the blue histogram on the [20,100] MeV range.

Understanding the Fast Neutron spectrum is not trivial. There are only few measurements, and predictions from simulations exhibit large variability. A dedicated neutron TPC (DCTPC) was developed and operated in both Near and Far Laboratories in order to study the ambient Fast Neutron in the laboratories. I took part in the installation, running phase and analysis of data from this detector. I was particularly in charge of the energy reconstruction and study of the α backgrounds in order to provide the pure neutron-induced nuclear recoil spectrum. This work is presented in Chapter 5.

Stopping Muons: Muons fully crossing the detector or travelling long distances in the Inner Detector are easily identifiable on an event-by-event basis as they deposit a high energy in the Inner Detector, and most likely in the Inner Veto as well. Any muon travelling more than $\sim 10\,\text{cm}$ in the Inner Detector will deposit an energy higher than the muon energy threshold of $20\,\text{MeV}$ and be rejected even if no energy is deposited in the Inner Veto. Muons can enter the Inner Detector through the Chimney without triggering either veto. The Chimney is filled with scintillator, the Michel electron from the muon decay will produce a delayed trigger to the prompt muon scintillation time related to the muon life-time of $2.2\,\mu\text{s}$ and with a high spatial correlation.

A major part of my work was dedicated to developing methods to identify Stopping Muons with low inefficiency on the IBD sample, while aiming for the highest possible Stopping Muon rejection power. This study will be presented in more details in Chapter 6.

Because of a high rejection power on Stopping Muons, the remaining Correlated Background is mostly composed of Fast Neutrons. The terms Correlated Background and Fast Neutrons will be used equally later on.

Remaining background rates after tagging and reductions are specific to an analysis, and will be summarized for each case in Table 4.2.

4.2.c Reactor-off Measurement

Due to the configuration of the CHOOZ power plant, with only two reactors, Double Chooz is the only θ_{13} reactor experiment able to have periods of data taking with both reactors off. This allows for a nearly neutrino-free sample, and an important cross-check of our background rejection capabilities.

Although both reactors are off, the $\bar{\nu}_e$ flux is not expected to be null as β-decays of long lived-isotopes are still occurring in the cores. However, because of the IBD threshold at 1.8 MeV, Double Chooz is mostly sensitive to high energy neutrinos, which are produced by short-lived decays. Turning off the reactors leads to a dramatic decrease of the effective neutrino flux seen by Double Chooz. Double Chooz is sensitive to two long-lived fission chains:

$$^{106}\mathrm{Ru} \xrightarrow[371.8\,\mathrm{d}]{\beta^-} {}^{106}\mathrm{Rh} \xrightarrow[30.1\,\mathrm{s}]{\beta^-(3.5\,\mathrm{MeV})} {}^{106}\mathrm{Pd} \tag{4.2}$$

$$^{144}\mathrm{Ce} \xrightarrow[284.9\,\mathrm{d}]{\beta^-} {}^{144}\mathrm{Pr} \xrightarrow[17.3\,\mathrm{min}]{\beta^-(3.0\,\mathrm{MeV})} {}^{144}\mathrm{Nd} \tag{4.3}$$

The flux prediction relies on the fission rate, based on the reactor thermal power, and thus cannot be used during reactor-off periods. Instead, a prediction from FISPACT is used, solving the Bateman equations from an input spectrum computed with BESTIOLE [83].

I have been in charge of producing the analysis of the reactor-off data for the H-III analysis [50]. Figure 4.10 shows the selection results for the ΔT, E_d and E_p distributions. The black shaded area, red and blue dots correspond respectively to all the IBD candidates before application of the vetoes, the rejected candidates, and the accepted candidates after application of the full selection.

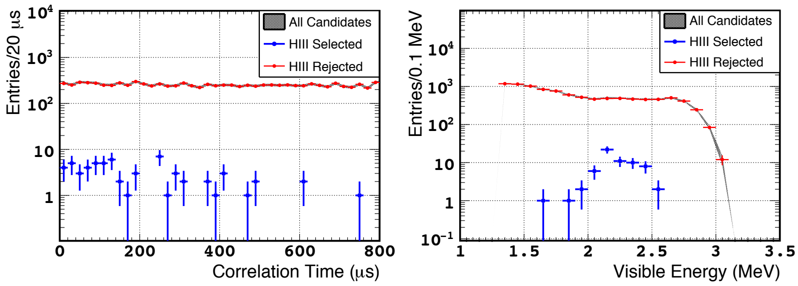


Figure 4.10: From left to right: ΔT, E_d, and E_p distributions for reactor-off data of the H-III analysis. The black shaded area corresponds to all the IBD candidates before application of the vetoes, the red points to the rejected candidates after application of all vetoes, and the blue points to the accepted candidates.

Table 4.2 summarizes the measured and expected rates after applying the full selection for

DC-III and DC-IV analyses. The backgrounds are estimated from the background models described earlier on in this chapter. Background rejection powers of 91.1% for DC-III n-Gd and DC-IV n-Gd analyses and 99.4% for DC-III n-H and DC-IV Gd++ analyses were achieved.

Although the reactor-off time is too short to increase significantly the sensitivity to θ_{13}, it allows to check the sanity of the estimations of the residual background rates.

Table 4.2: Total measured and predicted rates during reactor-off periods for DC-III and DC-IV analyses.

	DC-III n-Gd	DC-III n-H	DC-IV n-Gd	DC-IV Gd++
Live time	7.24 days	7.15 days	7.24 days	7.16 days
All Candidates (day^{-1})	8.56	1415.7	8.70	1413.00
IBD Candidates (day^{-1})	0.97	8.8	0.97	9.50
expected $\bar{\nu}_e$ (day^{-1})	0.21 ± 0.065	0.38 ± 0.11	0.21 ± 0.065	0.58 ± 0.17
expected Accidentals (day^{-1})	0.070 ± 0.003	4.32 ± 0.06	0.070 ± 0.002	3.55 ± 0.01
expected Cosmogenic (day^{-1})	$0.97^{+0.41}_{-0.16}$	$0.95^{+0.57}_{-0.33}$	0.86 ± 0.41	2.57 ± 0.60
expected FN/SM (day^{-1})	0.53 ± 0.09	1.45 ± 0.20	0.58 ± 0.030	2.89 ± 0.10
Total Prediction (day^{-1})	$1.78^{+0.43}_{-0.19}$	$7.10^{+0.62}_{-0.41}$	1.73 ± 0.43	9.6 ± 0.63

4.3 Detection Efficiency

Double Chooz's trigger efficiency reaches 100% above 0.5 MeV, meaning that any event above 0.5 MeV in the Inner Detector is guaranteed to be stored. However, not all IBD interactions end up in the candidate sample, effects such as Gd concentration, neutron scattering, energy scale and analysis selection can allow an IBD to pass undetected.

As will be discussed later in Section 4.4.a, the θ_{13} analyses is performed by comparing the data to an oscillated prediction of the anti-neutrino flux in each detector. The fit requires to know the detection efficiency for data and MC, the ratio of these two values is a correction factor applied to scale the MC to the data.

Gd fraction: Neutrons capture dominantly on H and Gd, the fraction of neutrons capturing on Gd depends directly on the relative concentrations of H and Gd. This relative concentration is measured using a ^{252}Cf neutron source. Neutrons from the source capture on both H and Gd and are coincident with an emission of γ's. The Gd fraction is then computed as:

$$\mathrm{Gd}_f = \frac{N_{3.5\,\mathrm{MeV}<E<10\,\mathrm{MeV}}}{N_{0.5\,\mathrm{MeV}<E<10\,\mathrm{MeV}}} \tag{4.4}$$

In addition to the usual IBD selection, two cuts are applied to reduce the contamination from Accidental coincidences:

- $E_p > 4\,\mathrm{MeV}$.
- More than one neutron is detected after a prompt signal (^{252}Cf emits 3.8 neutrons per fission in average).

The correction factor data to MC is $C_{\mathrm{Gd}} = 0.9767 \pm 0.0022$ for the FD and 0.9748 ± 0.22 for ND.

Gd fraction has no impact on the Gd++ as the full energy window is used for the delayed neutron capture. The detection is therefore 100% efficient.

Veto efficiency: Here, IBD selection efficiency refers to the efficiency of the vetoes applied to define the candidate sample (Light Noise, Muon Veto, and ANN when applicable). This efficiency is measured by applying the vetos to the ^{252}Cf source data.
Another measurement is performed in a data-driven manner, by comparing the impact of each analysis selection on the selected data without those vetos:

$$\epsilon_{Vetos} = \frac{N(\text{Vetos} \oplus \text{IBD Selection})}{N(\text{IBD Selection})} \tag{4.5}$$

The correction factors for the veto efficiency in the DC-IV, n-Gd analysis are 1.0010 ± 0.0018 for the FD and 0.9840 ± 0.0009 for the ND. In the DC-IV, Gd++ analysis, they are 1.0102 ± 0.0028 for FD and $0.99.61 \pm 0.0011$ for the ND.

IBD selection efficiency: In complement to the veto efficiency, the IBD selection efficiency is defined as the efficiency of all the cuts applied to the candidates after the vetos have been applied.

$$\epsilon_{IBD} = \frac{N(\text{Vetos} \oplus \text{IBD Selection})}{N(\text{Vetos})} \tag{4.6}$$

This efficiency includes efficiencies from all vetoes applied on the IBD candidates (IVV, OVV, Lithium Veto, and SM Veto).
The Correction factor for the IBD selection efficiency is found to be, for the DC-IV, n-Gd analysis : 0.9222 ± 0.0003 for FD-I, 0.9164 ± 0.0005 for FD-II, and 0.7586 ± 0.0002 for the ND. For the DC-IV, Gd++ analysis, these numbers are 0.9025 ± 0.0022 for FD-I, 0.8939 ± 0.0004 for FD-II and 0.6971 ± 0.0001 for ND.

Spill Effects: Spill-in and spill-out refer to neutrons capturing in a volume different from the one the neutrino interacted in. In the case of the n-Gd sample, neutrons capture in the Target, but, because of neutron scattering, the IBD could have taken place in the γ-Catcher, effectively increasing the detection volume. However, IBD occurring at the edge of the target may be undetected because the neutron may leave the target and be captured on H in the γ-Catcher.
For n-H, the Target is less efficient than the γ-Catcher, as $\sim 80\%$ of the captures occur on the Gd. This effect leads to an inefficiency on the γ-Catcher close to the target.
For the Gd++, there is no efficiency loss at the γ-Catcher/Target interface because all neutron

captures are considered.

All analyses however suffer from spill effects at the interface γ-Catcher/Buffer since interactions in the Buffer are invisible.

Spill effects are a geometrical effect measured with MC simulations, using Tripoli4 and Double Chooz's custom Geant4 neutron simulation package NeutronTH. The uncertainties are estimated as the discrepancies between the two methods.

The correction numbers for the spill effects are, for the DC-IV, n-Gd analysis: 1.0000 ± 0.0027 for both detectors, and 1.0000 ± 0.0020 for the DC-IV, nGd++ analysis.

Proton Number The number of targets for IBD interactions is the number of protons in the detection volume (Target for n-Gd analysis [61], and Target + γ-Catcher for n-H [50] and Gd++ analyses). The proton number is estimated by two different methods:

- The mass of Target scintillator is carefully measured upon filling.
- The volumes of the acrylic tanks are known, and the detectors are maintained at a constant temperature of $13.4 \pm 0.5\,^{\circ}\mathrm{C}$, thus the density and volume are known.

The chemical composition, the volumetric mass, and the molar mass are known for our scintillators. The Hydrogen fraction of the scintillator of the target is $f_H = 0.1360 \pm 0.0004$. The composition of the scintillators of both Targets are identical since they have been synthesized at the same time so they could age identically. It is then possible to measure the number of free protons inside the detector.

The proton number of the γ-Catchers is not known with the same accuracy as the one of the Target because the liquid scintillator was not weighed prior to filling. Indeed, at the time of the commissioning of the Far Detector, it was not envisioned to perform an analysis on Hydrogen, and thus a precise knowledge of the γ-Catcher was not mandatory.

Unfortunately, the proton number corrections have not been fully computed for the DC-IV analysis as these lines are written.

4.4 θ_{13} analysis in the single-detector phase

In this section, the methods used to perform the θ_{13} analysis in Double Chooz, for the single-detector phase, are presented. During this phase, only the Far Detector, in the FD-I configuration, was taking data. The results presented here correspond to the latest analysis performed in this phase, called DC-III.

The observed deficit between the IBD candidate and the sum of the predicted un-oscillated reactor neutrino signal and the estimated background contamination in each detector can be interpreted as the result of neutrino oscillations.

A first estimation of θ_{13}, in the 2 flavor scheme approximation, may be obtained by this expression, using only the observed rate of neutrinos:

$$\frac{N_{\mathrm{IBD}} - N_{\mathrm{BG}}}{N_{\nu,\ \mathrm{pred.}}} \simeq 1 - \sin^2(2\theta_{13}) \left(\frac{1.27\Delta m^2_{31}(\mathrm{eV}^2)\langle L\rangle(\mathrm{m})}{\langle E_\nu\rangle(\mathrm{MeV})} \right) \tag{4.7}$$

where N_{IBD} is the number of observed IBD candidates, N_{BG} the number of estimated remaining backgrounds, $N_{\nu,\ \mathrm{pred.}}$ the number of expected neutrino interactions without oscillation, $\langle L\rangle$ the average distance of each reactor, and $\langle E_\nu\rangle$ the average neutrino energy. In this expression, the ratio $\langle L\rangle/\langle E\rangle$ is weighted over the two reactors in order to account for their relative $\bar{\nu}_e$ flux contribution to the measurement.

In Double Chooz, two complementary analyses, referred to as *Rate+Shape* (R+S) and *Reactor Rate Modulation* (RRM), are performed. The R+S analysis is based on a fit of the observed energy spectrum with the spectral shapes of the expected neutrino spectrum and background contributions. The RRM analysis exploits the variability of the neutrino flux with the thermal power of the reactors while the background rate remains constant [55].

Both of these fit methods rely on a minimization of a χ^2 function, in which some measured parameters such as background rates, energy scale or $\Delta \mathrm{m}^2_{13}$ can be put in a "pull term" meaning that, during the fitting process, they are constraint by the measured central value and estimated error such that they are only allowed to vary within their estimated systematics, or left free to be adjusted without constraints (i.e. with $\sigma_x \to \infty$).

An additional constraint may be added using the residual $\bar{\nu}_e$ and the inclusive background estimation performed during the reactor-off periods.

4.4.a Rate and Spectral Shape

Double Chooz was the first of the three reactor θ_{13} experiments to perform a Rate+Shape analysis, not only taking into account the integral $\bar{\nu}_e$ flux as in a Rate-only analysis, but also the distortion of the spectral shape due to the energy dependence of $\bar{\nu}_e$ oscillation [54].

In order to perform such an analysis, one needs to know the expected un-oscillated neutrino rate for each energy bin of the spectrum. Equation 2.11 only provides the neutrino rate integrated over the full energy range, the neutrino rate per energy bin i is the integrated neutrino rate scaled by the weigh of bin i:

$$N_i^{\mathrm{pred}}(\mathrm{s}^{-1}) = \sum_{r=B_1,B_2} \left(P(\bar{\nu}_e \to \bar{\nu}_e, L_r, E_i) \times N_i^{\mathrm{pred},r} \right) + \sum_{k=1}^{N_{BG}} N_k^{\mathrm{BG,\ pred}} \tag{4.8}$$

$$\text{where}\quad N_i^{\mathrm{pred},r}(\mathrm{s}^{-1}) = \frac{1}{4\pi} N_p \epsilon \left(\frac{1}{L_r^2} \frac{P_{th,r}}{\langle E_f\rangle_r} \langle \sigma_f \rangle_r^i \right) \tag{4.9}$$

The last term in Equation 4.8 sums each background k in the background model with the expected contribution $N_k^{\mathrm{BG,\ pred}}$.

θ_{13} is measured by minimizing the χ^2 function defined as:

$$\begin{aligned}\chi^2 =&\sum_{i=1}^{B}\sum_{j=1}^{B}\left(N_i^{\text{obs}}-N_i^{\text{pred}}\right)M_{ij}^{-1}\left(N_j^{\text{obs}}-N_j^{\text{pred}}\right)+\sum_{k=1}^{5}\left(\frac{\epsilon_k}{\sigma_k}\right)^2\\ &+(\epsilon_a,\epsilon_b,\epsilon_c)\times\begin{pmatrix}\sigma_a^2 & \rho_{ab}\sigma_a\sigma_b & \rho_{ac}\sigma_a\sigma_c\\ \rho_{ba}\sigma_b\sigma_a & \sigma_b^2 & \rho_{bc}\sigma_b\sigma_c\\ \rho_{ca}\sigma_c\sigma_a & \rho_{cb}\sigma_c\sigma_b & \sigma_c^2\end{pmatrix}^{-1}\times\begin{pmatrix}\epsilon_a\\ \epsilon_b\\ \epsilon_c\end{pmatrix}\\ &+2\left(N_{\text{off}}^{\text{obs}}\times\ln\left(\frac{N_{\text{off}}^{\text{obs}}}{N_{\text{off}}^{\text{pred}}}\right)+N_{\text{off}}^{\text{pred}}-N_{\text{off}}^{\text{obs}}\right)\end{aligned}\tag{4.10}$$

where B is the number of energy bins considered, N_i^{obs} and N_i^{pred} refer respectively to the number of observed and predicted events in the bin i.

$\{M_{ij}\}$ is a 38x38 covariance matrix accounting for uncertainties and correlations between the B energy bins. $\{M_{ij}\}$ is defined as a sum of several covariance matrices accounting for statistical uncertainties ($\{M_{ij}^{\text{stat}}\}$) as well as systematic uncertainties from the reactor prediction ($\{M_{ij}^{\text{reactor}}\}$), the detection efficiency determination $\{M_{ij}^{\text{efficiency}}\}$, and background shape errors ($\{M_{ij}^{^9\text{Li}}\}$ and $\{M_{ij}^{\text{accidental}}\}$):

$$\{M_{ij}\}=\{M_{ij}^{\text{stat}}\}+\{M_{ij}^{\text{reactor}}\}+\{M_{ij}^{\text{efficiency}}\}+\{M_{ij}^{^9\text{Li}}\}+\{M_{ij}^{\text{accidental}}\}\tag{4.11}$$

Sensitivity studies have shown that the small amount of slope uncertainty present in the Correlated Background spectrum for DC-III n-Gd would have negligible impact on a Rate+Shape fit. Thus, no Correlated Background shape uncertainty $\{M_{ij}^{\text{correlated}}\}$ is included in the χ^2 function [84]. However, after the energy window has been opened up to 60 MeV for DC-III n-H and then 100 MeV for DC-IV, this matrix was added back to the χ^2 function.

The last term of the first line of Equation 4.10 is related to the rate and associated uncertainty of each background as well as the Δm_{31}^2 parameter and on the number of residual $\bar{\nu}_e$ emitted during the reactor-off periods $N_{\text{off-off}}$:

$$\begin{aligned}\sum_{k=1}^{5}\left(\frac{\epsilon_k}{\sigma_k}\right)^2 =& \left(\frac{\epsilon_{^9\text{Li}}-1}{\sigma_{^9\text{Li}}}\right)^2+\left(\frac{\epsilon_{\text{corr}}-1}{\sigma_{\text{corr}}}\right)^2+\left(\frac{\epsilon_{\text{acc}}-1}{\sigma_{\text{acc}}}\right)^2\\ &+\left(\frac{\epsilon_{N_{\text{off-off}}}-1}{\sigma_{N_{\text{off-off}}}}\right)^2+\left(\frac{\Delta m^2-\Delta m_{31}^2}{\sigma_{\Delta m_{31}^2}}\right)^2\end{aligned}\tag{4.12}$$

The parameters $\epsilon_{a,b,c}$ in the second line of Equation 4.10 characterize the correction for the systematic uncertainty on the energy scale. In order to account for MC/data discrepancies, the energy of the MC is corrected as follow:

$$E_{MC}\rightarrow\epsilon_a+\epsilon_b\times E_{MC}+\epsilon_c\times(E_{MC})^2\tag{4.13}$$

Uncertainties on these parameters are given by σ_a, σ_b, and σ_c, and the correlations between them are given by ρ_{ab}, ρ_{bc} and ρ_{ac}.

The third line in Equation 4.10 represents the contribution from the reactor-off measurements. Given the low statistics, the number of expected candidates is compared to the number of IBD candidates by a log-likelihood based on Poisson statistics.

4.4.b Reactor Rate Modulation

Double Chooz detectors are exposed to only two reactors, thus significant changes of the measured neutrino flux can be observed when one of the reactors is turned off.

Figure 4.11 shows the evolution of the observed IBD rates over time (data points), compared to the expected rate (red line), taking into account the thermal power and the data taking efficiency. Even though the measured IBD rates are affected by the data taking efficiency, the periods with one and two reactors on are clearly visible here. 7.6 days of reactor-off were acquired during the FD-I phase, the periods at zero for the other two configurations correspond to a period where no data have been acquired.

The Reactor Rate Modulation (RRM) analysis takes advantage of the frequent single-reactor and two-reactor periods alternation to compute the value of θ_{13}. Indeed, the observed IBD rate is proportional to the expected rate, with an offset corresponding to the background rate, independent of the expected IBD rate:

$$\begin{aligned} R^{\text{obs}} &= B + \left(1 - \sin^2(2\theta_{13})\eta_{\text{osc}}\right) R^{\bar{\nu}_e} \qquad (4.14) \\ \text{with} \quad \eta_{\text{osc}} &= \left(\frac{1.27\Delta m^2_{31}(\text{eV}^2)\langle L\rangle(\text{m})}{\langle E_\nu\rangle(\text{MeV})}\right) \end{aligned}$$

where $R^{\bar{\nu}_e}$ corresponds to the un-oscillated expected rate, and η_{osc} the average disappearance coefficient. η_{osc} is computed for each data point according to the relative thermal power of each reactor.

Data taking periods are separated in several bins of thermal power, one of which corresponding to the reactor-off period. A linear fit of $\sin^2(2\theta_{13})$ and of the background rate is performed. Three systematic error contributions need to be taken into account:

- Detection efficiency,
- Residual $\bar{\nu}_e$ prediction,
- Reactor flux prediction.

The later error depends on the thermal power level, as the in-core instrumentation is more precise at nominal power.

No error need to be accounted for concerning the background model, since the RRM analysis is independent of our understanding of the background, which is a constant contribution fitted at the same time as $\sin^2(2\theta_{13})$. The Rate+Shape method uses an exclusive model of each background contribution, whereas the RRM method performs an inclusive background estimation.

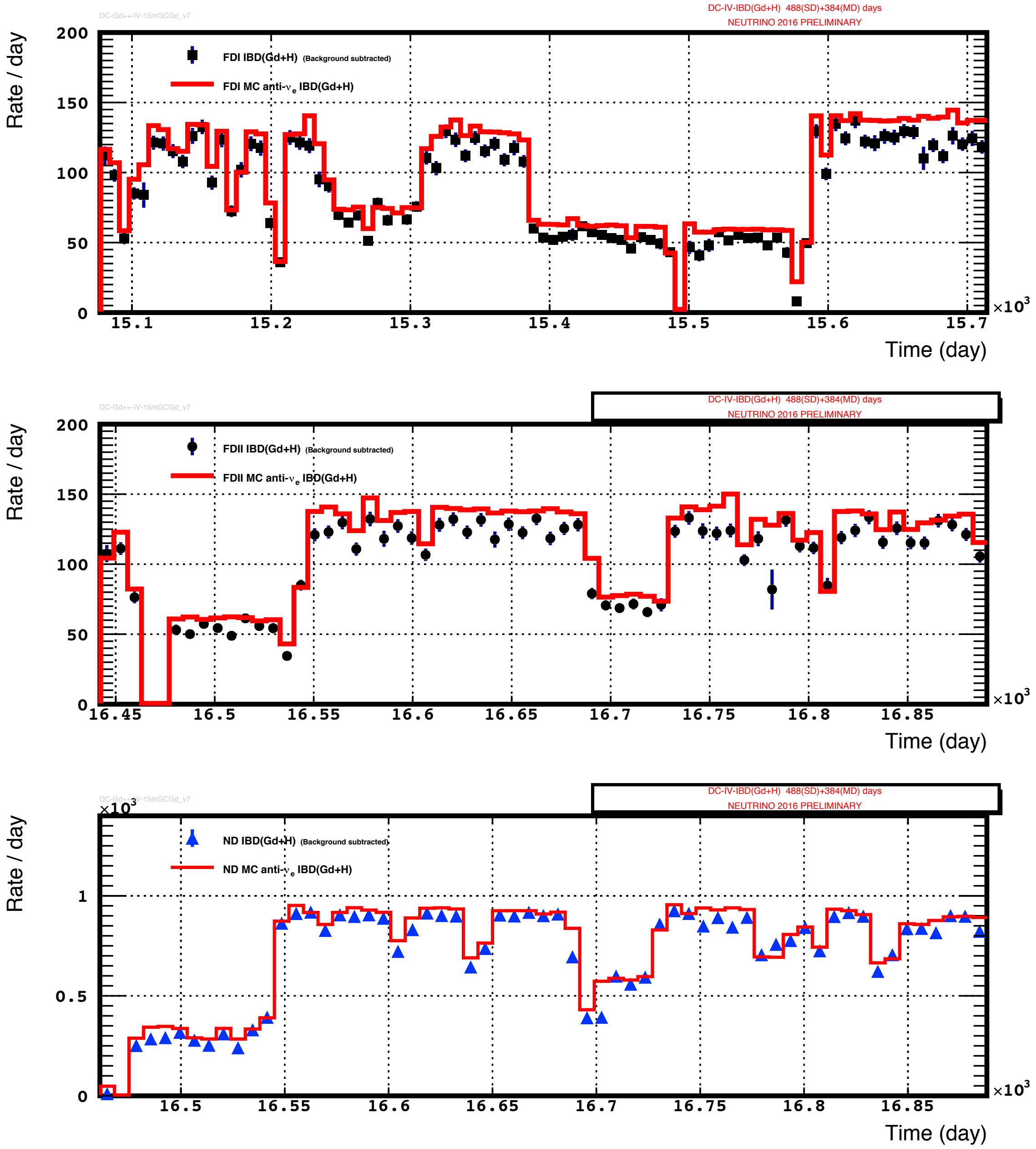


Figure 4.11: Evolution of the rate of IBD candidates over time for the Gd++ analysis, for FD-I (top), FD-II (middle) and ND (bottom). Periods with only one and two reactors on are visible. The red line shows the expected rate derived from the reactors thermal power and detector live time.

A χ^2 function for the fit can be defined as:

$$\begin{aligned}\chi^2 = \sum_{i=1}^{\mathrm{N_{bin}}} \frac{\left(R_i^{\mathrm{obs}} - R_i^{\mathrm{pred}} - B\right)^2}{\left(\sigma_{\mathrm{stat}}\right)^2} + \left(\frac{\epsilon_d}{\sigma_d}\right)^2 + \left(\frac{\epsilon_{\bar{\nu}_e}}{\sigma_{\bar{\nu}_e}}\right)^2 + \left(\frac{\epsilon_r}{\sigma_r}\right)^2 \\ + 2\left(N_{\mathrm{off}}^{\mathrm{obs}} \times \ln\left(\frac{N_{\mathrm{off}}^{\mathrm{obs}}}{N_{\mathrm{off}}^{\mathrm{pred}}}\right) + N_{\mathrm{off}}^{\mathrm{pred}} - N_{\mathrm{off}}^{\mathrm{obs}}\right) \qquad (4.15) \\ + \left(\frac{B - B^{\mathrm{pred}}}{\sigma_b}\right)^2\end{aligned}$$

The first line of Equation 4.15 sums over the single- and two-reactor on bins. R_i^{obs} and R_i^{pred} are respectively the observed and the predicted rates and B is the total background rate. Systematics are handled through the ϵ_d, $\epsilon_{\bar{\nu}_e}$ and ϵ_r parameters.
The second line corresponds to the reactor-off period. Here, as in the Rate+Shape case, the observed number of IBD candidates is compared to the expected one through a log-likelihood based on Poisson statistics.
The third line of the χ^2 definition is optional, and acts as an additional constraint using the background rate based on the background model estimations.

4.4.c DC-III Results for n-Gd and n-H Analyses

Here, the results of DC-III are described. They are based on a single-detector configuration with separate n-Gd and n-H IBD samples.
The results of DC-IV, latest Double Chooz results, are based on a two-detector configuration, with two phases of data taking for the Far Detector, and the full neutron capture spectrum (n-H and n-Gd together) is taken into account. DC-IV results will be presented in Section 4.5.

The DC-III analysis used data acquired 467.90 live days for the n-Gd and 462.72 live days for the n-H analyses. Two separate analyses were performed using neutron capture on Gd and H. A combined RRM analysis has been performed, improving the final result on $\sin^2(2\theta_{13})$.

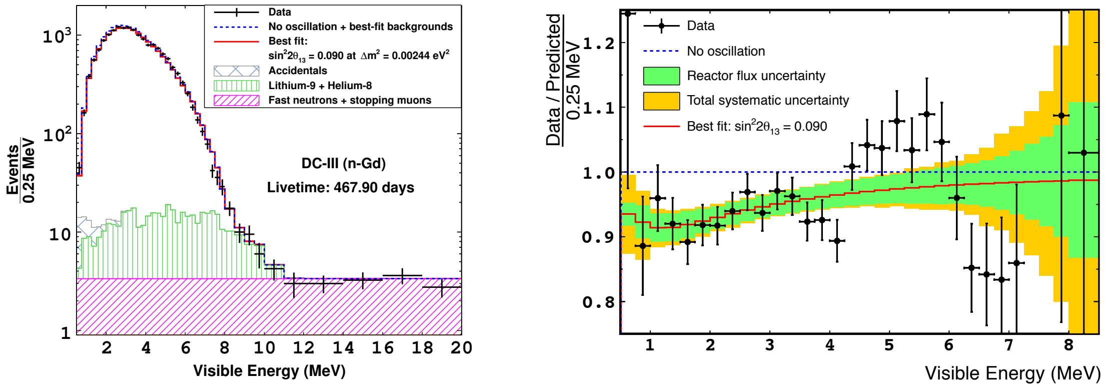
DC-III n-Gd

A fit of the prompt energy spectrum data shown in Figure 4.12(a) is performed using the χ^2 function defined in Equation 4.10. The fitted background shapes are superimposed, the grey crossed area corresponds to the Accidentals, the green stripped area corresponds to the Cosmogenic Backgrounds, and the purple hashed area corresponds to the Correlated Backgrounds, dominant at high energies.

The Δm^2_{31} used is $\Delta m^2_{31} = 2.44^{+0.09}_{-0.10} \times 10^{-3}\mathrm{eV}^2$ for the Normal mass Hierarchy (NH) and $\Delta m^2_{31} = 2.38^{+0.09}_{-0.10} \times 10^{-3}\mathrm{eV}^2$ for Inverted mass Hierarchy (IH), from MINOS measurements [28]. The energy spectrum of the positron signal was allowed to range from 0.5 MeV to 20 MeV although the IBD signal energy does not exceed 10 MeV. This allows to constrain the rate of Correlated Backgrounds (Fast Neutrons and Stopping Muons).

The ratio of the observed rate in each bin over the predicted rate with no oscillations and the fitted background model is presented in Figure 4.12(b). A negative deviation from the unity corresponds to a discrepancy between the no-oscillation hypothesis and the observed behaviour. The clear deficit below 3 MeV is interpreted as θ_{13}-driven oscillations.

Table 4.3 summarizes the input values and the best fit values for the fit parameters assuming a Normal mass Hierarchy. The number of residual $\bar{\nu}_e$ during reactor-off periods is computed over a live-time of 7.24 days. Values of $\sin^2(2\theta_{13})$ for both hierarchies are found to be:



(a)

(b)

Figure 4.12: 4.12(a): DC-III n-Gd analysis positron spectrum [56]. The black dots represent the data, the blue dashed line shows the no-oscillation hypothesis, and the red line shows the best fit with $\sin^2(2\theta_{13}) = 0.090$. The fitted background models are superimposed to the data. 4.12(b): Ratio of the data over the un-oscillated prediction [56]. Systematic uncertainties are shown in yellow (total), and in green (reactor flux contribution).

$$\sin^2(2\theta_{13})^{\mathrm{NH}} = 0.090^{+0.032}_{-0.029} \text{ with } \chi^2_{\mathrm{min}}/\mathrm{d.o.f} = 52.2/40 \tag{4.16}$$

$$\sin^2(2\theta_{13})^{\mathrm{IH}} = 0.092^{+0.033}_{-0.029} \text{ with } \chi^2_{\mathrm{min}}/\mathrm{d.o.f} = 52.2/40 \tag{4.17}$$

$$\tag{4.18}$$

A sanity check of this result is performed by removing the constraints on the normalizations of the cosmogenic and Correlated Backgrounds, yielding a value of

$$\sin^2(2\theta_{13}) = 0.088^{+0.030}_{-0.031} \tag{4.19}$$

with a ^{9}Li rate of $0.49^{+0.16}_{-0.14}$ events per day and a correlated rate of $0.541^{+0.052}_{-0.048}$ events per day. These results are consistent with the ones obtained in the standard analysis.

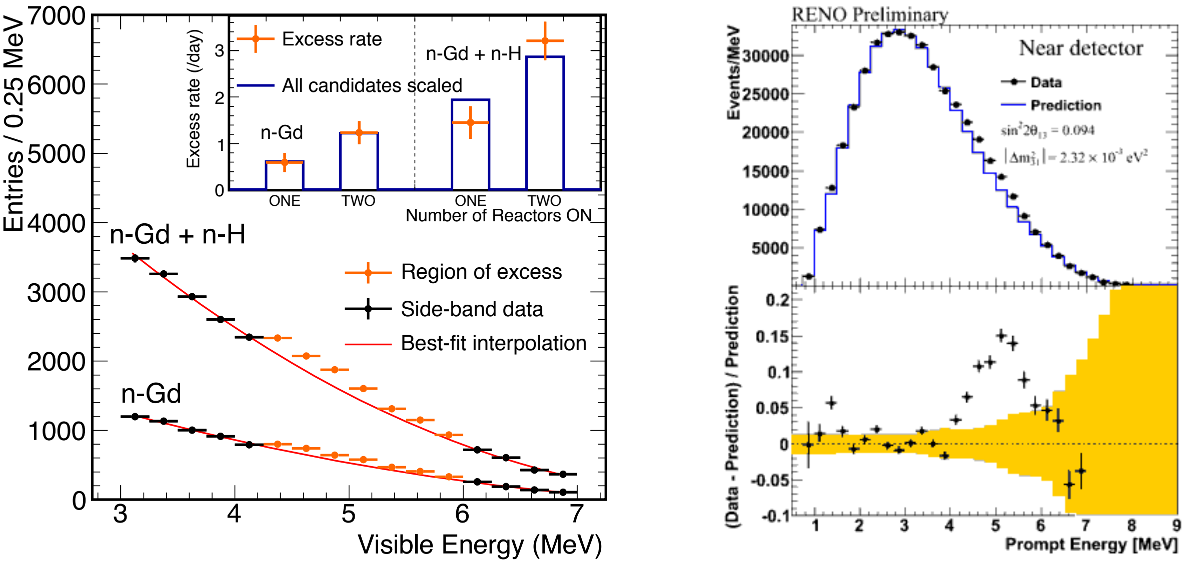
Spectrum Distortion

Above 4 MeV, an excess of events in the data compared to the oscillated predicted flux can be observed in Figure 4.12(b). This distortion has been studied in details by Double Chooz and reported with a 3σ significance [85]. It was confirmed by the RENO collaboration at the Neutrino 2014 conference [86], and later on by the Daya Bay collaboration at the ICHEP 2014 conference [87].

Investigations were carried out to understand the source of this distortion.
Among other disfavoured hypotheses, a problem in the energy scale was disproved based on the good agreement between data and MC at these energies for the calibrations.
Moreover, as shown in Figure 4.13(a), the excess is correlated to the thermal power and the

Table 4.3: Input and output parameters to the Rate+Shape fit for the DC-III n-Gd analysis [56], assuming a Normal mass Hierarchy.

Parameter	Input	Output
Energy Scale ϵ_a	0.27 ± 0.006	$-0.26^{+0.006}_{-0.005}$
Energy Scale ϵ_b	1.012 ± 0.008	$1.011^{+0.004}_{-0.006}$
Energy Scale ϵ_c	$(-1 \pm 6) \cdot 10^{-4}$	$(-6^{+7}_{-5}) \cdot 10^{-4}$
Rate Accidentals (day^{-1})	0.0701 ± 0.0026	0.0703 ± 0.0026
Rate Cosmogenics (day^{-1})	$0.97^{+0.41}_{-0.16}$	0.74 ± 0.13
Rate Correlated (day^{-1})	0.604 ± 0.051	$0.568^{+0.038}_{-0.037}$
Residual $\bar{\nu}_e$	1.57 ± 0.47	1.48 ± 0.47
$\Delta m^2_{31}(10^{-3}\text{eV}^2)$	$2.44^{+0.09}_{-0.10}$	$2.44^{+0.09}_{-0.10}$
$\sin^2(2\theta_{13})$		$0.090^{+0.032}_{-0.029}$
$\chi^2_{\text{min}}/\text{d.o.f}$		$52.2/40$



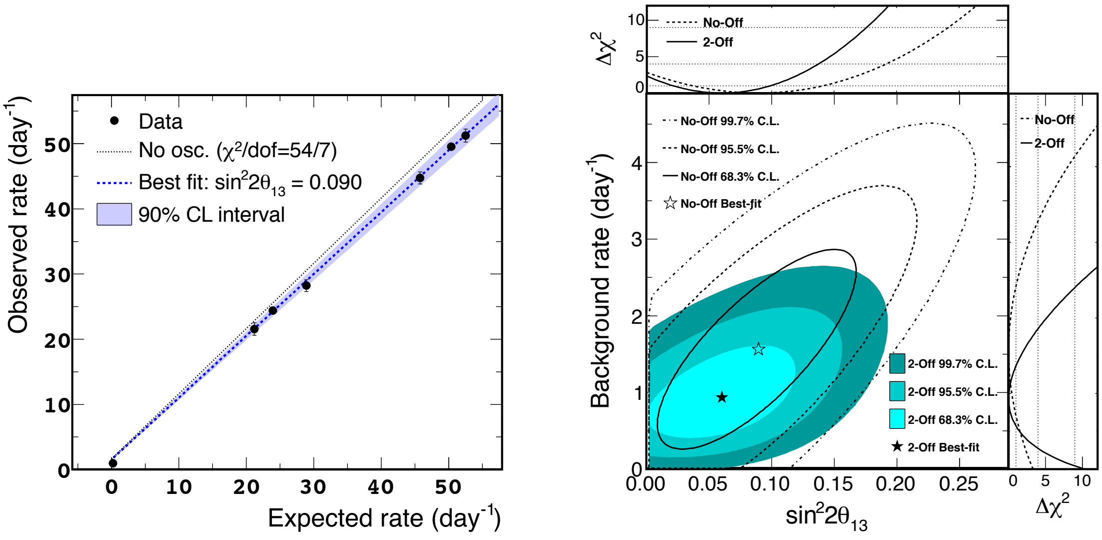
(a) (b)

Figure 4.13: 4.13(a): Close-up of the positron energy spectrum on the region of excess. The integrated number of candidates in that region is correlated with the total thermal power and the expected unmber of expected $\bar{\nu}_e$ as shown in the insert.
4.13(b): Prompt energy spectrum for the Near Detector of the RENO results shown at the Neutrino 2014 conference [86]. The bump at 5 MeV is clearly visible and is significantly inconsistent with their systematic errors (3.5σ).

expected rate of events. This behaviour could not be produced by a background.
Results from RENO and Daya Bay at this point had shown that this effect cancels when comparing near data to far data instead as data to MC as Double Chooz had to do in the single detector phase. This lead to suspect an issue with the expected spectra from reactor simulations.
The $\bar{\nu}_e$ spectra used in Double Chooz flux predictions originate from measurements of the integral β spectrum of the main isotopes (^{235}U, ^{239}Pu, and ^{241}Pu) at the ILL research reactor [65]. Experimental conditions made it difficult to measure decays that produce the neutrinos of higher energies that enter in the 5 MeV distortion because of their short life time.
All these observations tend to point toward an issue in the prediction of the neutrino flux emitted by the reactor cores, the flux appearing to be underestimated around 5 MeV.

The RRM analysis presented in Figure 4.14 compares data at different thermal power configurations. The amplitude of the spectral distortion is correlated to the reactor power, therefore, the RRM measurement is not affected by the spectral distortion.

Figure 4.14(a) shows the linear fit performed on the behaviour of the observed $\bar{\nu}_e$ rate versus the expected rate according to the thermal power. The black dotted line presents the no-oscillation hypothesis, and the blue dashed line presents the best fit with $\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035}$, with a χ^2/d.o.f= 4.2/6.



(a) (b)

Figure 4.14: DC-III n-Gd Reactor Rate Modulation analysis:
4.14(a): Observed $\bar{\nu}_e$ rate versus the expected rate for different thermal powers. The deviation of the slope of the linear fit from 1 is proportional to $\sin^2(2\theta_{13})$, while the intercept of the line provides a measurement of the background.
4.14(b): 2D representation of the evolution of the χ^2 function with the two fit parameters: $\sin^2(2\theta_{13})$ and the background rate.

The fit is performed with (Figure 4.14(a)) and without (Figure 4.14(b)) constraints on the background rate (i.e. with and without the the last term in the χ^2 function, constraining the background to the background model), providing cross-check of the background model and a background-independent measurement of θ_{13}.
Results on the background rate are presented in Equation 4.20:

$$\begin{array}{lcl} \text{BG input (day}^{-1}) & : & 1.64^{+0.41}_{-0.17} \\ \text{BG RRM}_{\text{constrained}}(\text{day}^{-1}) & : & 1.56^{+0.18}_{-0.16} \\ \text{BG RRM}_{\text{un-constrained}}(\text{day}^{-1}) & : & 0.93^{+0.43}_{-0.36} \end{array} \tag{4.20}$$

Figure 4.14(b) shows the 1, 2 and 3σ confidence intervals of the fit result with (coloured intervals) and without (clear intervals) using the reactor-off data. Adding the reactor-off constraints increases the sensitivity of the measurement, but both measurements are perfectly compatible, proof of the robustness of the technique.

DC-III n-H

Using neutron capture on Hydrogen instead of Gadolinium increases the active detection volume to include the γ-Catcher, thus increases the expected number of IBD candidate from the case of Target only. However using exclusively n-H captures lowers significantly the efficiency of the Target where $\sim 80\%$ of the neutrons capture on Gd. Moreover, the γ-catcher is more contaminated by Accidentals, and the vertex reconstruction is optimized for the Target. In addition, the liquid of the γ-Catcher has not been weighed before filling, thus the proton number in the γ-Catcher is not as well known as in the target, intrinsically limiting the sensitivity to θ_{13}.
Nevertheless, the proof of concept of the feasibility of an Hydrogen analysis have been demonstrated in [49], and was further improved in [50] by the development of the ANN to reject Accidentals.

In DC-III n-H, the Fast Neutron shape was found not to be flat, as was the case in n-Gd. Thus, the energy window of the prompt energy was opened up to 60 MeV, to provide a better handle on Correlated Background shape. Figure 4.15 shows the correlated spectrum, studied with both the Outer Veto tagged sample (blue dots) and the Inner Veto tagged sample (red dots), both scaled to the selected IBD candidates (black dots). The shape is described by a 3-parameter exponential function: $dN/dE = p_0 \times \exp(-p_1 \times E) + p_2$.

Those parameters enter as input parameters to the Rate+Shape fit presented in Table 4.4. Figure 4.16(a) shows the prompt energy spectrum and the fitted background models. Figure 4.16(b) shows the ratio of the background-subtracted data to the no-oscillation hypothesis. The red line corresponds to the best fit result for the n-H analysis. The deficit at low energy corresponds to the oscillation pattern. n-Gd (in blue) and n-H (in red) data are compared, along with their systematic uncertainties. The 5 MeV distortion is present and fully consistent with the one observed in n-Gd.

The robustness of the fit to the 5 MeV distortion was tested by performing cross-checks excluding the $[4.25; 5.75]$ MeV region, or fitting only the $[1; 4.0]$ MeV region. The values of $\sin^2(2\theta_{13})$ found were within 1σ of the initial result over the full energy window.

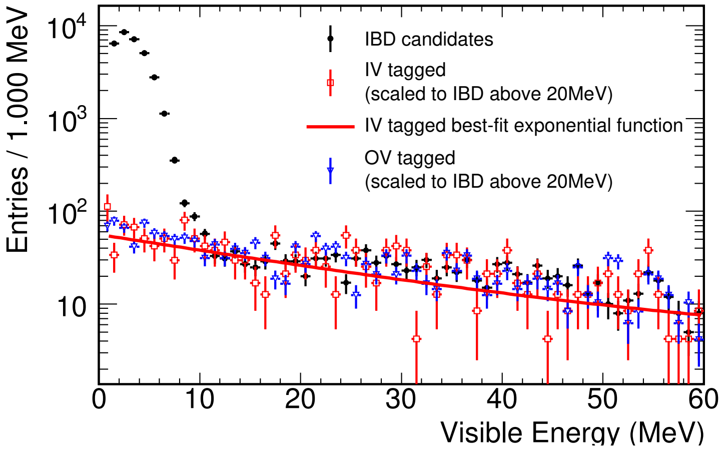


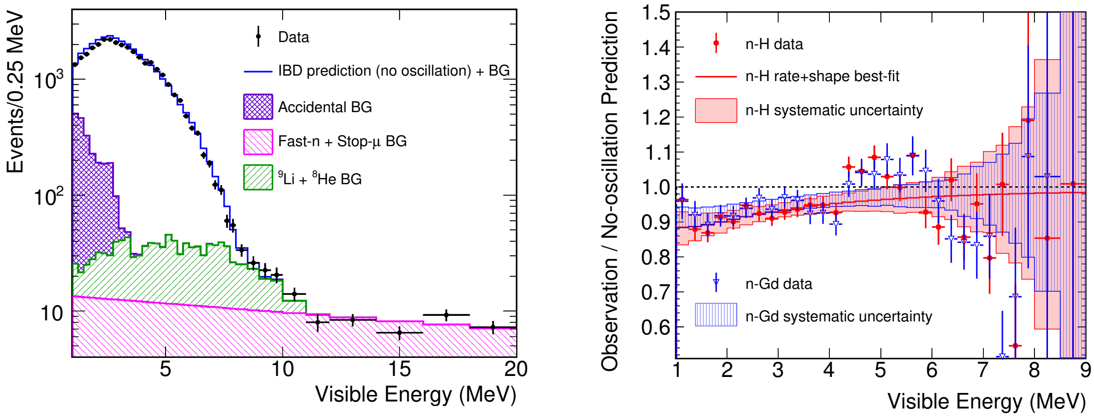
Figure 4.15: Increasing the energy range from 20 MeV in [56] to 60 MeV in [50] allows a better constraint on the Correlated Background shape, now defined with a 3-parameter exponential function.

Table 4.4: Input parameters to the Rate+Shape fit and their output values for the DC-III n-H analysis [50], assuming a Normal mass Hierarchy.

Parameter	Input	Output
Energy Scale ϵ_a (MeV)	0 ± 0.067	$-0.008^{+0.028}_{-0.020}$
Energy Scale ϵ_b	0 ± 0.022	$-0.007^{+0.007}_{-0.009}$
Energy Scale ϵ_c (/MeV)	0 ± 0.0006	$0.0005^{+0.0006}_{-0.0005}$
Rate Cosmogenics (day^{-1})	$0.95^{+0.57}_{-0.33}$	$1.60^{+0.21}_{-0.24}$
Rate Accidentals (day^{-1})	4.33 ± 0.011	4.33 ± 0.011
Rate Correlated (day^{-1})	1.55 ± 0.15	1.62 ± 0.10
FN shape p_0 (/MeV)	12.52 ± 1.36	12.33 ± 1.34
FN shape p_1 (/MeV)	0.042 ± 0.015	$0.037^{+0.015}_{-0.013}$
FN shape p_2 (/MeV)	0.79 ± 1.39	$0.39^{+1.48}_{-1.30}$
Residual $\bar{\nu}_e$	2.73 ± 0.82	2.81 ± 0.82
$\Delta m^2_{31}(10^{-3}\text{eV}^2)$	$2.44^{+0.09}_{-0.10}$	$2.44^{+0.09}_{-0.10}$
$\sin^2(2\theta_{13})$		$0.124^{+0.030}_{-0.039}$
χ^2_{min}/d.o.f		69.4/38

The reported value for θ_{13} for the Rate+Shape analysis is $\sin^2(2\theta_{13}) = 0.124^{+0.030}_{-0.039}$ with a $\chi^2_{\text{min}}/\text{d.o.f} = 69.4/38$ (instead of 30.7/32 when excluding the [4.25; 5.75] MeV region).

A RRM analysis was performed for the n-H sample. Figure 4.17 shows the results of the



(a) (b)

Figure 4.16: DC-III n-H Rate+Shape fit [50]:
4.16(a): Prompt energy spectrum with fitted background and no-oscillation expected spectrum. The black points are the data, the blue crossed area corresponds to Accidentals, the dominant background, the green striped area correspond to the Cosmogenic background, and the pink stripped area corresponds to the Correlated Background. The blue line presents the expected un-oscillated spectrum with the fitted background model.
4.16(b): Ratio of background-subtracted data to the no-oscillation prediction. n-Gd (in blue) and n-H (in red) data are compared along with their total systematic uncertainties.

RRM linear fit. The black dotted line presents the no-oscillation hypothesis, and the blue dashed line presents the best fit with $\sin^2(2\theta_{13}) = 0.095^{+0.038}_{-0.039}$, with a χ^2/n.o.f= 7.4/6.

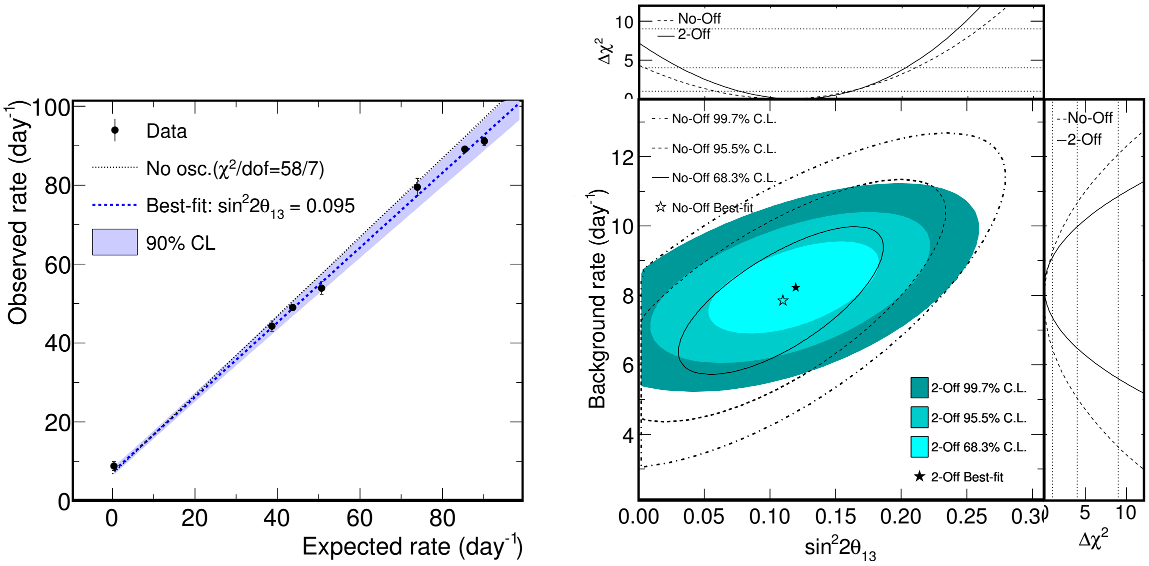
The fit is also performed without constraints on the background rate, and this with and without the reactor-off period, providing cross-checks of the background model and background-independent measurement of θ_{13}.
Results on the background rate are presented in Equation 4.21:

$$\begin{aligned} &\text{BG input (day}^{-1}) &&: \; 6.83^{+0.59}_{-0.36} \\ &\text{BG RRM}_{\text{constrained}}\ (\text{day}^{-1}) &&: \; 7.27 \pm 0.49 \\ &\text{BG RRM}_{\text{un-constrained}}\ (\text{day}^{-1}) &&: \; 8.23^{+0.88}_{-0.87} \end{aligned} \tag{4.21}$$

Figure 4.17(b) shows the 1, 2 and 3σ confidence intervals of the fit result without background constraints, and with (coloured intervals) and without (clear intervals) the reactor-off period. Adding the reactor-off period changes the central value slightly, but reduces the uncertainties by $\sim$ 20%.

Combined RRM Analysis

A combined RRM analysis was performed on both n-H and n-Gd samples. The two samples are independent. Although correlations exist on the detection systematics in the Target, detection systematic uncertainties are considered fully uncorrelated in order to make the most



(a)

(b)

Figure 4.17: DC-III n-H Reactor Rate Modulation analysis:
4.17(a): Observed $\bar{\nu}_e$ rate versus the expected rate for different thermal powers with background constraints and using reactor-off data.
4.17(b): χ^2 scan without background constraints, with and without reactor-off data.

conservative assumption. The reactor flux systematics uncertainties, however, are fully correlated as both datasets were taken at the same time and with the same detector.

Observed IBD candidate rates are plotted Figure 4.18(a) as a function of the expected rates for both n-Gd and n-H samples, as well as the best fit result for both with background constraints and using reactor-off period.

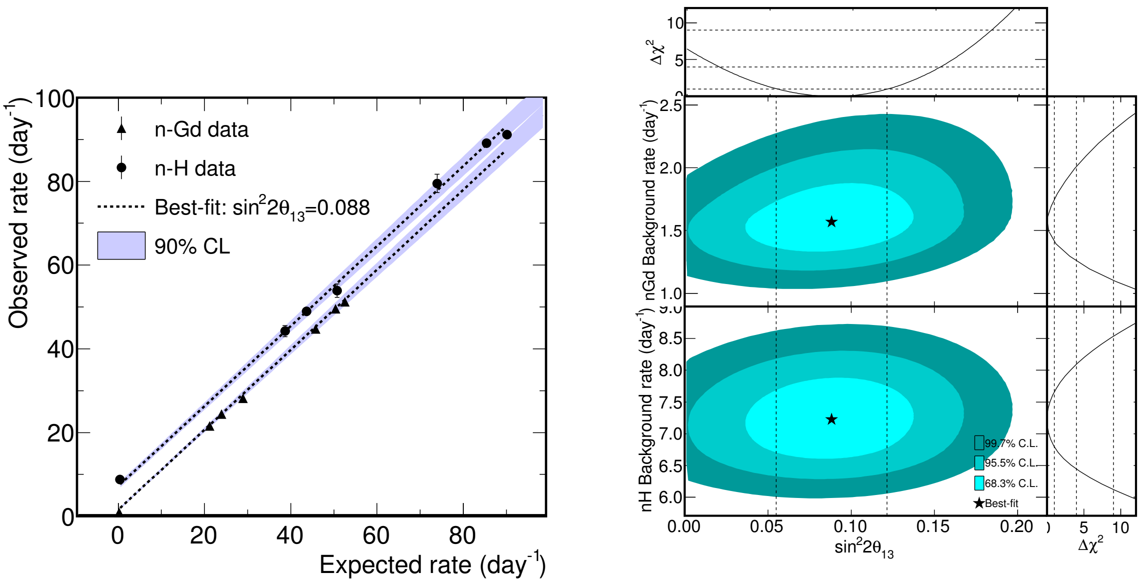
Figure 4.18(b) shows the χ^2 scans when releasing the background constraints and using the reactor-off data.

The best fit result for the joint analysis of DC-III n-H and n-Gd is $\sin^2(2\theta_{13}) = 0.088 \pm 0.033$ with a $\chi^2_{\text{min}}/\text{d.o.f} = 11.0/13$.

The n-H analysis achieved a level of sensitivity comparable to n-Gd such that adding the n-H information in a joint fit improves the uncertainty on the determination of $\sin^2(2\theta_{13})$.

4.5 θ_{13} analysis in the two-detectors configuration

DC-IV is the first Double Chooz analysis in the full 2-detectors configuration. This milestone is particularly important for the experiment, as the commissioning of the Near Detector allows to greatly reduce systematics linked to the reactor flux normalization and to the detection efficiencies.



(a)

(b)

Figure 4.18: Combined n-Gd+n-H RRM fit.
4.18(a): RRM fit with background constraints.
4.18(b): χ^2 scan for the combined fit without background constraints, but using the reactor-off period. The results for n-Gd are shown at the top, and those for n-H are shown at the bottom.

Table 4.5: Summary table of the different values of $\sin^2(2\theta_{13})$ measured in DC-III (assuming a Normal Hierarchy).

Analysis	$\sin^2(2\theta_{13})$ value
n-Gd R+S	$0.090^{+0.032}_{-0.029}$
n-Gd RRM	$0.090^{+0.034}_{-0.035}$
n-H R+S	$0.124^{+0.030}_{-0.039}$
n-H RRM	$0.095^{+0.038}_{-0.039}$
combined RRM	0.088 ± 0.033

Suppression of Reactor Systematics

The best case for a θ_{13} experiment with multiple reactors and detectors is to be able to position all the detectors on the iso-flux line, i.e. position the detectors in such a manner that the contribution of each reactor to the total flux seen by a detector is the same for each detector. In this particular case, and assuming the running periods of all detectors are identical, the Near Detector becomes a perfect monitor of the reactor $\bar{\nu}_e$ flux for the others, except for the geometrical effect of solid angle acceptance. $\sin^2(2\theta_{13})$ could then be extracted from a simple comparison between the two detectors, and the un-oscillated flux prediction then becomes irrelevant. Moreover, in that exact case, the reactor flux prediction systematics on flux normalization, formerly of the order of $\sim 1.7\%$ for the single-detector phase, are totally suppressed.

Systematic uncertainties on flux prediction in the Near and Far Detectors can be divided into two kinds: correlated and uncorrelated. Correlated uncertainties are linked to identical prediction biases in both detectors, and un-correlated uncertainties are linked to different prediction biases. In multi-detector configuration, correlated uncertainties are fully suppressed, leaving only the uncorrelated part of the uncertainty that could be partially reduced given the iso-flux conditions.

Concerning Double Chooz, the location of the detectors was optimized to best verify the iso-flux condition. However, because of the difficulties to install an experiment on a nuclear power plant site, it was not possible to place the two detectors exactly on the iso-flux line. Therefore, the prediction flux from each reactor will remain necessary, and systematic uncertainties cannot be fully suppressed.

The effects of iso-flux conditions on the full systematic uncertainty on the un-oscillated $\bar{\nu}_e$ prediction has been studied in [88]. In this article, a generic study was performed in order to compute the suppression factor (SF) on flux prediction systematics in multi-detector configuration for the three reactor $\bar{\nu}_e$ experiments for short baselines Double Chooz, RENO and Daya Bay.

The SF accounts only for the error suppression provided by the iso-flux status of the experimental setup and reflects the ability of each experiment to reduce the overall reactor uncertainty relative to the simplest case of one detector with one reactor configuration where no cancellation is expected ($\sigma_{\phi^{pred}} \simeq$1,7%). The convention used is that the smaller the SF, the smaller the final reactor flux error systematic. SF ranges in the interval [0,1], a null value for SF stands for a total cancellation of reactor systematics while a value of 1 stands for no reactor systematics reduction.

Figure 4.19(a) shows the SF for the experimental set-up of Double Chooz as a function of the anti-symmetry between the fluxes from reactors B1 and B2 and of the Uncertainty Type Asymmetry defined as:

$$\text{Flux Asymmetry} = (\phi_{B2} - \phi_{B1})/(\phi_{B2} + \phi_{B1}) \tag{4.22}$$

where ϕ_{B1} and ϕ_{B2} are the fluxes emitted by reactors B1 and B2 respectively, and :

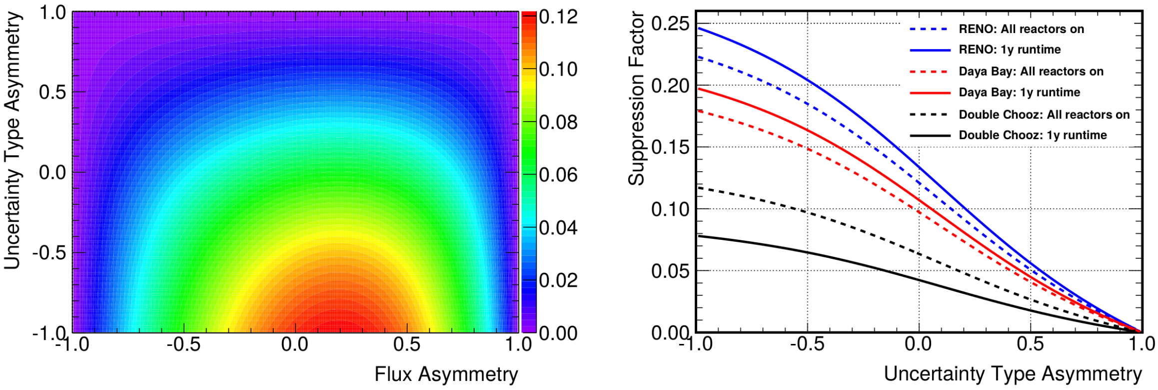
$$\text{Uncertainty Type Asymmetry} = (\delta^c - \delta^u)/(\delta^c + \delta^u) \tag{4.23}$$

where δ^c and δ^u correspond respectively to the correlated and uncorrelated components of the total reactor flux uncertainty per individual reactor.

On this plot, a SF of zero means that the systematic uncertainty on the reactor flux is fully suppressed. This is only achievable when one of the two reactors is off (flux asymmetry = -1), or when only the correlated component of the uncertainty is left (uncertainty type asymmetry = 1). For intermediary cases, the two reactors have different thermal power levels and thus different flux contributions. The least favourable case is when the Uncertainty Type Asymmetry $= -1$ and the Flux Asymmetry $\simeq 0.2$, which leads to a maximal SF of ~ 0.12.

Figure 4.19(b) shows the evolution of SF for Double Chooz, Daya Bay and RENO set-ups as a function of the Uncertainty Type Asymmetry. Two cases are taken in consideration, the first one, for a data taking period with all reactors at nominal power, and the second one, considering a data taking period of one year, while assuming a two-month fuel re-processing period. It can be noted that:

- Thanks to its simple configuration, Double Chooz is the experiment with the lowest SF.
- Despite having numerous reactors, Daya Bay shows a SF higher than RENO because of the use of multiple Near experimental halls, while RENO only uses one to monitor six reactors.



(a) SF for the DC multidetector configuration.

(b) SF comparison for DC, Daya Day and RENO experiments.

Figure 4.19: 4.19(a): Double Chooz suppression factor on reactor flux uncertainty in multiple detector configuration as a function of the reactor flux uncertainty type asymmetry (y-axis) and the reactor power flux asymmetry (x-axis) defined as $(\phi_{B2} - \phi_{B1})/(\phi_{B2} + \phi_{B1})$ (i.e. the flux difference between the reactor B1 and B2).
4.19(b): The variation of SF to the reactor uncertainty type asymmetry, defined as $(\delta^c - \delta^u)/(\delta^c + \delta^u)$, is shown for DC-II, RENO and Daya Bay. All experiments have assumed, so far, their errors to be reactor uncorrelated; i.e. $(\delta^c - \delta^u)/(\delta^c + \delta^u) = -1.0$. Double Chooz benefits the best SF (~ 0.1) due to its almost iso-flux site. RENO has a large spread in the distances between the reactors and detectors and therefore has the highest SF. Both full reactor power (dashed lines) and simple refuelling scenario (solid lines) are shown.

Reactor systematics in single-detector configurations for FD-I, FD-II and ND are summarized in Table 4.6. The SF for DC-IV FD-II and ND was computed to 0.08, corresponding to a reduction of the reactor systematics by a factor ~ 20 from the single-detector configuration. This leads to a remaining reactor flux normalization of $\sim 0.07\%$ out of the initial $\sim 1.7\%$.

4.5.a Reduction of Detection Systematics

In addition to suppressing reactor systematics, the multi-detector configuration allows to lower the uncertainty linked to the signal normalization because of the correlations in the detection systematics: a detection bias correlated between the Near and Far Detectors does not affect the ratio of events in both detectors, provided that the bias is identical.

Table 4.6: Reactor uncertainties for DC-IV analyses. Numbers in red are correlated in 2-detector analysis, and numbers in blue cancel when the two detectors run at the same time.

Origin		FD-I (%)	FD-II (%)	ND (%)
Bugey-4 normalization	$\langle \sigma_f \rangle^{B4}$	1.4	1.4	1.4
Energy per fission	$\langle E_f \rangle$	0.16	0.16	0.16
Spectrum⊕Cross Section	$S_k(E) \oplus \sigma_{IBD}(E)$	0.20	0.20	0.20
Baselines	L	< 0.01	<0.01	0.01
Fission Fraction	α_k	0.82	0.74	0.73
Thermal Power	Pth	0.44	0.44	0.44
Total (single detector)		1.70	1.66	1.66
Correlation (FD-I:FD-II)	ρ(FD-I : FD-II)	0.72		
Correlation (FD-II:ND)	ρ(FD-II : ND)		>0.99	

Table 4.7 summarizes the detection systematics for each of the FD-I, FD-II and ND detectors. It can be noted that the detector systematics are suppressed by a factor of ~ 2 in the two-detector configuration compared to the single-detector configuration.

Table 4.7: DC-IV detection uncertainties. Numbers in blue correspond to uncorrelated uncertainties in multi-detector analysis.

Origin	FD-I (%)	FD-II (%)	ND (%)
Background Vetoes	0.11 (0.11)	0.09 (0.09)	0.02 (0.02)
IBD selection	0.21 (0.21)	0.16 (0.16)	0.07 (0.07)
Gd fraction	0.25 (0.14)	0.26 (0.15)	0.28 (0.19)
Spill effects	0.27 (0)	0.27 (0)	0.27 (0)
Proton Number	0.30 (0)	0.30 (0)	0.30 (0)
Total	0.49 (0.26)	0.47 (0.22)	0.38 (0.15)

4.5.b DC θ_{13} results for DC-IV analysis

Double Chooz presented the first preliminary results for θ_{13} analysis in the two-detector configuration (DC-IV) at the Moriond 2016 Conference [58]. Those results bore on the n-Gd analysis only and a data taking period of approximately nine months. An update was presented at the NEUTRINO2016 conference [89], two new analyses (n-Gd and Gd++) were presented, bearing on a 15 months live time period in the two-detector configuration. However, the final θ_{13} result associated with to these analyses could not be presented, but a release of the results in a CERN seminar in September 2016 is announced.

In this section, we will focus on the fit results presented in the Moriond Conference [58].

4.5.c θ_{13} Fit

Different approaches can be used to perform the θ_{13} fit in a multiple detector configuration. A first method consists in building a χ^2 function considering the Near Detector data as prediction for the Far Detector (data-to-data method). The main advantage of this technique is to reduce the dependency to the MC, and therefore to reduce the potential bias due to the MC. The main drawback is that there is no Near Detector data to compare to the Far Detector in the single-detector phase (FD-I). A hybrid method is therefore necessary, to compare data to data for the two-detector phase, and data to MC for the single-detector phase. An additional constraint is introduced on the background shapes in the Near to Far Detector comparison.

For the results presented in Moriond, a data-to-MC approach similar to the one used in the single-detector phase was used to fit θ_{13}. In this method, the χ^2 function compares the data of each detector to a reactor flux prediction:

$$\begin{aligned}
\chi^2 = & \sum_{i=1}^{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} \Big((N^{\text{pred}}_{\text{FD-I}})_i - (N^{\text{obs}}_{\text{FD-I}})_i, (N^{\text{pred}}_{\text{FD-II}})_i - (N^{\text{obs}}_{\text{FD-II}})_i, (N^{\text{pred}}_{\text{ND}})_i - (N^{\text{obs}}_{\text{ND}})_i \Big) \times M^{-1}_{ij} \begin{pmatrix} (N^{\text{pred}}_{\text{FD-I}})_j - (N^{\text{obs}}_{\text{FD-I}})_j \\ (N^{\text{pred}}_{\text{FD-II}})_j - (N^{\text{obs}}_{\text{FD-II}})_j \\ (N^{\text{pred}}_{\text{ND}})_j - (N^{\text{obs}}_{\text{ND}})_j \end{pmatrix} \\
& + \frac{(\Delta m^2 - \Delta m^2_{31})^2}{\sigma^2_{\Delta m^2_{31}}} \\
& + \sum^{\text{FN,Acc}} (B_{FD-I} - (B_{\text{FD-I}})_0, B_{FD-II} - (B_{\text{FD-II}})_0, B_{ND} - (B_{\text{ND}})_0) \\
& \times \begin{pmatrix} (\sigma_{B_{\text{FD-I}}})^2 & \rho_{B_{\text{FD-I}}B_{\text{FD-II}}}(\sigma_{B_{\text{FD-I}}}\sigma_{B_{\text{FD-II}}}) & \rho_{B_{\text{FD-I}}B_{\text{ND}}}(\sigma_{B_{\text{FD-I}}}\sigma_{B_{\text{ND}}}) \\ \rho_{B_{\text{FD-I}}B_{\text{FD-II}}}(\sigma_{B_{\text{FD-I}}}\sigma_{B_{\text{FD-II}}}) & (\sigma_{B_{\text{FD-II}}})^2 & \rho_{B_{\text{FD-II}}B_{\text{ND}}}(\sigma_{B_{\text{FD-II}}}\sigma_{B_{\text{ND}}}) \\ \rho_{B_{\text{FD-I}}B_{\text{ND}}}(\sigma_{B_{\text{FD-I}}}\sigma_{B_{\text{ND}}}) & \rho_{B_{\text{FD-II}}B_{\text{ND}}}(\sigma_{B_{\text{FD-II}}}\sigma_{B_{\text{ND}}}) & (\sigma_{B_{\text{ND}}})^2 \end{pmatrix}^{-1} \\
& \times \begin{pmatrix} B_{FD-I} - (B_{\text{FD-I}})_0 \\ B_{FD-II} - (B_{\text{FD-II}})_0 \\ B_{ND} - (B_{\text{ND}})_0 \end{pmatrix} \\
& + 2\left(N^{\text{obs}}_{\text{off}} \times \ln\left(\frac{N^{\text{obs}}_{\text{off}}}{N^{\text{pred}}_{\text{off}}} \right) + N^{\text{pred}}_{\text{off}} - N^{\text{obs}}_{\text{off}} \right)
\end{aligned} \tag{4.24}$$

The first term of the χ^2 definition compares the data and MC spectrum for each detector k, with k = FD-I, FD-II and ND. N^{obs}_k and N^{pred}_k are respectively the number of measured and expected events in the detector k. The matrix $\{M_{ij}\}$ is a covariance matrix that contains systematic uncertainties on the reactor flux, the energy scale, the detection efficiency, and the Accidental and ^{9}Li shapes, as well as the statistical uncertainties. It also accounts for the correlation between systematic uncertainties, and thus, the SF previously described.
The second term corresponds to the pull term associated to the Δm^2_{13} parameter using the MINOS measurement.
The third term accounts for the constraints on Fast Neutrons and Accidental background rates, B_k and $(B_k)_0$ are respectfully the fitted background rate and the estimated background rate for the detector k. Uncertainties on background rates are noted σ_{B_k} and the inter-detector correlation on background rates are noted $\rho_{B_k B_j}$.
The fourth term represents the contribution from the reactor-off measurements for the FD-I

and is the same as in the Rate+Shape fit of DC-III.

The advantage of this method is that it is very similar to the one used in DC-III, as being a simple extension of the single-detector analysis, and allows to keep the statistics of the single-detector phase, which would not be possible in a simple Near data to Far data comparison. However, the measurement still relies heavily on the MC, and a possible normalization bias could remain if the two detectors are not normalised to the MC the same way.

4.5.d θ_{13} Results

The event selection is the same as presented in Section 4.2. At the Moriond Conference, the Stopping Muon rejection used the Chimney Pulse Shape variable (CPS). The CPS variable, defined in more details later in Section 6.4.f, is the ratio of the event's pulse shape likelihoods computed at the reconstructed vertex, and assuming that the event took place in the Chimney. The Stopping Muon rejection was achieved by the following cut:

$$CPS_{\text{prompt}} + CPS_{\text{delayed}} > 2 \tag{4.25}$$

The distribution of CPS prompt versus delayed can be found in Figure 4.20. The red line materializes the cut position. Events below this line are rejected as Stopping Muons.

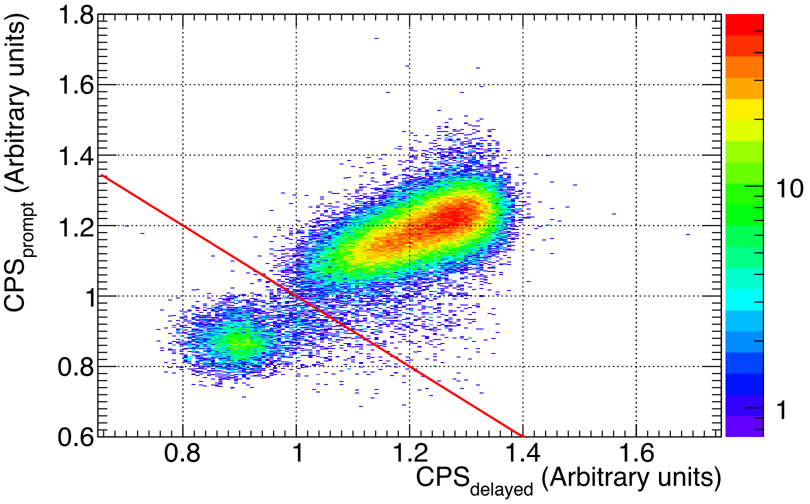


Figure 4.20: Distribution of the CPS variable for the prompt event versus the delayed event. Events below the red line are rejected.

The remaining background rates are summarized in Table 4.8 for the results presented at the Moriond Conference. New background studies have been performed since then, and a new analysis using neutron captures on Gd and H was developped in parallel, called Gd++. The latest background rates, that I had the opportunity to present in a poster at the NEUTRINO2016 Conference [90], are summarized in Table 4.9. The new analyses consistently lowered the uncertainty on the individual background rates, in an effort to reduce the systematic uncertainties in the output of the θ_{13} fit.

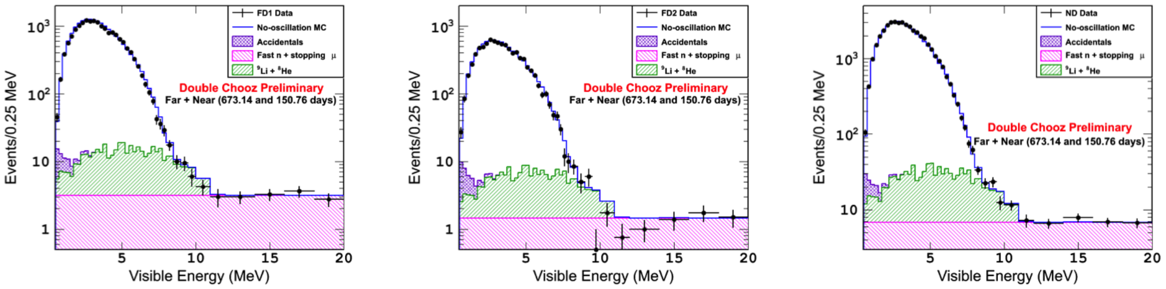
Table 4.8: DC-IV Backgrounds summary for the n-Gd analysis presented at the Moriond 2016 conference [58].

BG source	FD-I	Reactor off	FD-II	ND
Accidental (day^{-1})	0.070 ± 0.003		0.106 ± 0.002	0.344 ± 0.002
Fast Neutrons (day^{-1})	0.586 ± 0.061			3.42 ± 0.23
Cosmogenics (day^{-1})	$0.97^{+0.41}_{-0.16}$			5.01 ± 1.43

Table 4.9: DC-IV Backgrounds summary for n-Gd and Gd++ analyses presented at NEUTRINO2016.

	n-Gd			Gd++		
BG source	FD-I	FD-II	ND	FD-I	FD-II	ND
Accidental (day^{-1})	0.069 ± 0.002	0.118 ± 0.006	0.303 ± 0.007	3.994 ± 0.004	4.284 ± 0.039	3.104 ± 0.004
Fast Neutrons (day^{-1})	0.43 ± 0.03	0.51 ± 0.04	4.00 ± 0.15	2.60 ± 0.11	2.48 ± 0.10	20.77 ± 0.43
Cosmogenics (day^{-1})	0.87 ± 0.42		4.67 ± 1.42	2.59 ± 0.61		11.11 ± 2.96
Reactor-OFF (day^{-1})	0.76 ± 0.38	———	———	$8.90 \pm .20$	———	———
Signal/ΣBG	23.80	27.00	31.50	10.30	11.10	20.60

The χ^2 function defined in Equation 4.24 includes the three detectors and their correlations, i.e. the three spectra are fitted at the same time. Figure 4.21 shows the prompt energy spectrum for each detector for the n-Gd analysis presented at Moriond 2016. The fitted spectra of individual backgrounds are superimposed to the data and the blue line corresponds to the un-oscillated best fit.



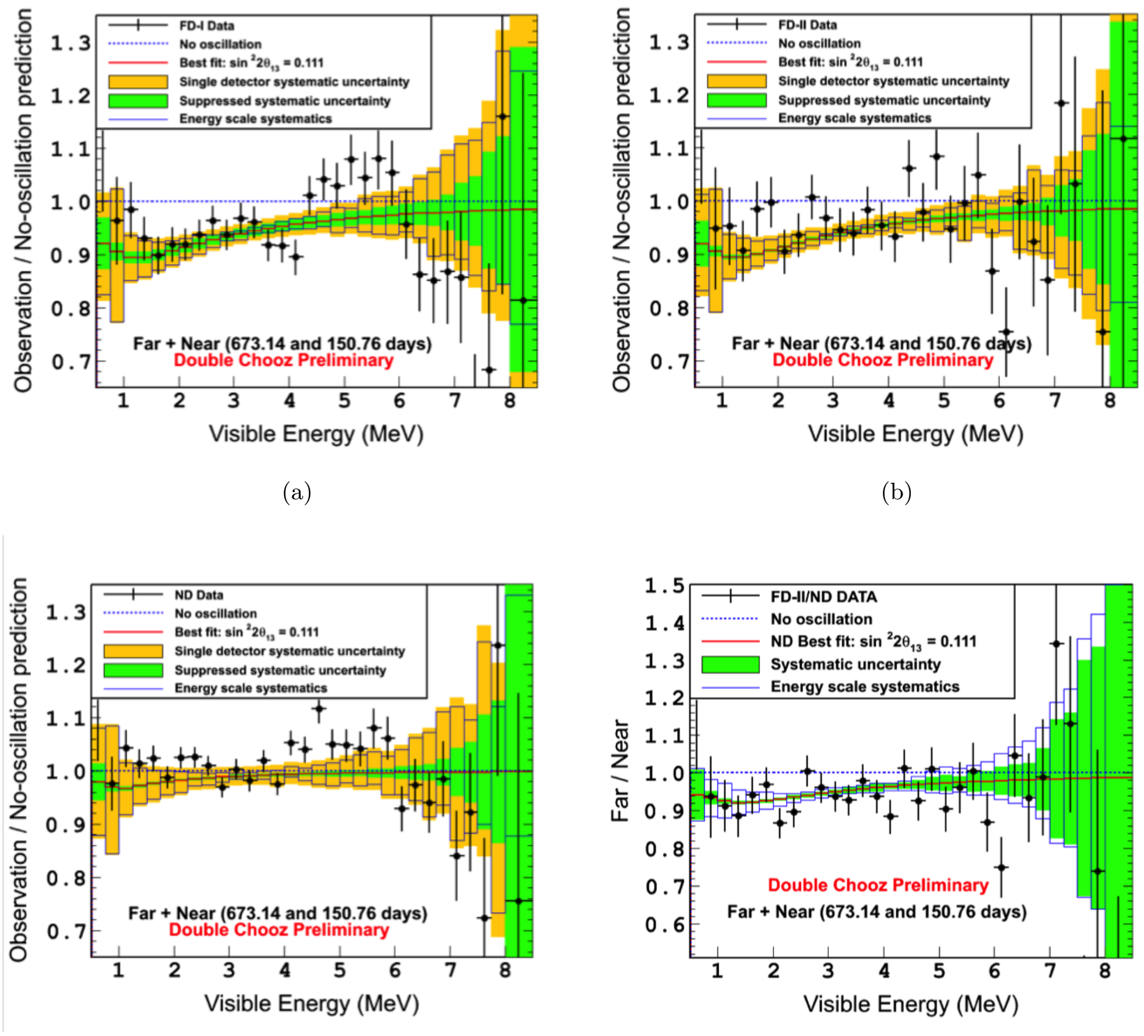
(a)

(b)

(c)

Figure 4.21: Prompt energy spectra of the DC-IV n-Gd analysis presented at Moriond 2016. The black dots are the data points, the blue line is the un-oscillated prediction, and the coloured histograms are the individual backgrounds.

Figure 4.22 shows the ratio of the data with subtracted backgrounds to the un-oscillated prediction. The yellow boxes show the systematic uncertainties with the assumption of only one detector and the green boxes show the fraction of the systematic uncertainties that can be reduced in the multi-detector treatment. The Near Detector is slightly impacted by θ_{13}, thus, a small deficit should be visible when comparing to the un-oscillated prediction. The deficit is not observed, hinting towards a possible bias in the flux normalization of the MC in the Near Detector.



(a) (b)

(c) (d)

Figure 4.22: Ratio of the data to un-oscillated prediction after subtraction of the background model:
4.22(a): Ratio of the data to prediction FD-I.
4.22(b): Ratio of the data to prediction FD-II.
4.22(c): Ratio of the data to prediction ND.
4.22(d): Ratio of the Far to Near data for the two-detector configuration (FD-II/ND) superimposed to the best fit function and errors in the data-to-MC case.

Figure 4.22(d) shows the ratio of the Far data to the Near data during the two-detector phase (FD-II/ND). It is to be noted that the distortion at 5 MeV disappears when comparing data-to-data, consistent with the behaviour observed by RENO and Daya Bay. The fit and errors displayed on this plot are computed by the three data-to-MC fits, and are simply overlayed here for reference.

The measured value for $\sin^2(2\theta_{13})$ is $\sin^2(2\theta_{13}) = 0.111 \pm 0.018$ with a $\chi^2/d.o.f. = 128.8/120$, corresponding to an exclusion of the no-oscillation hypothesis at 5.8σ C.L.

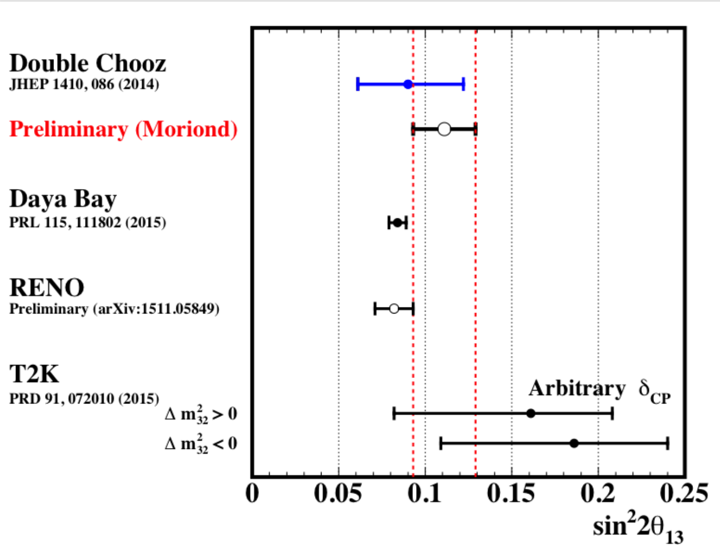


Figure 4.23: Comparison of the latest Double Chooz result to the other θ_{13} experiments in the world.

The measured rates of Fast Neutrons and ^{9}Li backgrounds are:

Fast Neutrons	$0.534 \pm 0.035\,\mathrm{d}^{-1}$ (FD)	$3.53 \pm 0.16\,\mathrm{d}^{-1}$ (ND)
Cosmogenic (^{9}Li)	$0.75 \pm 0.14\,\mathrm{d}^{-1}$ (FD)	$4.89 \pm 0.78\,\mathrm{d}^{-1}$ (ND)

Figure 4.23 shows the comparison of the latest value of $\sin^2(2\theta_{13})$ measured by Double Chooz to the other θ_{13} experiments in the world. A tension with the current best measurement by Daya Bay of 1.6σ appears. This situation is likely to evolve with the up-coming DC-IV analysis with more statistics and a better handling of the MC normalization.

4.6 Conclusion

Statistical uncertainties can only be reduced with increasing the live time, however, by demonstrating the feasibility of a n-H based analysis with a final uncertainty of a similar order than the n-Gd analysis, the possibility to increase the size of the active detection volume allowed to double the IBD rate. This was made possible by an important effort on understanding and rejecting backgrounds, especially the Accidental background, predominant in the n-H analysis. The knowledge acquired on the backgrounds, as well as on the detection efficiency, allowed to lower these systematic uncertainties even faster than the reduction of the statistical uncertainty, proportional to $1/\sqrt{\text{live-time}}$. The Rate+Shape fit helped constrain the background uncertainty by fitting simultaneously the backgrounds and the IBD rates and spectral shapes. These behaviours are shown in Figure 4.24 where the evolution of the different sources of uncertainties is illustrated. Each vertical series of points correspond to an analysis and their corresponding live-time and uncertainty level. The latest series of points correspond to the DC-IV n-Gd analysis presented at the Moriond 2016 conference. Introducing the inter-detector and -reactor correlations allows to greatly reduce the uncertainties compared to their un-correlated treatment.

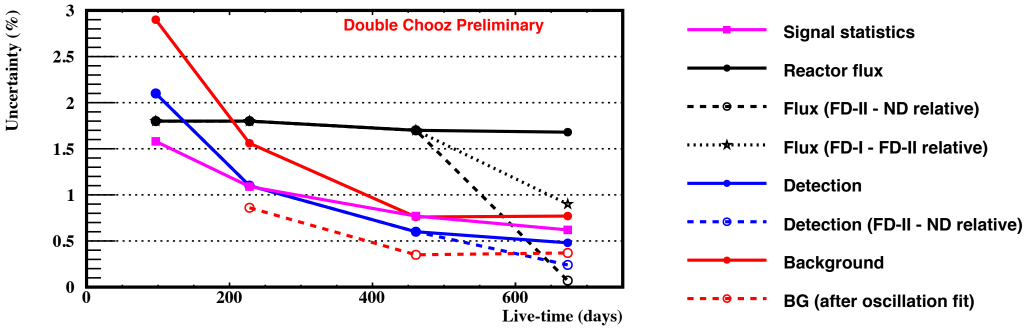


Figure 4.24: Evolution of the various sources of uncertainties as a function of the live-time for the successive n-Gd analyses of Double Chooz.

The table 4.10 summarizes the successive values of $\sin^2(2\theta_{13})$ measured by Double Chooz.

Table 4.10: Summary table of the different values of $\sin^2(2\theta_{13})$ measured in DC-III and DC-IV (assuming a Normal Hierarchy).

Analysis	$\sin^2(2\theta_{13})$ value
DC-III n-Gd R+S	$0.090^{+0.032}_{-0.029}$
DC-III n-Gd RRM	$0.090^{+0.034}_{-0.035}$
DC-III n-H R+S	$0.124^{+0.030}_{-0.039}$
DC-III n-H RRM	$0.095^{+0.038}_{-0.039}$
DC-III combined RRM	0.088 ± 0.033
DC-IV n-Gd R+S	0.111 ± 0.018

Double Chooz strived to increase the quality of the analysis in order to increase the sensitivity to θ_{13} by presenting new analyses methods (Rate+Shape, n-H and RRM). All methods developed are in agreement, and the different estimations of our backgrounds allow to achieve the best relative uncertainties on our background model.

The proposal for the experiment suggested a sensitivity for $\sin^2(2\theta_{13})$ from 0.2 to 0.03 with two detectors [6]. The latest result for the single-detector phase reached the proposal sensitivity with an uncertainty of 0.033, but with a single detector. Using both detector decreases the systematic uncertainties even further to 0.018.

Chapter 5

Fast Neutron Background measurement : DCTPC

As explained in Section 4.2.b, a major source of background for low background experiments such as neutrino oscillations or dark-matter detection consists in fast neutrons, mostly produced by atmospheric muons interacting in surrounding materials, or by (α, n) decays from natural radioactivity of the ambient rock. Double Chooz Time Projection Chamber (DCTPC), a portable directional neutron detector, has been designed to provide a measurement of the neutron flux, spectra and direction in both Double Chooz detector laboratories [91].

5.1 Production of Fast Neutrons

Neutrons are generated underground by two sources; the first through the interactions of muons and the second from natural radioactivity. At depths greater than a few tens of meters, radioactive processes are the dominant source of neutron production. The neutrons produced via radioactivity are low energy neutrons, typically less than 10 MeV and therefore can be well shielded against. The neutron spectrum from muon spallation, however, is harder and the resulting neutrons have energies than can extend up to several GeV. The muon-induced component is far harder to shield, since there is an equilibrium with the muon and hadron flux in the rock, the shielding material can produce more neutrons than it removes.

5.1.a Muon induced neutrons

The mechanisms through which cosmic ray muons produce neutrons are:

Negative muon capture on nuclei: a low energy negative muon can be attracted by the Coulomb field of a nucleus forming a 'muonic atom' bound state. This bound muon quickly cascades to the 1s state where it can either undergo decay or nuclear capture.

$$\mu^- + \mathrm{A(Z,N)} \rightarrow \nu_\mu + A(Z-1, N+1)$$

The number of neutrons from muon capture is a combination of the flux of stopping μ^- at the given depth, the capture probability, which depends on the target material and is proportional to Z^4_{target}, and the neutron multiplicity following capture. The muon-induced neutron rate is more important for shallow depth laboratories, such as Double

Chooz laboratories, with 150 and 400 m.w.e., as the muon energy spectrum is softer and thus the rate of stopping muon is higher.

Electromagnetic showers: as a muon passes through matter, photons and electrons from bremsstrahlung and pair production are emitted and produce particle showers. Typically this contribution is small for rock, but the cross-sections increase with Z^2 and become important for high-Z targets such as lead, often used as detector shielding, producing neutrons from the detector shielding.

Muon spallation: muon spallation via virtual photon exchange, as illustrated in Figure 5.1. This process is one of the least understood processes for high energy muons.

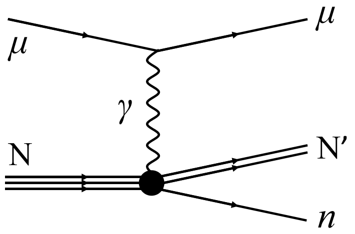


Figure 5.1: Diagram of the muon spallation process, in which the muon knocks a neutron out of a nucleus by exchange of a virtual photon.

Muon-nucleon quasi-elastic scattering: neutrons are knocked-out of the nucleus by high energy muons.

Secondary neutron production from previous processes: primary particles produced in the above processes also interact with the dense media such as rock, shielding and detector materials to produce secondary neutrons. Such processes quickly dominate at large depths with contributions becoming as high as 75% at deep sites, i.e. with a hard muon spectrum.

Of interest to us are measurements made at shallow sites; Palo Verde, Stanford Underground Facility and Russian facilities. Measurements are usually performed using Gadolinium-loaded liquid scintillators. Theoretical models show that muon-induced neutron rate should grow as $\sim E_\mu^{0.75}$ [92], which is compatible with the behaviour observed in Figure 5.2, from [93], that shows the evolution of the total muon-induced neutron yield in liquid scintillator as a function of the mean muon energy for sites at various depths, thus with harder spectra as their depth increases.

However, although the neutron production from muons is well understood in liquid scintillator, the neutron background in Double Chooz is not produced in the detector itself. Indeed, the muon tagging efficiency is very high, and a veto of 1 ms is applied upon detection of a muon. The neutron background is actually produced outside the detector, in the surrounding rock, either induced by muons or from natural radioactivity (α, n) reactions. This rock-produced neutron flux is more difficult to compare from one site to another because of high variability in rock compositions.

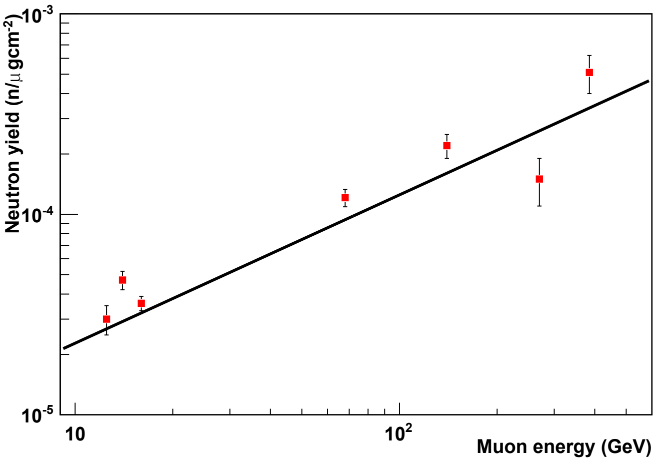


Figure 5.2: The total muon-induced neutron yield in liquid scintillator as a function of the mean muon energy. The line shows the fitting result of the FLUKA-1999 simulation study [94]. The data points, from left to right, are from respectively, the Stanford Underground Facility [95] (20 m.w.e.), a gypsum mine [96] (25 m.w.e), the Palo Verde experiment [97] (32 m.w.e.), a salt mine [96] (316 m.w.e), the Artemovsk Scientific Station [98] (750 m.w.e.), the Large Volume Detector (LVD) in Gran Sasso [94] (3650 m.w.e.), and the Large Scintillation Detector (LSD) in Mont Blanc [99] (5200 m.w.e.).

5.1.b Natural Radioactivity

In addition to muon induced production, neutrons can be emitted by (α,n) reactions with low Z materials and also via spontaneous fission. Although muon-induced production is dominant at shallow sites, the muon flux quickly decreases with depth, and the emission from radioactivity becomes predominant, as shown in Figure 5.3.

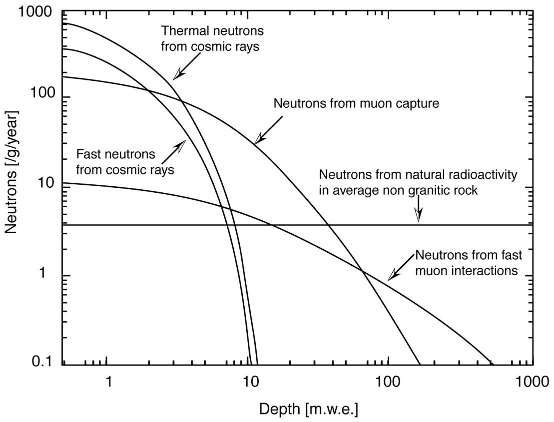


Figure 5.3: Relative importance of neutron production processes as a function of the depth [100].

Direct fission neutrons from uranium and Thorium (dominantly ^{238}U) are sub-dominant to the alpha-n process.

Alpha-emitting daughters from the trace isotopes of ^{238}U,^{235}U and ^{232}Th in the laboratory surrounding rocks can produce neutrons through the (α,n) reaction. For the ^{232}Th chain, the important alpha emitters are ^{216}Po and ^{212}Po, and for ^{238}U, they are ^{218}Po, ^{214}Po, and ^{210}Po.

Neutrons arise from the interaction of alphas on less abundant elements such as aluminium or sodium. The reaction is exoergic on light elements such as ^{9}Be, ^{13}C, ^{17}O, ^{25}Mg, or ^{43}Ca, present in the surrounding rock.

Modane and LNGS laboratories observe a 1.6 to $2.6 \cdot 10^{-6}$ n cm^{-2} s^{-1} fast neutron flux that corresponds to 2.6 ppm of ^{238}U, and 0.74 ppm of ^{232}Th in Gran Sasso rock, and 1.05 ppm of ^{238}U and 0.67 ppm ^{232}Th in Gran Sasso concrete walls [101].

Figure 5.4 shows a prediction of the neutrino energy spectrum in Gran Sasso, assuming a standard rock composition, and with the measured local rock composition. The two predictions show significant differences, hinting the importance of rock composition and thus, the difficulty to compare rock-produced neutrons from one experimental site to another.

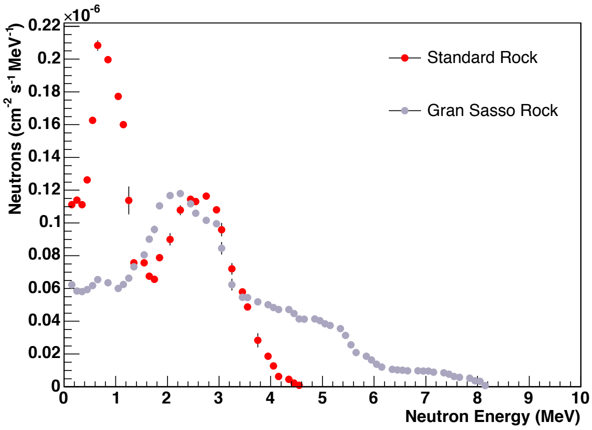
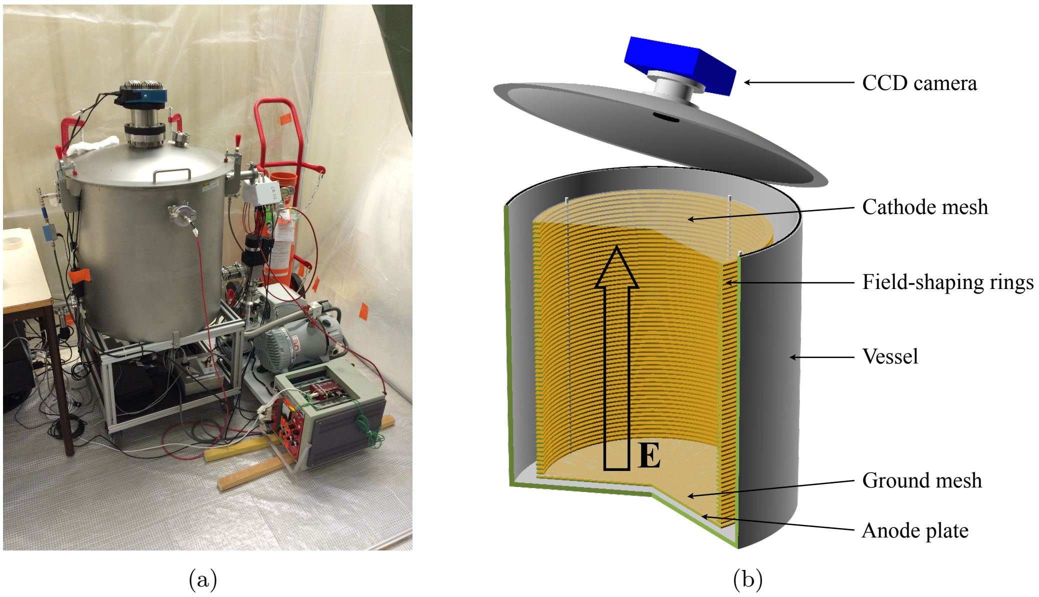


Figure 5.4: Neutron flux predictions by Mei et al. [101] for (α,n) neutron emission for different assumptions on the rock composition.

5.2 Presentation of DCTPC

Double Chooz, and other θ_{13} experiments RENO and Daya Bay, are located at intermediate depths of hundreds of meters of water, where muon-induced and radioactivity-induced neutrons compete, depending on depth and rock composition. However, neutrons that constitute the background in such experiments are mostly of high energy, in order to be able to cross the shielding and the different volumes and interact in the Target, and therefore created by muons in the surrounding rocks or in the shielding itself.

DCTPC, presented in Figure 5.5, is a gaseous TPC, which was installed and operated for 6 months in each of the Double Chooz underground laboratories. It is a larger version of a previous first prototype [91] which was tested at MIT and ran for a period in the DC Far



(a) (b)

Figure 5.5: 5.5(a): DCTPC installed in the far detector laboratory.
5.5(b): Sectional view of the organization principle of the TPC.

Laboratory. Valuable experience was gained running this prototype, leading to the design of DCTPC. Many improvements were made, in particular the size of the chamber was increased. The slow control was also rewritten to allow a better operation at distance.

The active detection volume is composed of a 60 L time projection chamber (TPC) filled with a mixture of ^{4}He and CF_4 (respectively 87.5% and 12.5%), at a pressure of 600 Torr (0.8bar). The ^{4}He nuclei act as targets for fast neutron elastic scattering. The recoiling ^{4}He nuclei in turn ionize the CF_4, leaving a primary ionization track. Inside the chamber, a drift field of $160\,\mathrm{V\,cm^{-1}}$ is applied between the cathode mesh at the top and the ground mesh at the bottom (Fig 5.6). The field is kept uniform throughout the volume by a series of field shaping copper rings coupled together and to the meshes by identical resistors.

An amplification region is created between the ground mesh and the anode plate, separated by insulating cylinders, called spacers, of 440 µm diameter and placed 2.41 cm apart. A field of $1.75\,\mathrm{kV\,cm^{-1}}$ amplifies the primary ionization by a factor of approximately 10^5. Electrons from primary ionizations drift down towards the ground mesh, and an amplification avalanche takes place between the ground mesh and the anode plane. The light from the avalanche of the projected ionization track on the amplification plane is imaged by a CCD camera at the top of the detector. The charge of the primary ionization is read-out on the ground mesh by a fast amplifier, and that of the amplification avalanche is read-out by a charge-sensitive preamplifier.

5.2.a CCD Readout

The CCD camera continuously takes images of the amplification plane (Fig. 5.7), with integration times of one second. The camera used is the Apogee Alta F6 low noise camera

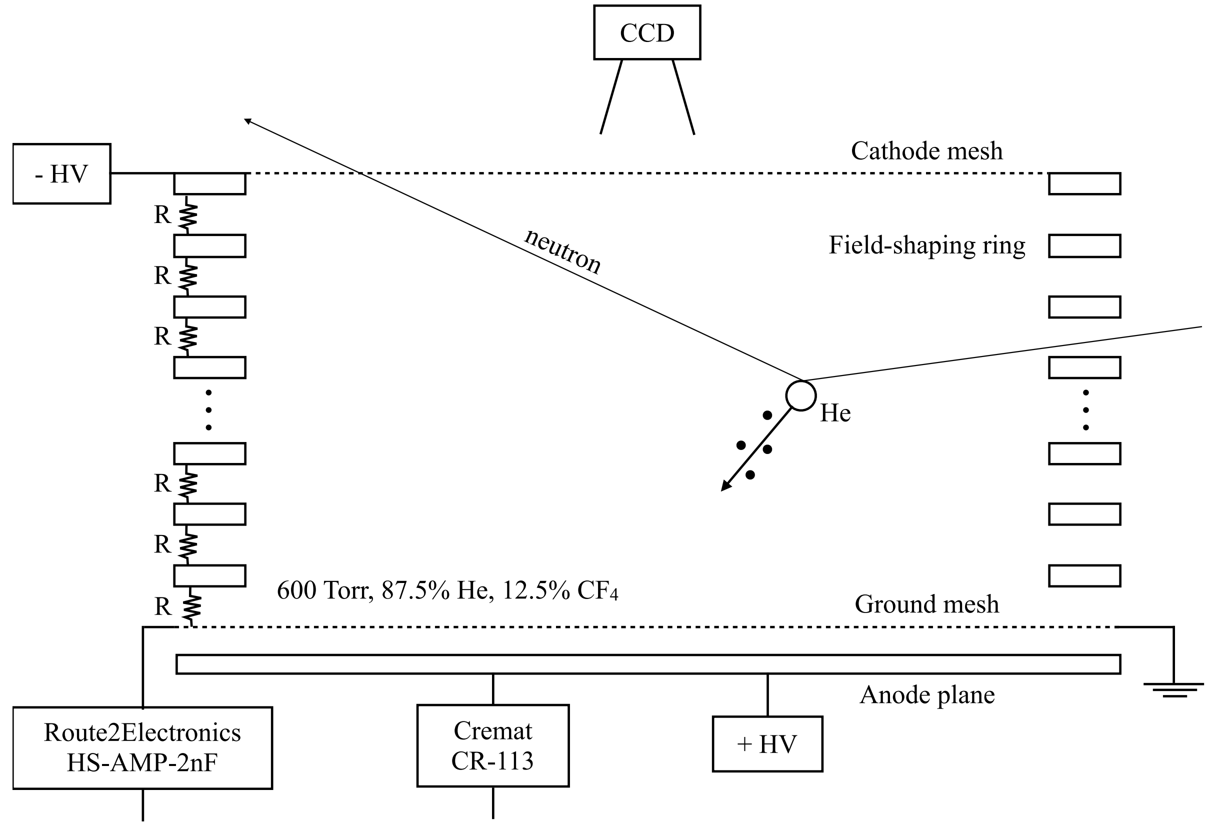


Figure 5.6: A schematic of the detector: the drift field is created by a cathode mesh, field-shaping rings attached to a resistor chain, and a ground mesh. Primary ionization from a recoiling nucleus drifts down to the ground mesh. The high-field amplification region is formed by the ground mesh and the anode plane. The ground mesh is read out with a fast amplifier and anode are readout with a charge-sensitive preamplifier. Scintillation light from the amplification region is recorded with the CCD camera.

with a hydroelectrically cooled grade 2 Kodak KAF 1001E CCD chip of 1024 × 1024 pixels (24 µm × 24 µm). In order to minimize dark current, we keep the CCD chip at a temperature of −20°C. A Nikon NIKKOR 50mm F1.2 lens is used with the camera.
A sample of 100 bias frames are taken at the beginning of each run, with the shutter closed. The mean bias frame is then subtracted from the image read-out during data taking in order to prevent impact from pixels with unusually high values. To increase the signal-to-noise ratio, the image is rebinned by grouping 4 by 4 pixels. Clusters of more than 2 pixels with a weight of more than 3.7σ above the mean value of the image are considered to be tracks from nuclear recoils.

The CCD image provides the position of the projected track on the anode plane, as well as its horizontal angle ϕ (see Figure 5.7). Moreover, the amount of light collected is proportional to the charge deposited on the anode plate, and therefore to the energy of the recoil.

Clearly visible on the image of the track is the Bragg peak, the increase in light at the end

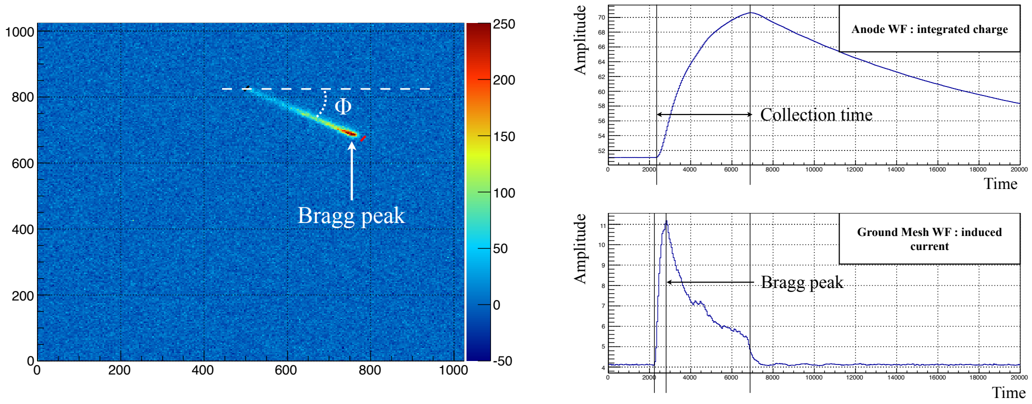


Figure 5.7: Event read-out : CCD image of a nuclear recoil track and the corresponding waveforms (top : anode readout, bottom : ground mesh readout).

of the track due to the increase in dE/dX of the recoiling nucleus as it loses energy. Identifying the Bragg peak of a track gives the direction of recoil.

5.2.b Waveform Readout

The current induced in the ground mesh and the charge deposited on the anode plate are amplified by fast pre-amplifiers (respectively Route2Electronics HS_AMP_2nF and Cremat CR112) and digitized with a Flash-ADC CAEN $N6720B$ (Fig. 5.7).

The waveforms are coded over 12 bits, with a sampling rate of 250 MS/s. The acquisition is done in auto-trigger mode. There is no external trigger system, the Flash-ADC records a waveform when the anode waveform crosses a fixed threshold of 40 mV.

The total deposited charge can be obtained in either of two ways, by integrating either the full ground mesh waveform or the anode waveform up to its maximum. The fall time of the anode waveform is dominated by the pre-amplifier's discharge and is of no interest for our study.

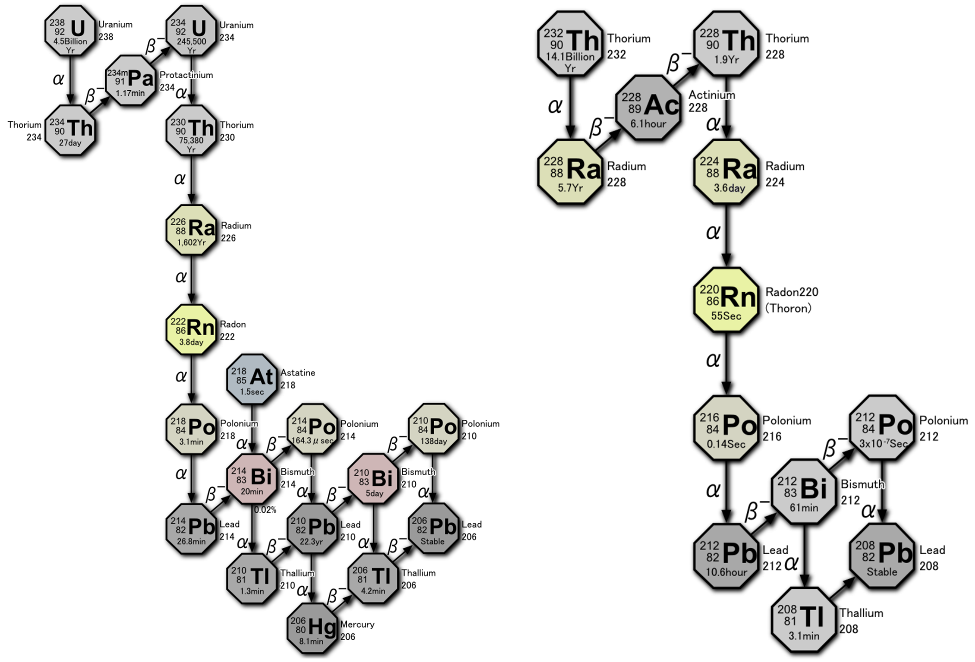
The mesh rise time is defined as the necessary time for the mesh waveform to go from the baseline to the maximum amplitude of the waveform. The fall time is the necessary time for the waveform to fall from the maximum to 10% of the maximal amplitude. The sum of the mesh rise and fall times gives the collection time, the time necessary for all the charges to be collected on the amplification plane.

5.3 Backgrounds in DCTPC

DCTPC is a low pressure gaseous detector, with a pressure of 0.8 bar. The low density of the detecting volume is one of the strengths of this detector as it is insensitive to Minimum Ionizing Particles (MIP) such as atmospheric muons, ambient radioactivity γ or β. These interacting particles have low dE/dx and so will deposit energies below the sensitivity threshold of the detector. For low energy depositions, the signal-to-noise ratio of the CCD renders track reconstruction impossible. Nuclear recoil however, have a high dE/dX and those depositing more energy than 0.5 MeV are well detectable.

DCTPC is optimized to detect recoils of helium nuclei, thus any α emitter is a source of background for the experiment. α of unknown energies can be emitted by decays within the materials of the TPC, such as its stainless steel vessel, or the field-shaping copper rings. These sources are discarded by a simple requirement that the event's track has to start within the volume of the TPC.

The materials, however, are naturally contaminated with traces of ^{238}U and ^{232}Th. Among the daughter nuclei of these isotopes are ^{222}Rn and ^{220}Rn (Figure 5.8). Those two isotopes are gaseous at ambient conditions. They can out-gas into the detection volume and decay inside the sensitive region. DCTPC is then sensitive to all α decays in the rest of the chain, summarized in Table 5.1.



(a) (b)

Figure 5.8: ^{238}U and ^{232}Th decay chains. The ^{238}U and ^{232}Th naturally present in the steel vessel decay into ^{222}Rn and ^{220}Rn. Those are gaseous at room temperature and pressure, and can out-gas into the TPC volume. DCTPC is sensitive to α decays of all the daughter isotopes of Radon in the rest of the decay chains.

Additional contaminations may arise from opening the TPC for maintenance and letting in some air, containing natural levels of Radon. To reduce this contamination, the vessel is pumped down to $\sim 10^{-6}$bar during 24 hours, to remove Radon as much as possible. During data taking at the Far Laboratory, the TPC operated very steadily and was kept closed, suppressing this contamination source. At the Near Laboratory however, the TPC required maintenance and so

Table 5.1: Major α emitter isotopes in the Radon decay chains.

	Isotope	Energy	Life-time
^{238}U	^{222}Rn	5.490 MeV	3.8 d
	^{218}Po	6.002 MeV	3.1 min
	^{214}Po	7.687 MeV	164.3 µs
	^{210}Po	5.305 MeV	138 d
^{232}Th	^{220}Rn	6.288 MeV	55 s
	^{216}Po	6.778 MeV	0.14 s
	^{212}Bi	6.089 MeV	61 min
	^{212}Po	8.785 MeV	0.3 µs

had to be opened.

α from Radon and daughter isotopes decays are indistinguishable from genuine neutron-induced nuclear recoils, and happen randomly in the detection volume. Their energy ranges from 5 MeV to 10 MeV and lowers our neutron sensitivity in this region. The contamination below 4.5 MeV however is negligible, allowing a background-less neutron measurement. Future developments of the TPC should address the Radon problem, by changing the target nuclei from Helium to Neon (see Section 5.7) or by allowing for filtering-out of the Radon through active charcoals.

Since the energies of each α line is well known, they constitute a powerful calibration source, and a tool to verify the homogeneity and stability of the response of the TPC.

Radon isotopes ^{222}Rn and ^{220}Rn and their successive daughter nuclei are short lived compared to their respective parent nuclei (respectively 1602 years for ^{226}Ra and 3.6 days for ^{224}Ra). An equilibrium is thus established between Radon production and outgassing (slow process), and the radon and daughter nuclei decays (fast process). The rate of alphas in the TPC volume is therefore not expected to vary with time.

5.4 Data Analysis

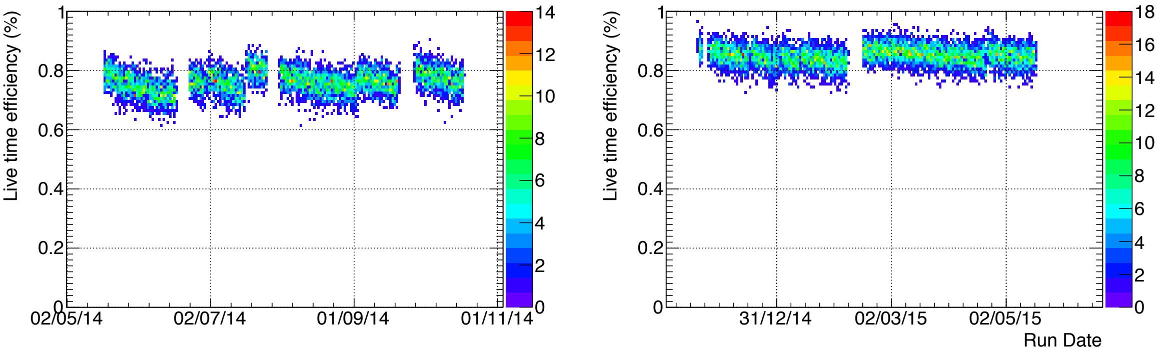
Data taking periods are separated in sequences, corresponding to a running period between gas renewal. When renewing the gas, the old gas is pumped out thanks to a conventional and a turbo pump, bringing the pressure down to 1×10^{-7} bar, and filling with new gas up to 0.8 bar with a tolerance of 8×10^{-3} bar.

5.4.a Spark Removal

Sparks can happen between the ground mesh and the anode plate. An event is identified as a spark if the weight of the higher pixel in an identified track is greater than 1000 ADU (Analog to Digital Unit), or if one of the tracks in the image is comprised of more than 1500 pixels. The six events immediately following a spark are removed in order to allow the amplification voltage to revert to its nominal value.

The six first events of each run are rejected as well to account for the possibility of a spark occurring just before the beginning of data taking. This spark veto introduces a dead time of $(23 \pm 5)\%$ when running in the Near Laboratory, and $(15 \pm 3)\%$ when running in the Far Laboratory. Figure 5.9 shows the live-time efficiency (in %) as a function of the run date. The decreasing pattern is due to the increasing spark rate, due to impurities in the gas depositing in the amplification region. The spark rate increased faster when running in the Near Laboratory; this was due to the ground mesh plastically bending towards the anode plate, thus effectively increasing the amplification and the spark rate. After the mesh had stabilised, the spark rate increased much more slowly. Steps in the spark rate evolution correspond to a new sequence, where the gas is pumped out and changed, and the amplification field is turned off.

To increase stability of the data taking, the size of the drift chamber was reduced by $\sim 25\%$ when moving to the Far Laboratory. This allowed to bring the amplification plane closer to the CCD camera, thus improving the light collection, and allowing to lower the amplification field, therefore reducing the spark rate. Because the spark rate was more stable, sequences were allowed to be longer in the Far Laboratory, increasing the data taking stability.



(a)

(b)

Figure 5.9: Data taking efficiency as a function of the data taking time for the TPC in the Near (5.9(a)) and the Far (5.9(b)) laboratories. Loss of efficiency over a same sequence is dominated by the increasing spark rate.

5.4.b Radial Selection

The copper of the field-shaping rings is expected to be an intense source of α. All the events starting close to the rings are rejected. Due to the high rate of α from the rings compared to the 1 s exposure of the CCD, it is very likely that two α can overlap. The reconstruction algorithm can fail to identify the two tracks, falsely reporting one track with a fake Bragg peak closer to the actual origin of the tracks. In this case, the track origin is wrongly reconstructed as starting well inside the field cage, and ending near the rings. To prevent this behaviour, all events coming from or to the rings are discarded.

Figure 5.10 shows that the ratio of energies reconstructed using the ground mesh signal and the light collected by the CCD decreases with the radius. The amplification field depends on the distance between the ground mesh and the anode plane. This distance is kept constant by parallel spacers, so an inhomogeneity in the amplification field due to that distance would not be a radial effect, and the same light yield is expected throughout the plane. The effect is mostly due to the edges of the plane being further away from the camera lens than the center, and the light collection is less efficient as the distance to the center increases. As a consequence, the tracks appear brighter at the center, and low energy tracks have a better chance of being reconstructed in the center than in the edges.

The tracks projected on the amplification plane need to be within a distance ρ_{max} from the center, with $\rho_{max} = 10.3\,\text{cm}$ for the Near Laboratory and 8.8 cm for the Far Laboratory, corresponding to 300 and 400 pixels respectively, in order to limit the light collection efficiency loss to less than 25%.

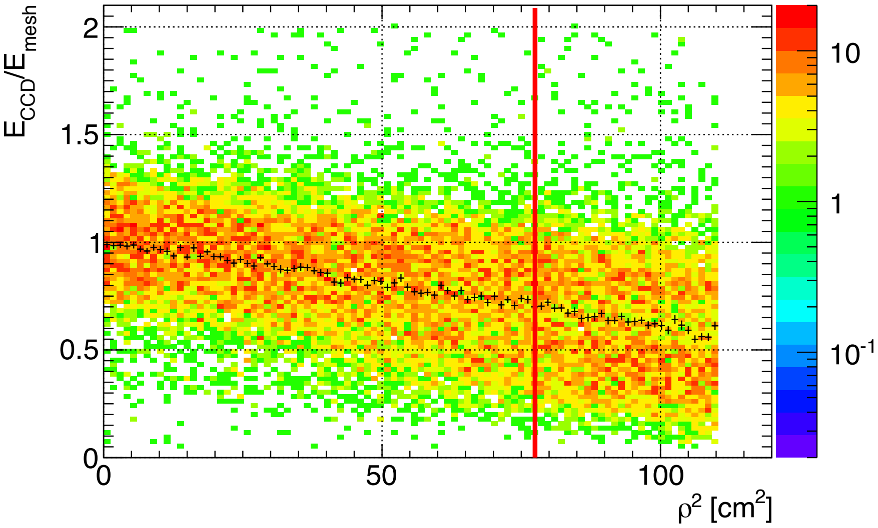


Figure 5.10: Evolution of the ratio of the energies reconstructed using the ground mesh signal (E_{mesh}) and the CCD collected light (E_{CCD}) as a function of the squared radius of the track position for DCTPC in the Far Laboratory. The mesh energy is constant over the radius, only the light collection changes. The limit is set at 8.8 cm (red line) so that the light collection efficiency does not decrease by more than 25%, to ensure a good reconstruction efficiency at low energies.

5.4.c 3D Length Reconstruction

As explained in Section 5.2, an event is

- a read-out of the current induced in the ground mesh when the primary ionization track drifts to the amplification plane,

- a read-out of the charge collected on the anode plane after amplification,
- a read-out of the light collected by the CCD during a 1 s exposure time.

The sketch on Figure 5.11 defines the variables used thereafter. The CCD provides spatial information on the projected track ($\mathrm{L}_{\mathrm{horizontal}}$). The full 3D track length ($\mathrm{L}_{\mathrm{track}}$) can be obtained by reconstructing the vertical information thanks to the ground mesh signal: if the drift speed ($\vec{V}_{\mathrm{drift}}$) of the primary ionization track is constant, the vertical length ($\mathrm{L}_{\mathrm{vertical}}$) can be computed by measuring the time it took for the ionization track to be fully deposited on the amplification plane, called later on collection time ($t_{\mathrm{collection}}$).

$$\mathrm{L}_{\mathrm{vertical}} = t_{\mathrm{collection}} \cdot \vec{V}_{\mathrm{drift}} \tag{5.1}$$

$$\mathrm{L}_{\mathrm{track}} = \sqrt{\mathrm{L}^2_{\mathrm{horizontal}} + \mathrm{L}^2_{\mathrm{vertical}}} \tag{5.2}$$

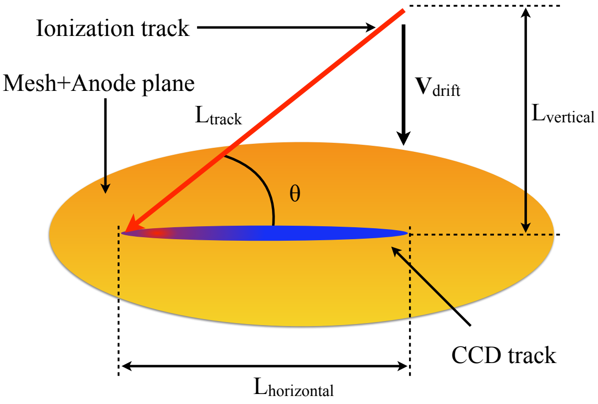


Figure 5.11: Track length reconstruction : The horizontal length is obtained by the CCD image, the vertical length is obtained by the collection track time if the drift velocity is known.

The horizontal length is calibrated by summing all CCD images, thus lighting the full anode plane. The spacers are insulating, and no light is provided at their position, thus they show-up as parallel lines of lower gain. The distance between the spacers is 2.41 ± 0.01 cm. We can then access the corresponding size of a pixel on the anode plane: 0.3447 ± 0.0007 mm/pixel for the Near Laboratory and 0.2203 ± 0.0006 mm/pixel for the Far Laboratory.

The collection time is the sum of the rise time and the fall time of the ground mesh waveform. Some background events originate from the mesh itself, or the anode plate. Those events have an anomalous waveform with very short rise and fall times, dominated by the behaviour of the pre-amplifcators. To reject those events, it is requested that the collection time is higher than a time offset. Figure 5.12 shows the distribution of the collection times for DCTPC in the Far Laboratory. The dotted line corresponds to the minimum value required. The offset is set at 1.726 µs for the Far Laboratory, and 1.826 µs for the Near Laboratory.

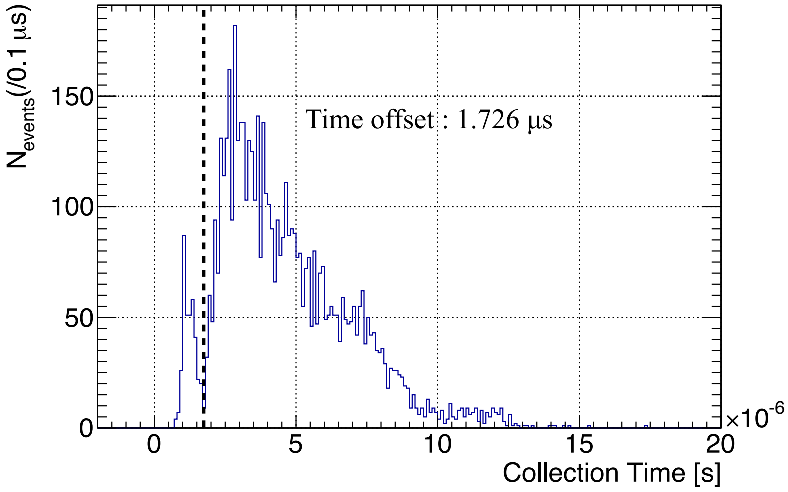


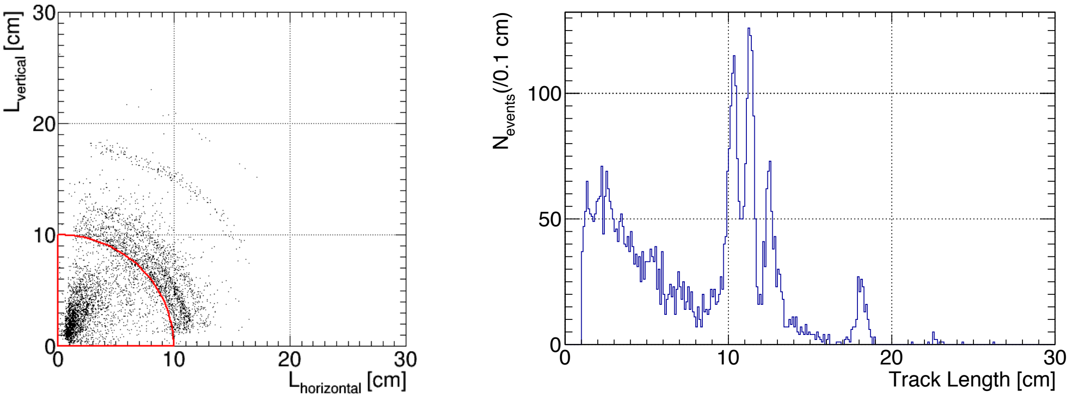
Figure 5.12: Collection time of the mesh waveform for DCTPC running in the Far Laboratory. Events below 1.725 µs are dominated by the rise and fall times of the pre-amplifiers. Their collection time does not have physical meaning. The collection time is shifted by a 1.725 µs offset.

The drift speed is unknown and has to be measured with the data. For a given value of the drift speed, a certain value of the vertical length is obtained. If the events are happening isotropically in the volume, for a given $\mathrm{L_{track}}$, any set of $\mathrm{L_{vertical}}$ and $\mathrm{L_{horizontal}}$ that satisfies Eq. 5.2 can happen. If $\vec{V}_{\mathrm{drift}}$ is correctly set, the events with the same $\mathrm{L_{track}}$ (i.e. with the same energy) will form a circular arc in the ($\mathrm{L_{vertical}}$, $\mathrm{L_{horizontal}}$) plane with radius $\mathrm{L_{track}}$. Figure 5.13(a) shows the distribution of horizontal and vertical lengths in the Far Laboratory. Circular arcs are visible, corresponding to mono-energetic α emission from the ^{222}Rn and ^{220}Rn decay chains. Figure 5.13(b) shows the reconstructed $\mathrm{L_{track}}$ distribution, corresponding to the distance of the points in Figure 5.13(a) to the origin. The clear peaks correspond to α emitted by different isotopes in the Radon decay chains.

If the assumed drift speed differs from the true drift speed, the circular arcs are deformed, and the α peaks in Figure 5.13(b) will be wider. A Gaussian fit is performed for multiple drift speeds on the peak located around 18 cm, corresponding to the ^{214}Po emission. The value of the width of the Gaussians for each speed is shown in Figure 5.14. The optimal drift speed, the one that minimizes the width of the peak, is found to be $V_{\mathrm{drift}} = 1.682 \pm 0.003\mathrm{cm\,\mu s^{-1}}$ for the Far Laboratory, and $V_{\mathrm{drift}} = 1.347 \pm 0.004\mathrm{cm\,\mu s^{-1}}$ for the Near Laboratory.

5.4.d Energy Reconstruction

An event energy can be reconstructed using the TPC's read-outs, with the current induced in the ground mesh when the primary ionization track drifts towards the amplification plane (E_{mesh}), the charge collected on the anode (E_{anode}), or the light emitted by the avalanche



(a)

(b)

Figure 5.13: 5.13(a): Event distribution in the (L_{vertical}, $L_{\mathrm{horizontal}}$) plane. Events of same energies form a circular arc if the drift speed is correctly tuned. Here a circular arc is shown in red to guide the eye.
5.13(b): Track length distribution for DCTPC in the Far Laboratory of Double Chooz.

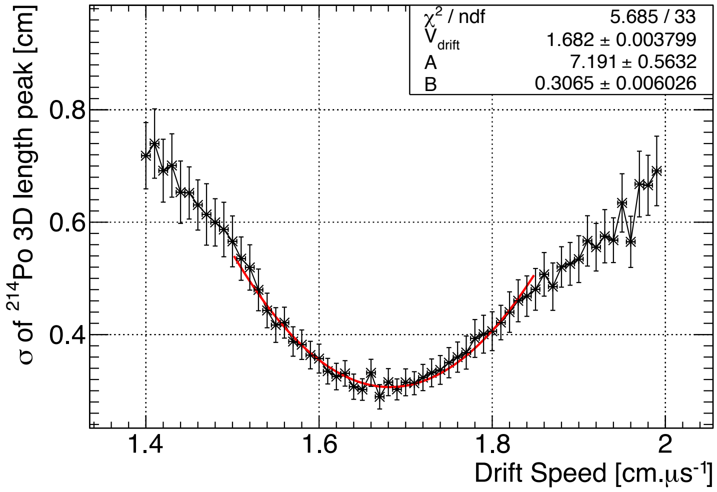


Figure 5.14: Ionization track drift velocity estimation. For each drift speed, the width of the ^{214}Po is measured, the actual drift speed is the one that minimizes the width. It is found to be $1.682 \pm 0.003\mathrm{cm}\,\mathrm{\mu s}^{-1}$.

in the amplification plane (E_{CCD}). Figure 5.15 shows the three spectra. The peak around 6 MeV corresponds to all α emissions from Radon and its daughter nuclei decays, from 5 MeV to 10 MeV.

The length distribution in Figure 5.13(b) shows a better resolution of the α peaks than the energies E_{mesh}, E_{anode}, or E_{CCD}. Figure 5.16(a) shows the correlation between the track length of an event and the corresponding mesh energy E_{mesh}. The red line is a computation of the

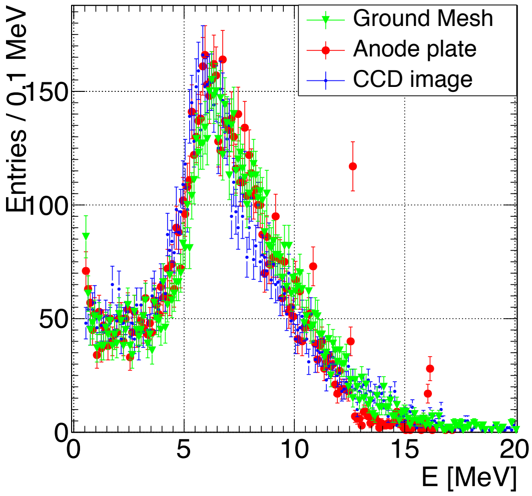
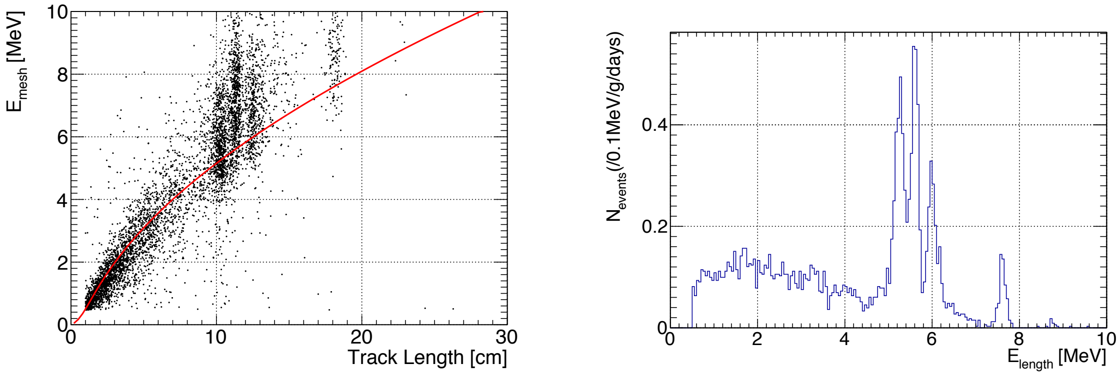


Figure 5.15: Energy reconstructed from the built-in read-outs in the Far Laboratory : the CCD light collection (blue), the anode charge collection (red), and the current induced in the ground mesh (green). The peak around 6 MeV corresponds to the collective α emission from the Radon decay chains.

energy of a recoil for a given length performed using the SRIM software [102]. The low energy neutron induced nuclear recoil seem to follow the SRIM prediction, however the α peaks appear to be leaking at higher E_{mesh}.
Figure 5.16(b) shows the energy of the events evaluated from their track length using the SRIM prediction. There again, the α peaks are well separated, with a relative width of $\sigma/\mu \sim 3\%$.

As the recoils have a higher energy, their tracks get longer, and they are more likely to exit the fiducial volume. If the recoils exit the volume from the sides, then they are rejected by the radial veto, yielding a loss of efficiency, but without biasing the measured energy. If they leave from the top or bottom of the TPC, but their exit point is still within the accepted radius, the events are kept, but only a part of the track is in the TPC and will be collected, leading to a loss in measured energy. For low energy track, this effect is negligible as the boundary effect are small compared to the fiducial volume, but this is no longer the case for tracks of more than 15 cm, i.e for recoils of more than 6 MeV. Similarly, events from above the TPC can enter from the top, this time again with only a fraction of their energy collected.

This behaviour can be simulated by generating events of different energies isotropically in the fiducial volume, and propagating the energy through the reconstruction processes. The result of this simulation for the Far Laboratory configuration is shown in Figure 5.17(a). This simulation also takes into account the resolution in time of the collection time reconstruction, and the spatial resolution of the CCD track reconstruction. It also accounts for the efficiency loss at low energy due to the CCD track reconstruction threshold, and the electron absorption length in the drift cage. Figure 5.17(b), on one hand, shows the expected measured energy distribution for events with true energy of 5 MeV uniformly distributed through-out the detec-



(a)

(b)

Figure 5.16: 5.16(a): E_{mesh} versus track length. The red line is a prediction of the length per energy from SRIM simulations for our compound gas. Neutron events at low energy are in good agreement with the predictions, but alpha lines seem better resolved in length than in energy.
5.16(b): Energy reconstructed from $\mathrm{L}_{\mathrm{track}}$. Each individual α emitter is visible, with a better resolution than the other energy measurements presented in 5.15.

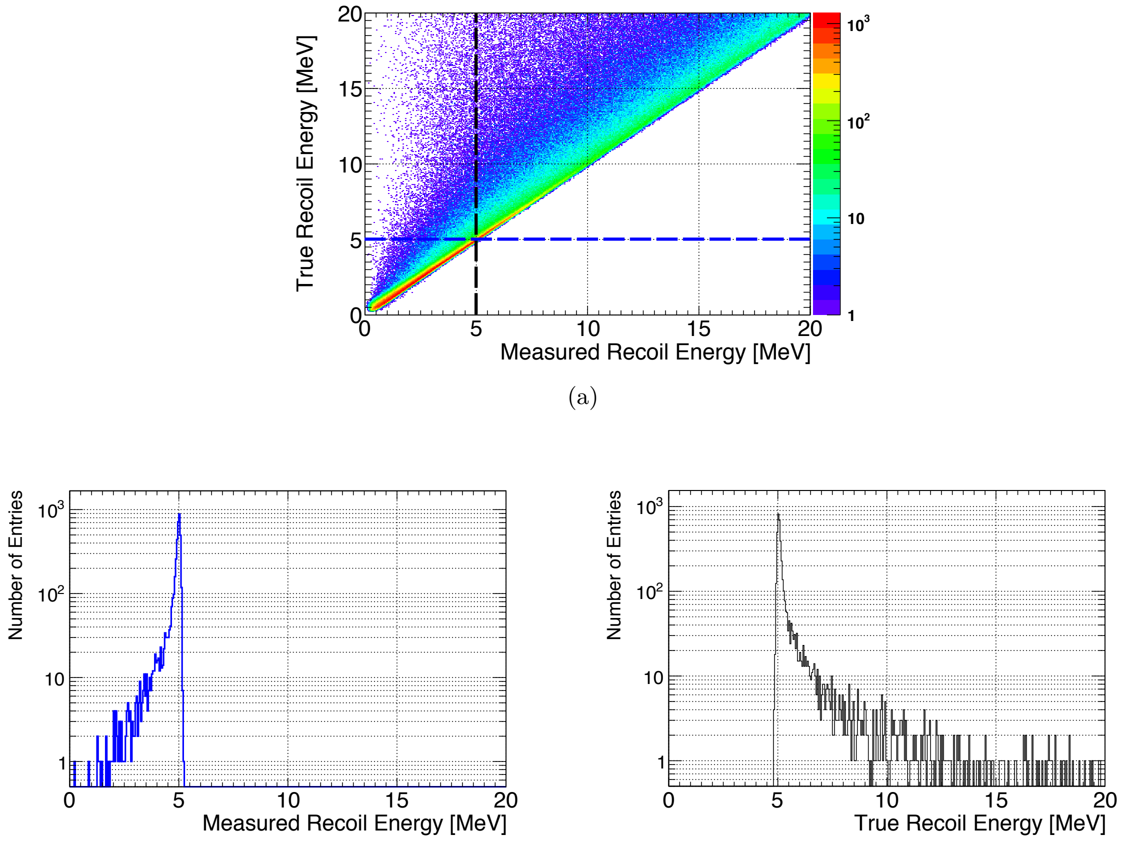
tor. An α emission from a decay in the volume cannot have a measured energy higher than its true energy. However, it can exit the TPC, or lose part of its ionization track due to electron diffusion and be reconstructed at lower energy. This the behaviour is modelled by reading one line of the matrix (blue dotted line). On the other hand, Figure 5.17(c) shows that an event with a given measured energy can have a higher true energy, this behaviour is described by reading one column of the matrix (black dotted line).

It is observed that these behaviours are well described by a Gaussian convoluted to a Landau function in the energy range from 5 MeV to 10 MeV where the α lines are. The Gaussian accounts for the well contained events, spread-out by the energy resolution of the detector. The Landau accounts for the partially contained events, leaking out or in the fiducial volume.

Figure 5.18 shows a fit to the energy spectrum at the Far Laboratory. The Fit function is a sum of a Gaussian to describe the shape of the underlying neutron-induced spectrum, and 8 α peaks, corresponding to the isotopes listed in Table 5.1.

Each α peak is described by a Gaussian convoluted to a Landau. The mean energy of the peaks, and their normalization is left free. All the peaks are characterized using a single parameter σ_{res} for the resolution, and a common parameter σ_{leak} for the non-contained events, to simplify and speed-up the convergence of the fit. The red line corresponds to the best fit of the histogram. The red shaded area corresponds to the 1σ confidence interval induced by the α parameters. The blue line represents the contribution of the α lines to the overall spectrum. The fit outputs are summarized in Table 5.2 for the Near and Far Laboratories.

Figure 5.19(a) and 5.19(b) respectively show the correlation and covariance matrix of the fit performed in Figure 5.18. The clusters of parameters are materialized by lines; the first three



(a)

(b)

(c)

Figure 5.17: 5.17(a): Energy collection matrix.
5.17(b): Expected measured energy for a 5 MeV event - blue dashed horizontal slice from the collection matrix (a).
5.17(c): True recoil energy for events producing a measured recoil energy of 5 MeV (black dashed vertical slice from (a).)

parameters characterize the Gaussian shape attributed to the neutron-induced nuclear recoils, the fourth and fifth parameters are respectively the energy resolution σ_{res} and the leak parameter σ_{leak}, the other couples of parameters correspond to the normalization and the mean energy of each α peak. It is to be noted that the parameters of the nuclear recoil spectrum are mostly independent from the α peaks determination, allowing a clean measurement of the background, and allowing the α peaks to be subtracted without biasing the nuclear recoil spectrum.

5.4.e Energy Linearity

Figure 5.20 shows the fractional difference between each α peak mean energy and the expected value. The non-linearity of DCTPC length-based energy measurement is at the percent level in the 5 MeV to 9 MeV region. Except for ^{222}Rn, all isotopes are reconstructed less than 1% away from their true energy, and the behaviour of the energy reconstruction is flat over the 5 MeV to 9 MeV region.

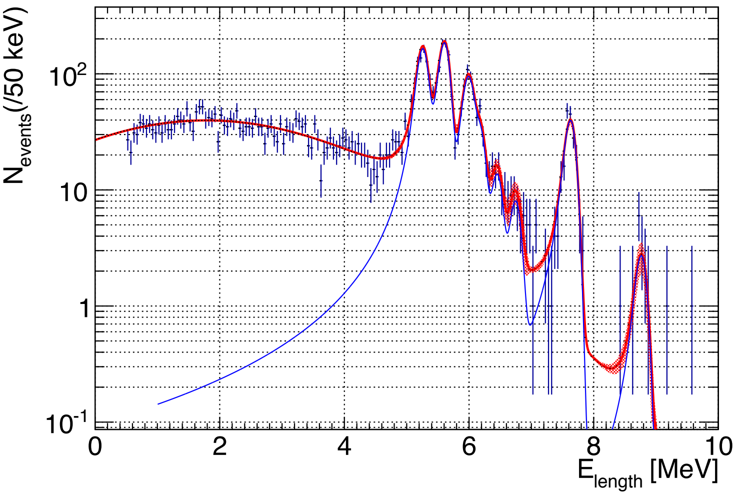
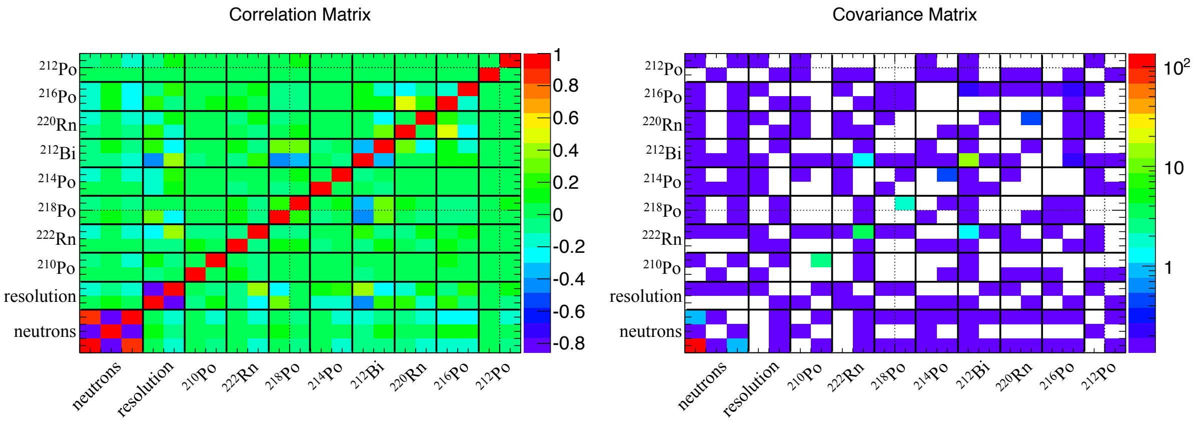


Figure 5.18: E_{length} spectrum with the α lines fitted (red line) for the runs in the Far Laboratory. The red area corresponds to the 1σ confidence interval of the best fit function imputed to the parameters of the α. The blue line is the contribution of the α lines to the overall distribution.



(a)

(b)

Figure 5.19: 5.19(a): Correlation Matrix of the parameters for the fit shown in 5.18.
5.19(b): Covariance Matrix of the fit parameters.

The linearity at low energies, however, cannot be quantified using this method since no low energy source is naturally present in the gas. In order to estimate the sanity of the energy reconstruction, we used an ^{241}Am source. ^{241}Am decays to ^{237}Np by emitting α of 5.5 MeV. However, the isotopes are embedded in a bulk material, and manufacturers usually place a layer of coating for safety. In addition, the source is collimated, so we do not have access to the be-

Table 5.2: Fit output for data sets acquired in the Far and Near Laboratories.

	Parameter	Far Laboratory	Near Laboratory
	N_{n}	199 ± 11	106.43 ± 4.02
	μ_{n} (MeV)	1.77 ± 0.14	3.23 ± 0.13
	σ_n (MeV)	2.01 ± 0.10	2.19 ± 0.08
	Resolution @5 MeV (%)	1.26 ± 0.10	1.41 ± 0.21
	α_{spill} (MeV)	0.024 ± 0.003	0.032 ± 0.004
^{210}Po	N	37.31 ± 1.67	35.61 ± 2.248
	E (MeV)	5.270 ± 0.005	5.292 ± 0.008
^{222}Rn	N	43.95 ± 1.81	88.03 ± 2.86
	E (MeV)	5.618 ± 0.004	5.619 ± 0.003
^{218}Po	N	18.92 ± 1.38	21.00 ± 2.94
	E (MeV)	5.995 ± 0.008	5.9899 ± 0.019
^{214}Po	N	9.69 ± 0.67	20.61 ± 1.23
	E (MeV)	7.641 ± 0.007	7.689 ± 0.006
^{212}Bi	N	3.08 ± 3.84	75.47 ± 10.78
	E (MeV)	6.18 ± 0.015	6.167 ± 0.001
^{220}Rn	N	3.13 ± 0.63	7.46 ± 4.56
	E (MeV)	6.460 ± 0.027	5.990 ± 0.004
^{216}Po	N	1.88 ± 0.46	4.54 ± 0.78
	E (MeV)	6.760 ± 0.043	6.509 ± 0.004
^{212}Po	N	0.66 ± 0.17	3.203 ± 0.47
	E (MeV)	8.777 ± 0.026	8.778 ± 0.001
	χ^2/d.o.f.	178.12/159	176.61/159
	Probability	0.143	0.156

ginning of the tracks. For those reasons, we could not know precisely the energy of the emitted α.

The source was placed on a holding stick, with a collimator, and a pierced 4 µm-thick Nickel foil was placed to cover the collimator's opening as shown in the sketch Figure 5.21(a). With this set-up, a fraction of the α will go through the hole in the foil, not being attenuated from it, while the rest will go through the foil, and therefore will be attenuated. The stopping power of α can be obtained using the SRIM software [102] and is shown in Figure 5.21(b). If we cannot access the true energy of the α, the energy difference between the two configurations is measurable. The energy difference ΔE between the attenuated and non-attenuated α only depends on the incident energy of the α and the thickness of the foil.

Figure 5.21(c) shows the length energy E_{length} measured for the non-attenuated α at 3.75 MeV and the attenuated α at 1.5 MeV. As expected the source does not deliver 5.6 MeV α, a ^{241}Am-only run would have concluded to a misbehaviour of the detection technique. However, a cross-check allowed thanks to the two correlated energies for the attenuated and non-attenuated α determined that the measurement was in agreement with the known width of the Nickel foil of 4 µm.

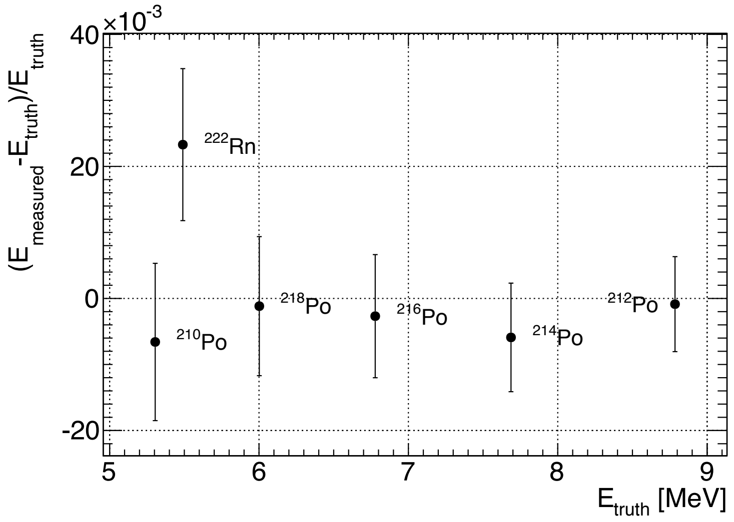


Figure 5.20: Fractional difference between the measured α peak mean energy and the expected value.

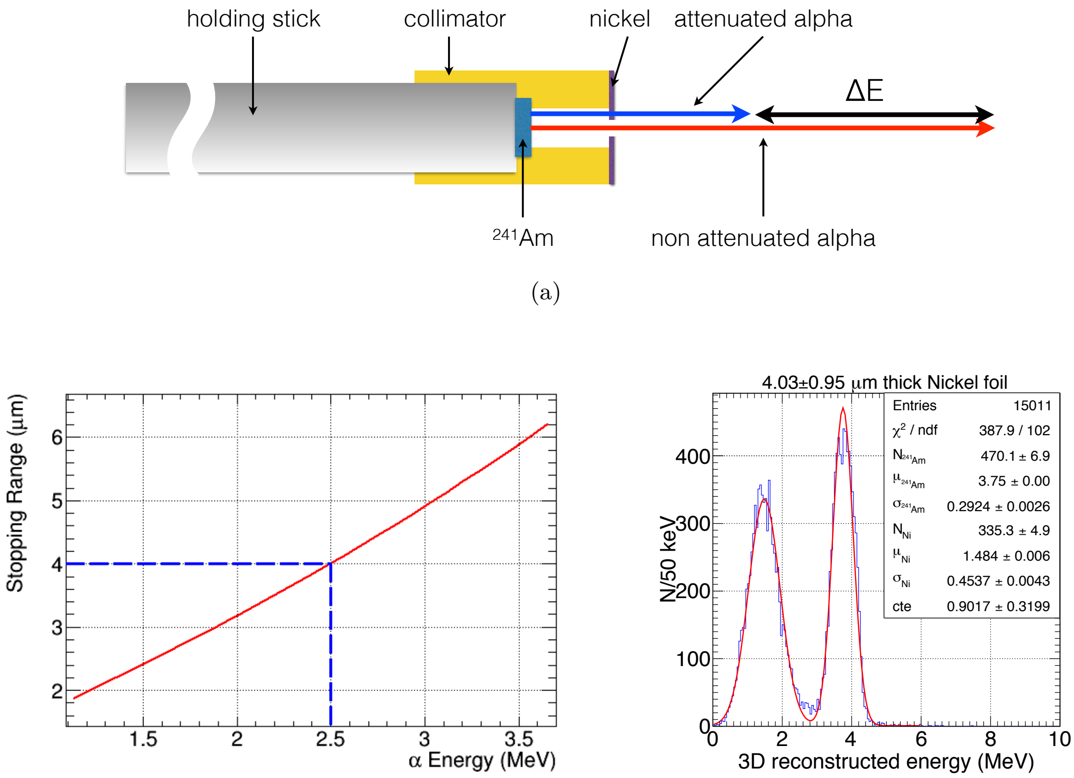
This setup allows to claim that the energy measurement method is valid at high energies thanks to the good linearity of the position of the Rn and Po peaks, as well at low energy thanks to the agreement between the energy difference measured between α of 3.75 MeV and 1.5 MeV and the stopping range of α in Nickel.

5.5 Time Stability of DCTPC

Figure 5.22 shows the evolution of the pressure as a function of the time in the TPC for runs in the Near Laboratory and in the Far Laboratory. The pressure changes between sequences because of tolerance of 1% of the filling procedure, but remains stable within one sequence.

Figure 5.23 shows the evolution of the drift voltage during data taking. The drift voltage, and therefore the drift field, remains very stable throughout data taking. It is to be noted that the drift voltage is reduced for the runs in the Far Laboratory to keep the drift field constant, as it was decided to change the size of the TPC by $\sim 1/4$ to improve light collection and decrease the absorption and diffusion of primary tracks.

Figure 5.24(a) shows the evolution of the energy of the events with $E_{\text{length}} > 4\,\text{MeV}$ for runs in the Far Laboratory. The more populated horizontal lines correspond to the α peaks. Their horizontality shows the stability of energy definition over time, consistent with a constant gas pressure and drift field, and therefore implies the stability of the vertical length definition. The horizontal length definition is not expected to change as the horizontal calibration is done using the image of the anode plane, of fixed dimensions. However, a small effect can be expected at the beginning of the track, where the primary ionization track is the faintest: if electron diffusion and absorption in the gas is too high, the beginning of the track can be too much attenuated, and can become more difficult to reconstruct. This effect however is constant in time as the electron diffusion and absorption depend on the gas pressure, which was found to be very stable.



(a)

(b)

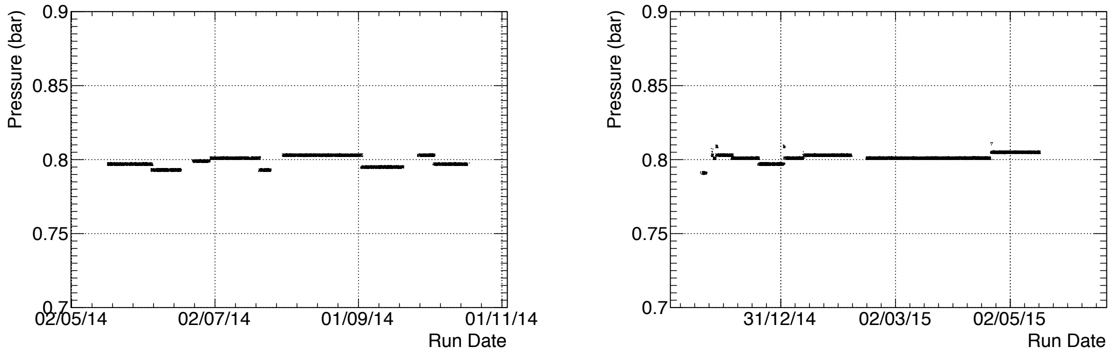
(c)

Figure 5.21: 5.21(a): Setup of the ^{241}Am source with a pierced Ni foil to attenuate some α.
5.21(b): Stopping range α in Ni from SRIM [102]. For a 4 µm-thick foil, an energy decrease of 2.5 MeV is expected.
5.21(c): E_{length} reconstructed from the runs with the ^{241}Am and Ni foil source. The energy difference $\Delta E = 2.27$ MeV is perfectly compatible with an incoming 3.75 MeV α and a 4 µm Ni foil.

Figure 5.24(b) shows the rate of selected events (in black) for runs in the Far Laboratory. The red line corresponds to the events with $E_{\text{length}} > 5$ MeV, dominated by the α emissions, and the blue line corresponds to the events with $E_{\text{length}} < 4$ MeV, dominated by neutron-induced nuclear recoils. The rates of both events appear stable within statistical fluctuations.

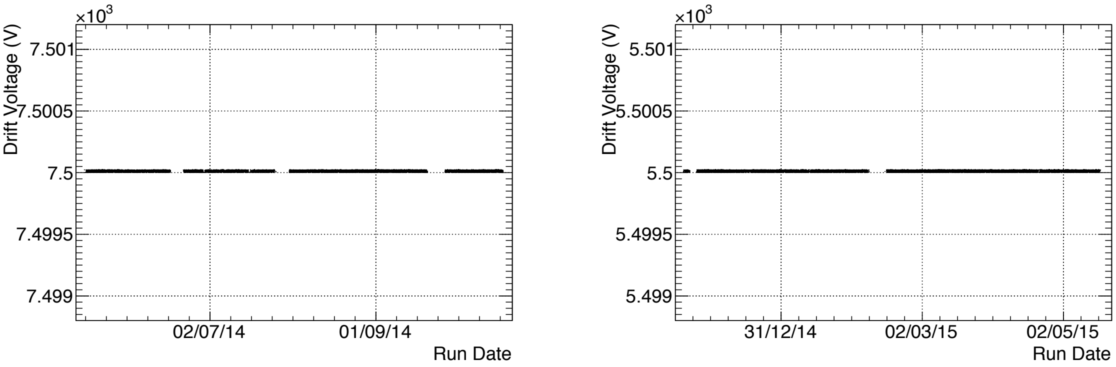
5.6 Event Rates at Double Chooz laboratories

Figure 5.25(a) shows the energy spectrum after the α shape has been subtracted. The error bars for each bin are the sum of squares of the statistical errors of the original measured bin value and of the 1σ confidence interval of the fit due to the α parameters. Large errors are expected in the energy ranges where the fit is poorly constrained due to poor statistics in the α peak, or to an error in the shape assumed to describe the distribution. This leads to a possibility of over-subtracting, and potentially more than 100% of error in certain bins.



(a) (b)

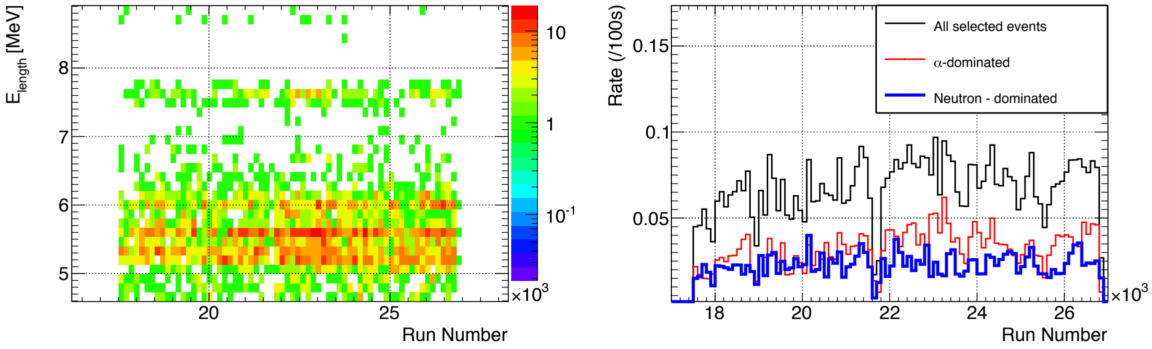
Figure 5.22: Gas pressure in the volume as a function of the data taking time for the TPC in the Near (5.22(a)) and Far (5.22(b)) laboratories. The pressure is very stable in both sites, even for long sequences.



(a) (b)

Figure 5.23: Drift voltage as a function of the data taking time for the TPC in the Near (5.23(a)) and the Far (5.23(b)) laboratories. The drift voltage is kept very stable over time, and was decreased in the Far Laboratory to keep the same drift field when the TPC was scaled down.

According to the discussion around Figure 5.17(a) some of these events originate from higher energy events for which only a part of the energy is recovered. In order to account for that behaviour, we need to unfold this matrix from the spectrum, to account for the probability of each energy bin to leak to higher energies. Let M_{ij} be an element of the energy collection matrix in Figure 5.17(a). For each entry in bin j in the $\mathrm{E_{length}}$ histogram, the true recoil energy histogram $\mathrm{E_{recoil}}$ is filled according to the pdf contained on the column j of the matrix M. Thus each column in the matrix M should be normalized to 1. Let m_{ij} be the normalized weight of the element M_{ij}:



(a) (b)

Figure 5.24: Stability of the reconstructed energy and of the rate as a function of the run number.

$$m_{ij} = \frac{M_{ij}}{\sum_{j=1}^{N} M_{ij}} \tag{5.3}$$

$$\mathrm{E}_{\mathrm{recoil}}[i] = \sum_{j=1}^{N} m_{ij} \mathrm{E}_{\mathrm{length}}[j] \tag{5.4}$$

When decomposing a single bin $\mathrm{E}_{\mathrm{length}}[j]$ over the $\mathrm{E}_{\mathrm{recoil}}[i]$, the error on the integrated $\mathrm{E}_{\mathrm{recoil}}$ should remain the same than that of the bin $\mathrm{E}_{\mathrm{length}}[i]$, thus, the error $\sigma_{\mathrm{recoil}}[i]$ on bin $\mathrm{E}_{\mathrm{recoil}}[i]$ can be expressed:

$$\sigma_{\mathrm{recoil}}[i] = \sum_{j=1}^{N} m_{ij}\, \sigma_{\mathrm{length}}[j]^2 \tag{5.5}$$

Figure 5.25(b) shows the recoil energy spectrum corrected from the α-subtracted length energy spectrum of Figure 5.25(a). Events have been pushed to high energies, with the appearance of a tail at $\mathrm{E}_{\mathrm{recoil}} > 10\,\mathrm{MeV}$. The peaked shape at low energy is more pronounced since low energy events are better constrained, and less likely to come from a higher energy recoil.

The energy containment matrix M only accounts for events that exit the TPC from the top, without being rejected by the radial cut, or the collection time offset condition. However, for the Far Laboratory runs, the fiducial volume is a cylinder of only 8.8 cm of radius and 31.5 cm high. For the Near Laboratory runs, the fiducial volume is a cylinder of 1.34 cm of radius for 52.8 cm high. A significant fraction of high energy recoils exit the fiducial volume from its sides. These events must be taken into account as an acceptance inefficiency.

Moreover, some low energy recoils are not detected, due to electron diffusion during the primary ionization track drift, and due to the low signal/noise ratio of the CCD images at low energies. This must be accounted for as a detection inefficiency.

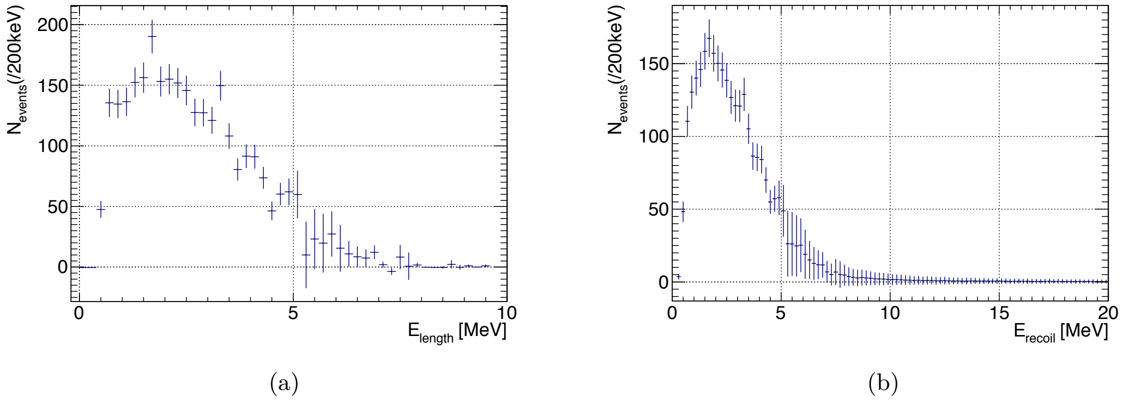


Figure 5.25: 5.25(a): Far Laboratory E_{length} spectrum of recoil nuclei once the alpha lines have been subtracted.
5.25(b): Far Laboratory unfolded spectrum, some events are pushed to higher energies due to the transfer matrix.

Figure 5.26(a) shows the evolution of both acceptance effect at high energies, and detection efficiency effects at low energies. The detection efficiency reaches its maximum around 1 MeV. At this point the acceptance begins to dominate the inefficiency, with a decrease down to 8 MeV, then enters a plateau at high energies.

The decrease in acceptance comes simply from a geometric effect. As the tracks get longer, it is harder to contain them in the fiducial volume. The plateau at high energies corresponds to tracks $\gtrsim 30$ cm. Those tracks cannot be contained in the volume, only tracks that are almost vertical can exit from the top without being rejected, and this fraction of accepted tracks does not depend any more on the track length.

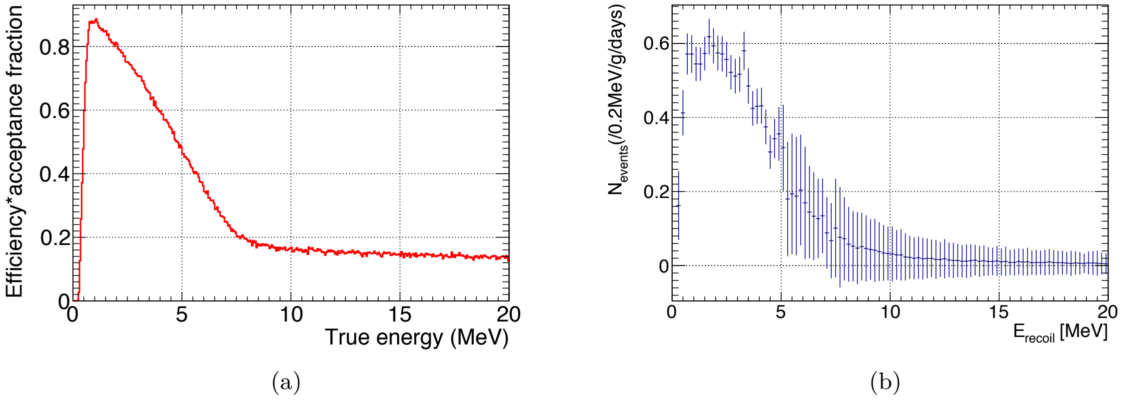


Figure 5.26: Far Laboratory efficiency- and acceptance-corrected spectrum.

Figure 5.26(b) shows the recoil spectrum corrected for those acceptance and efficiency effects. The high energies are amplified with respect to the lower energies since the acceptance is lower for the high energies.

Neutron rates and spectral shapes were acquired in both experimental laboratories of Double Chooz for a total exposition time of 102.11 days in the Near Laboratory, and 89.63 days in the Far Laboratory.

Figure 5.27 compares the raw E_{length} (top) and final corrected E_{recoil} (botton) spectra for Near (left) and Far Laboratories (right).

The spectral shapes for the two different experimental locations are significantly different, be it in recoil spectral shape, or α-emitter isotopes contamination.

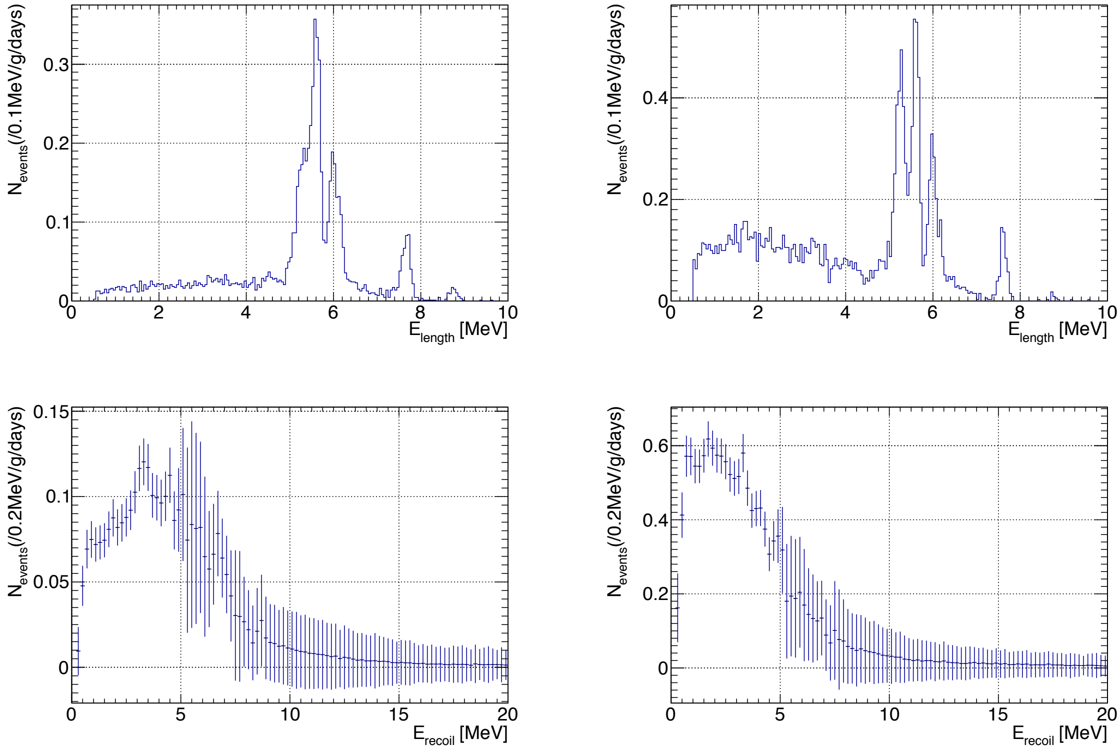


Figure 5.27: Comparison of the raw (top) and fully corrected recoil spectra (bottom) for Near (left) and Far (right) laboratories.

5.7 Prospects for DCTPC

5.7.a Reconstruction of Neutron Energy

Reconstructing the neutron energy from a recoil energy is not a trivial task as the recoil energy depends on the recoil angle, which is not isotropic when the recoiling nucleus is not Hydrogen. In most experiments, the impacting neutron flux is usually simulated and the recoil

Table 5.3: Number of events, and the corresponding rates for both experimental laboratories of Double Chooz.

	step	number of events	rate ($\mathrm{d}^{-1}\,\mathrm{g}^{-1}$)
Near Laboratory	before α subtraction	9692 ± 98.45	11.17 ± 0.11
	after α subtraction	3683.37 ± 149.53	4.24 ± 0.13
	after unfolding	3683.37 ± 149.53	4.24 ± 0.13
	after efficiency correction	8626.59 ± 481.08	9.94 ± 0.41
Far Laboratory	before α subtraction	5418 ± 73.61	16.34 ± 0.22
	after α subtraction	2903.06 ± 96.92	8.75 ± 0.24
	after unfolding	2903.06 ± 96.92	8.75 ± 0.24
	after efficiency correction	4710.48 ± 297.84	14.20 ± 0.69

and detector transfer matrices are applied to the simulation to simulate the observed recoil data. Unfortunately Double Chooz did not maintain a neutron simulation in both laboratories, as Fast Neutrons are quantified directly from IBD candidates.

Therefore the understanding of the shape of the recoil spectrum is still in progress. The current reconstruction method is based on Bayesian unfolding using cross sections and angular distributions from [60] to reconstruct the possible neutron spectrum impacting the TPC.

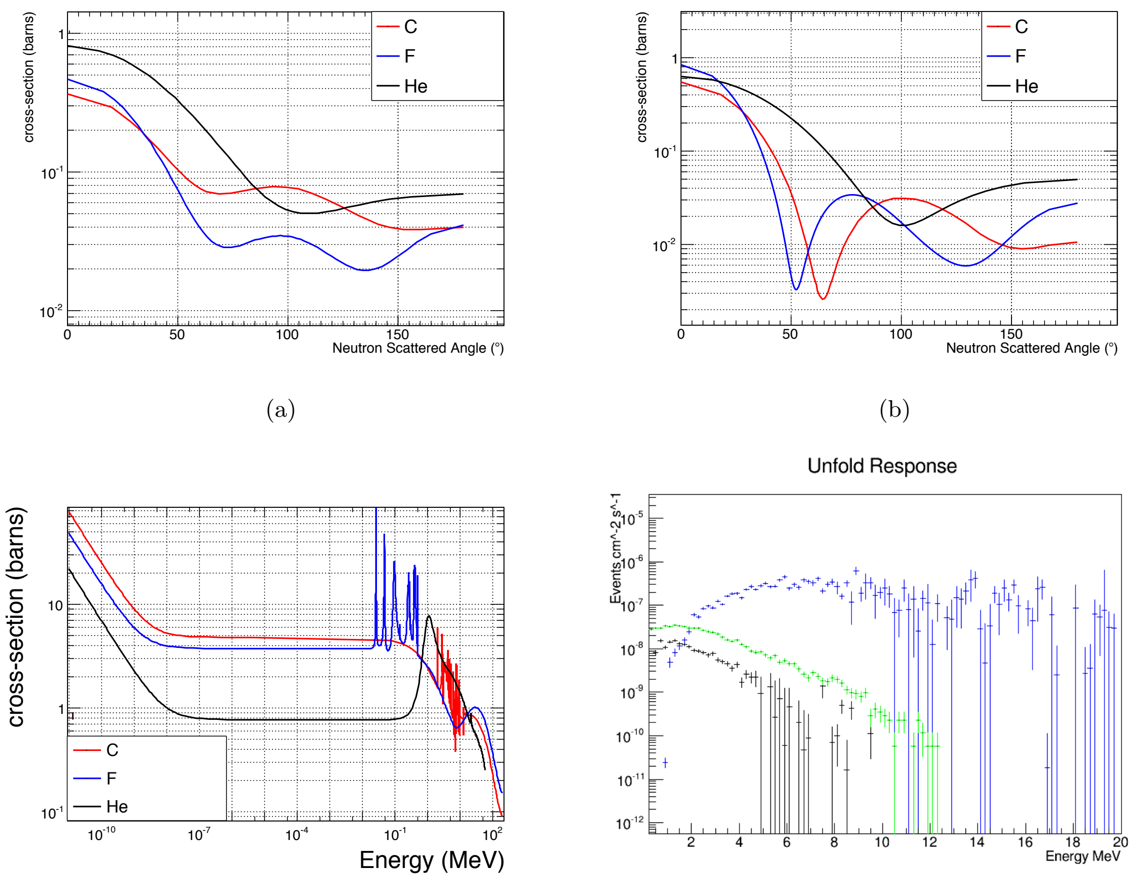
Figures 5.28(a) and 5.28(b) show the angular distribution for elastic scatterings of neutrons of respectively 5 MeV and 10 MeV on the three dominant nuclei in the gas: He, F and C. Figure 5.28(c) shows the elastic scattering cross section on each nucleus. DCTPC is sensitive to Helium recoils of 1 MeV to 10 MeV, hence to neutrons of energies higher than 4 MeV. At these energies, the Helium cross section is barely dominant, but Helium constitute 87.5% of the target nuclei, and the dE/dX of recoils on Carbon and Fluorine do not fit with the SRIM simulation used to reconstruct the energy, as such only Helium recoils are taking into account.

Figure 5.28(d) shows a preliminary result from the Bayesian unfolding technique. The original recoil spectrum is shown in black and the reconstructed neutron spectrum in blue. The green dots correspond to the re-folded recoil spectrum, neglecting efficiencies, to cross-check the method. The green dots follow the well the shape of the original spectrum, showing the consistency of this technique. The shape of the reconstructed neutron spectrum is not fully understood at this time, it is believed that the rising part at low energies is not physical and corresponds to a threshold effect of our detection and reconstruction process.

5.7.b Neutron directionality

Neutrons showering from muon should follow a preferred direction according to the muon track. The directionality of muons in Double Chooz laboratories is strongly determined by the profile of the overburden, shown in Figure 5.29.

Figure 5.30(a) shows the simulated angular distribution of muons in the Far Laboratory of Double Chooz in the (ϕ, θ) plane. Figure 5.30(b) shows the observed angular distribution



(a)

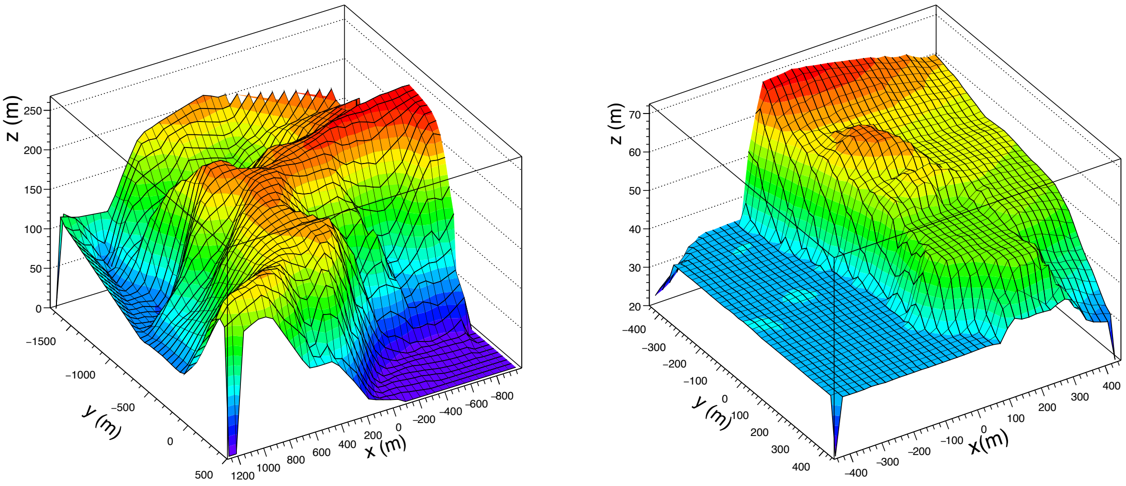
(b)

(c)

(d)

Figure 5.28: 5.28(a): Scattering angle distribution for a 5 MeV neutron on He, F, and C.
5.28(b): Scattering angle distribution for a 10 MeV neutron on He, F, and C.
5.28(c): Neutron elastic scattering cross section on He, F, and C.
5.28(d): Result from a Bayesian unfolding taking the previous distributions into account. The black dots correspond to the original recoil spectrum, the blue dots correspond to the unfolded neutron spectrum, and the green dot correspond to the re-folded recoil spectrum, to cross-check the method.

of the recoil tracks, with $E_{length} < 4\,\mathrm{MeV}$, in DCTPC running in the Far Laboratory. A hot spot appears in the DCTPC distribution in agreement with the position of the hot spot of the Muons in Double Chooz. The preferred direction in DCTPC is not as well defined as the direction of the muons because angular information is lost during the generation of neutrons, their scattering in the rocks, and the angular distribution of elastic scatterings on the TPC gas nuclei. Moreover, muon-induced neutrons represent only a fraction of the events in DCTPC as some neutrons from (α, n) reactions in the rock may appear, without any correlation with the muon tracks.



(a) (b)

Figure 5.29: 5.29(a): Profile of the hill above the Far Laboratory.
5.29(b): Profile of the hill above the Near Laboratory.

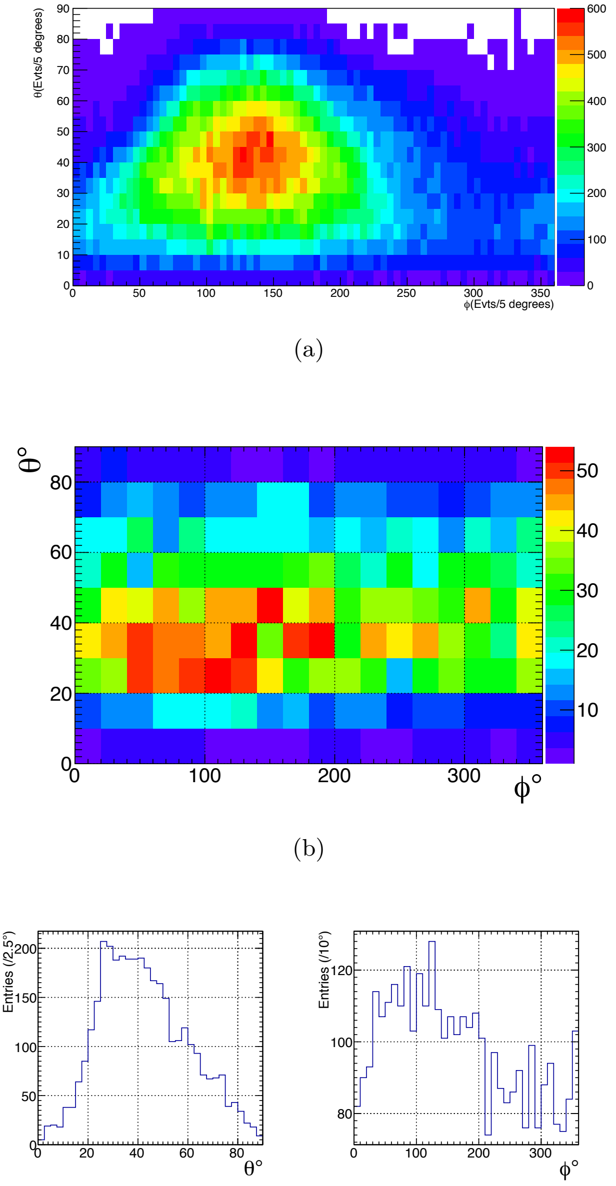
5.8 Conclusion

DCTPC was scheduled for two six-months data taking periods, one for each laboratory. For the first data taking period, DCPTC ran at the Near Laboratory. This phase was still a transitory period, during which we could understand the operations of the TPC and its maintenance. Operation of the TPC became much more stable in the Far Laboratory, and quality data taking became possible.

During both run periods, we developed a measurement of energy relying on the length of the tracks, and obtained an energy resolution at the percent level in both cases. This length base energy measurement was found to be very stable in time, and the validity over the full energy range of interest was tested by using a partially attenuated ^{241}Am source.

The directionality of a recoil is also available when reconstructing the its 3D-length. Although there is no information on an event-by-event basis, it was possible to find a preferred direction for the recoils, and therefore the incoming neutrons correlated to the direction of the muons in the Far Laboratory.

Although He seemed like a good target choice in the energy range of DCTPC because of its lightness, and therefore low energy threshold, it became clear that the ability to differentiate neutron-induced nuclear recoils from the α-background through a difference in dE/dX would have been appreciable. The data taking at Double Chooz laboratories is now over, the TPC has been sent to Fermilab where it measures the rate of beam-induced neutrons for the MicroBooNE experiment. For this new running period, we chose to replace He by Ar. Tracks are shorter



(a)

(b)

(c)

Figure 5.30: 5.30(a): Simulation of the angular distribution of the muons in Double Chooz for the Far Detector.
5.30(b): Angular distribution of recoils in DCTPC for the runs in the Far Laboratory.
5.30(c): Individual θ and ϕ distribution for DCTPC runs in the Far Laboratory.

for a same energy, but the α-background is now easily identifiable. Moreover, the coincidence with the beam time allows for a good background rejection, while still allowing background measurements during beam-off periods.

Chapter 6

Reducing the Stopping Muon Contamination in the IBD Candidates

One of my contributions to the Double Chooz analysis has been to develop Stopping Muons rejection techniques with high rejection power while keeping a low inefficiency.

In this chapter, I will describe my work in developing rejection strategies and quantifying their efficiency and impact on the IBD sample.

6.1 Stopping Muons in Double Chooz

Atmospheric muons in Double Chooz are tagged with the Outer Veto, Inner Veto, and a high energy muon threshold in the Inner Detector ($E_{p,\text{max}}$). The Inner Veto offers a nearly 4π-coverage of the Inner Detector, so most muons are tagged thanks to this volume. However, both the Inner and Outer Vetoes are pierced by the Chimney, the access tube to the Target and the γ-Catcher. The Chimney allows for a small acceptance hole in the veto volumes, allowing muons, among other backgrounds, to penetrate the detector unseen.

Out of the muons that enter through the chimney, muons that deposit an energy in the Inner Detector higher than the muon threshold are tagged and rejected, and a muon veto of 1 ms is applied. Muons that decay while still at the top of the Inner Detector may deposit a lower energy than $E_{p,\text{max}}$ and not be tagged. These muons are called Stopping Muons (SM).

Muons can decay, each producing a Michel electron and two neutrinos:

$$\begin{aligned} \mu^- &\to e^- + \bar{\nu}_e + \nu_\mu, \\ \text{and } \mu^+ &\to e^+ + \nu_e + \bar{\nu}_\mu, \end{aligned} \tag{6.1}$$

thus producing a prompt signal, the muon scintillation, and a delayed signal, the Michel electron scintillation. These signals are highly correlated in time and space.

μ^- can be captured in an atomic orbital, and rapidly cascade down to the $1s$ state. From this state, it can either decay, following the process described in Equation 6.1, or may undergo nuclear capture, a process analogous to the electron capture:

$$\mu^- + {}^{\mathrm{A}}_{\mathrm{Z}}\mathrm{X} \rightarrow {}^{\mathrm{A-1}}_{\mathrm{Z-1}}\mathrm{Y} + n + \nu_\mu \tag{6.2}$$

However, due to the large muon mass, a wider range of final states is available, including the emission of a free nucleon. Such process also creates a prompt/delayed coincidence with a time correlation dominated by the neutron capture, hence irreducible from an IBD candidate. However, a dedicated study [103] demonstrated this background to be negligible in Double Chooz.

The overburden in the Near Detector is smaller than in the Far Detector, thus the Muon background rate in the Near Detector is ~ 5 times higher than in the Far Detector [104]. However, the muon spectrum is also softer, with more muons at low energies, thus the Stopping Muon rate is ~ 20 times higher in the Near Detector than in the Far Detector, making the Stopping Muon reduction a priority in the Near Detector.

The Outer Veto, and especially the Upper Outer Veto, that covers the Chimney, is a very useful tool to tag muons and Stopping Muons. However, because of different running periods, a significant amount of data does not benefit from it and we need other ways to reject Stopping Muons.

6.2 Correlation Time Study

Muons decay with a life time of $\tau_\mu = 2.197\,\mu\mathrm{s}$ [31]. τ_μ is relatively short, and as such, Stopping Muons contribute to the correlated background along with Fast Neutrons. To quantify the Stopping Muon contamination, it is useful to understand the time distribution of the IBD candidates.

The cross section of neutron capture on Gd follows a power law, with a sharp decrease between 10^{-1} and 10 eV as shown in Figure 6.1(a). IBD neutrons are emitted with an energy of a few keV [51]. The capture cross section at these energy is much lower than at thermal energies, thus the neutrons need to lose energy, and thermalize in order to be efficiently captured on Gd.

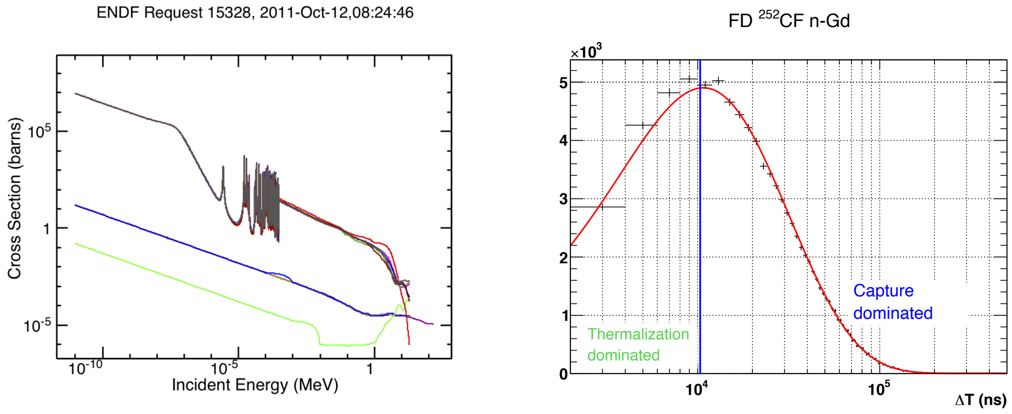
The abundance ratio of Gd and H in the Target is

$$\frac{n_H}{n_{Gd}} = 1.72 \times 10^4 \tag{6.3}$$

Thus, the ratio of cross sections, for capture on Gd to be dominant (at least 10 times more probable) needs to be:

$$\frac{\sigma_{Gd}}{\sigma_H} > 10^5 \tag{6.4}$$

This condition is achieved for thermal energies, lower than $\sim 0.1\,\mathrm{eV}$. At higher energies, the captures on H and Gd are in competition.



(a) (b)

Figure 6.1: 6.1(a): Cross sections of neutron captures on Gadolinium (grey), Hydrogen (blue), and Carbon (green) [60].
6.1(b): Correlation time distribution for neutron capture on Gadolinium for ^{252}Cf in the center of the Target in the Far Detector. Two regimes are distinguishable, at high ΔT, the neutron capture is dominant, but at low ΔT, the neutrons have too high an energy to capture on Gd, and need to lose energy, mostly by elastic scattering on Hydrogen. This regime is dominated by the thermalization process.

The need for thermalization effectively lowers the capture rate at low delayed-prompt time difference ΔT since most neutrons have too high an energy to be available for capture.

Figure 6.1(b) shows the ΔT distribution of neutron captures on Gd for a ^{252}Cf source in the center of the target. Two regimes are distinguishable:

- $\Delta T < 10\,\mu\text{s}$: regime dominated by thermalization. As ΔT increases, more and more neutrons reach thermal energies and become available for capture, thus the capture rate effectively increases.
- $\Delta T > 10\,\mu\text{s}$: regime dominated by capture. Most neutrons have reached thermal energies, the ΔT distribution is dominated by the exponential behaviour of the capture process.

The thermalization at low energy (epithermal to thermal energies) happen mostly through elastic scattering on Hydrogen nuclei. As protons and neutrons have similar masses, the neutron loses up to half its energy at each collision. Scattering on other nuclei may occur, but are less efficient for thermalization.

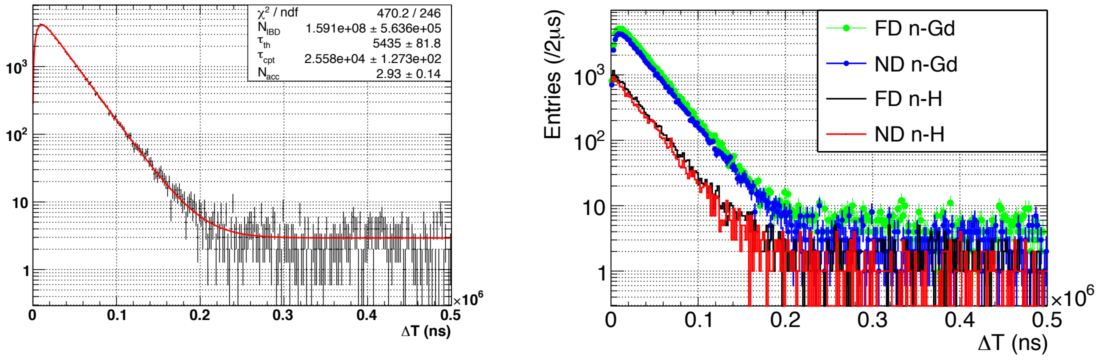
Two characteristic times appear, $\tau_{th} = \mathcal{O}(5\,\mu\text{s})$ and $\tau_{cpt} = \mathcal{O}(25\,\mu\text{s})$ respectively describing the thermalization and capture processes. They can be included in a model to fit the ΔT distribution. Each process can be described, with a good approximation, by an exponential. The processes being consecutive, the neutrons need to thermalize to be able to capture on a nucleus, the rate of captures is the convolution of the thermalization process and the capture of the thermalized neutron:

$$
\begin{aligned}
\mathrm{rate}(t) &= \mathrm{thermalization}(\tau_{\mathrm{th}}) \otimes \mathrm{capture}(\tau_{\mathrm{cpt}}) \\
\mathrm{rate}(t) &= \frac{N}{\tau_{cpt} - \tau_{th}}\, e^{\dfrac{-t}{\tau_{cpt}}} \times \left(1 - \; e^{\dfrac{-t(\tau_{cpt} - \tau_{th})}{\tau_{cpt}\tau_{th}}} \right)
\end{aligned}
\tag{6.5}
$$

The red line in Figure 6.1(b) shows the best fit of this model to the calibration data.

For neutron captures on the Hydrogen and carbon however, there is no transition in the cross section. The neutrons do not need to reach a certain energy to be efficiently captured. They can be captured as soon as they are produced, thus there is no need for a thermalization time in the model, the ΔT distribution can be described by a simple exponential.

Figure 6.2(a) shows the ΔT distribution and best fit result for a ^{252}Cf source in center of the Near Detector Target for neutron captures on Gd. A constant term N_{Acc} is added to account for Accidental background. The captures on H and Gd are studied separately.
Figure 6.2(b) shows the distributions for both captures in both detectors (ND and FD-II). The n-Gd samples are shown as green dots for the Far Detector and blue dots for the Near Detector, and the n-H samples are shown as black line for the Far Detector and red line for the Near Detector.

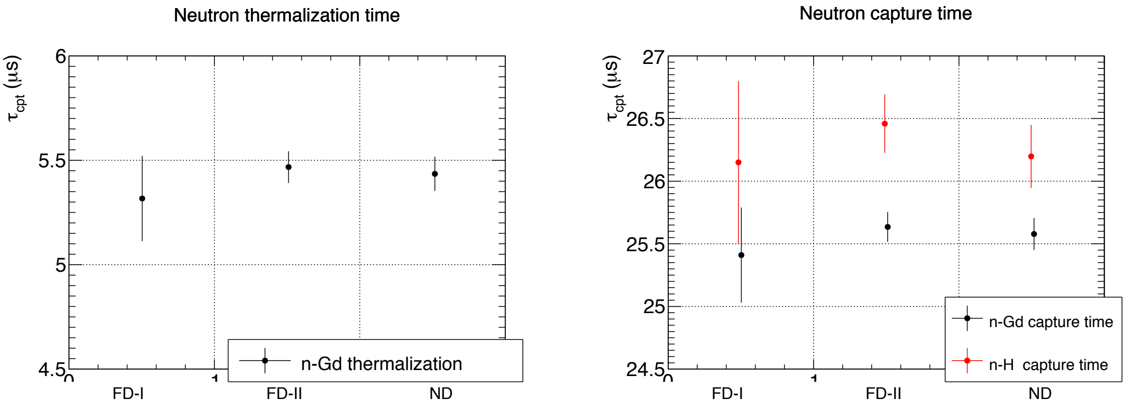


(a)

(b)

Figure 6.2: Calibration data in the Near and Far Detectors for ^{252}Cf source in the middle of the Targets. 6.2(a): Example of fit to the ΔT distribution of the n-Gd capture in the Near Detector. 6.2(b): ΔT distributions for n-Gd and n-H captures in the Far Detector (respectively green dots and black line) and the Near Detector (respectively blue dots and red line).

The fit results for each distribution from Figure 6.2(b) are presented in Figure 6.3, along with a similar study performed during the calibration of the FD-I configuration. The good agreement between the three configurations shows a good consistency of the physics in the various configurations, and especially, the consistency of the composition of the liquid scintillators since the characteristic times are identical. Especially, the FD-I and FD-II configurations of the Far Detector are expected since only the electronics changed, which has no impact on neutron capture.



(a)

(b)

Figure 6.3: Thermalization and capture times for FD-I, FD-II and ND. The good agreement between the three configurations shows a good consistency of the physics in the various configurations, and especially, the consistency of the Hydrogen component in the liquid scintillators.

The capture time of neutrons depends on the composition of the volume in which they travel, as $\tau_{\text{cpt}} \propto \sigma_{\text{cpt}} \times n_{\text{targets}}$ where σ_{cpt} is the effective capture cross section of the volume and n_{targets} is the number of target nuclei in the volume.

Previous results were shown for captures in the target, and exhibited a capture time of $\sim 25.5\,\mu\text{s}$. For studies of the n-Gd sample, this clarification is irrelevant as Gd is only present in the target. However, for the n-H sample, or the Gd++ analysis, neutrons can capture in the γ-Catcher as well, which is not loaded with Gd, and thus has a lower effective cross section, since it is dominated by capture on H. Therefore, a longer capture time is then expected for events capturing in the γ-Catcher.

Figure 6.4 shows the ΔT distribution for IBD candidates in the Near Detector, for Gd++ analysis (black line) where neutrons can capture in both the Target and the γ-Catcher. The Accidental contamination is reduced by rejecting events with a prompt energy higher than 4 MeV. Two regimes can be observed here: a short-lived sample, corresponding to the captures in the Target, with a capture time of $\sim 25.5\,\mu\text{s}$, and a long-lived sample, corresponding to the captures in the γ-Catcher, with a capture time of $\sim 180\,\mu\text{s}$.

The blue line corresponds to the sample with captures on Gd, i.e. in the Target, scaled to the black histogram between 5 and 20 µs. The green line corresponds to the sample with captures on H, with a fiducial cut on the delayed event at $\rho_d > 1.5\,\text{m}$, to make sure that the distribution is dominated by captures in the γ-Catcher, and scaled to the black histogram between 200 and 800 µs.

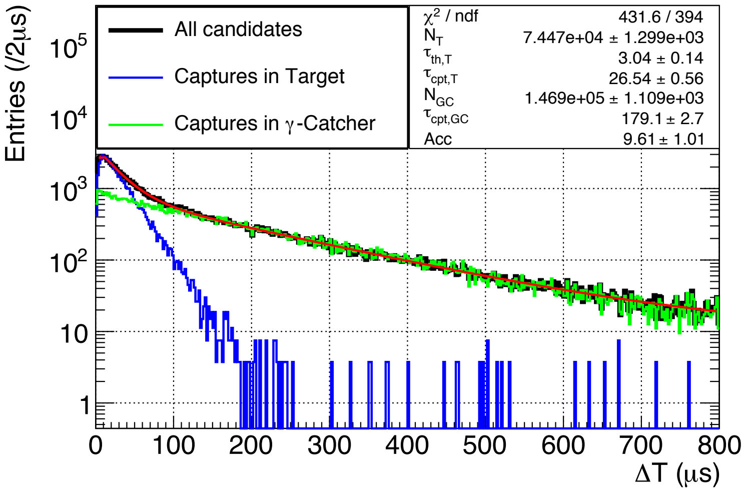


Figure 6.4: ΔT distributions for capture in different volumes for the Gd++ analysis in the Near Detector. The black line corresponds to all IBD candidates, the blue line corresponds to neutron captures in the Target, and the green line corresponds to neutron captures in the γ-Catcher. The red line corresponds to a fit to the black histogram with the full-volume model.

The model previously exposed is only valid in one of the volumes, as only one capture time is included. It is however possible to add a new contribution:

$$\text{rate}(t) = \frac{N_{\text{T}}}{\tau_{\text{cpt, T}} - \tau_{\text{th, T}}}\, e^{\frac{-t}{\tau_{\text{cpt, T}}}} \times \left(1 - e^{\frac{-t(\tau_{\text{cpt, T}} - \tau_{\text{th, T}})}{\tau_{\text{cpt, T}}\tau_{\text{th, T}}}}\right) + \frac{N_{\text{GC}}}{\tau_{\text{cpt, GC}}} \times e^{\frac{-t}{\tau_{\text{cpt, GC}}}} + N_{\text{acc}} \tag{6.6}$$

The red line in Figure 6.4 shows the best fit of the model in Equation 6.6. The capture time in the Target is compatible with the one measured in the ^{252}Cf source.

Using this model, we were able to quantify the difference between two Hydrogen packages in the detector simulation for [54], a default Geant4 package assuming free Hydrogen in the target, and a custom package introducing bonding of the Hydrogen in heavier molecules called NeutronTH, leading to a slightly less efficient neutron thermalization and a longer thermalization time. Figure 6.5 shows the ΔT distribution of events from a ^{252}Cf source in the Target (black dots), compared to the standard Geant4 neutron model (blue), and NeutronTH (yellow). The bonding of Hydrogen modifies the energy transfer during the elastic scatters, and thus changes the thermalization time.
The standard MC package and NeutronTH show noticeable discrepancies below 10 µs, where the Stopping Muons are located.Therefore, a precise knowledge of the thermalization rising edge of the distribution is crucial when attempting to quantify the Stopping Muons contamination. We decided to use the custom model, as its thermalization time was closer to the observed data.

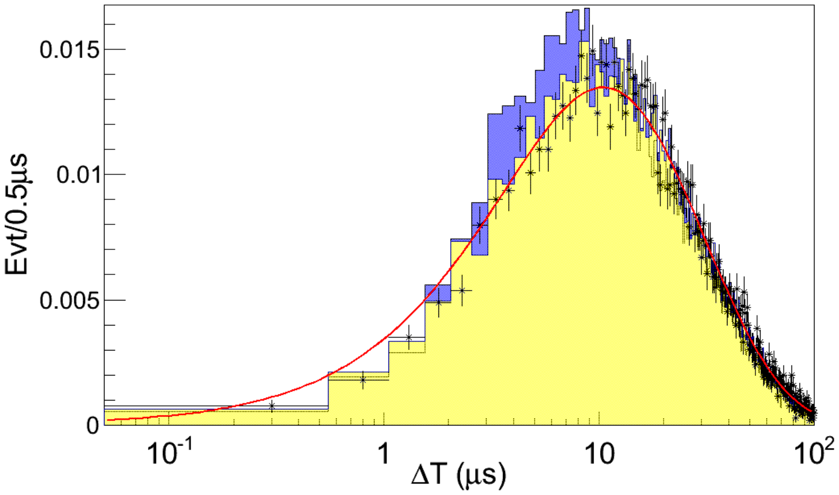


Figure 6.5: Discrimination between two MC models for Hydrogen:
The blue histogram corresponds to the default Geant4 package, assuming free Hydrogen atoms. The yellow histogram corresponds to Double Chooz's custom Hydrogen package, assuming that Hydrogen atoms are bound to heavier molecules. The black dots correspond to ΔT distribution of events from a ^{252}Cf source in the Far Detector. The bound-H model leads to a higher thermalization time, closer to the observed behaviour of data.
The red line corresponds to the neutron capture model from equation 6.5 fitted to the data points.

The model can be refined to better describe the data by adding an exponential term $\dfrac{N_{SM}}{\tau_{SM}}\, e^{(-t/\tau_{SM})}$ to account for Stopping Muon contamination.

6.3 Functional Value

The use of the Functional Value to reject Stopping Muons was first developed for the DC-III n-Gd analysis [61] and was later used in the DC-III n-H analysis. I was in charge of optimizing the cut for the n-H analysis and later on, I applied the method to the DC-IV analysis for both n-Gd and Gd++ analyses.

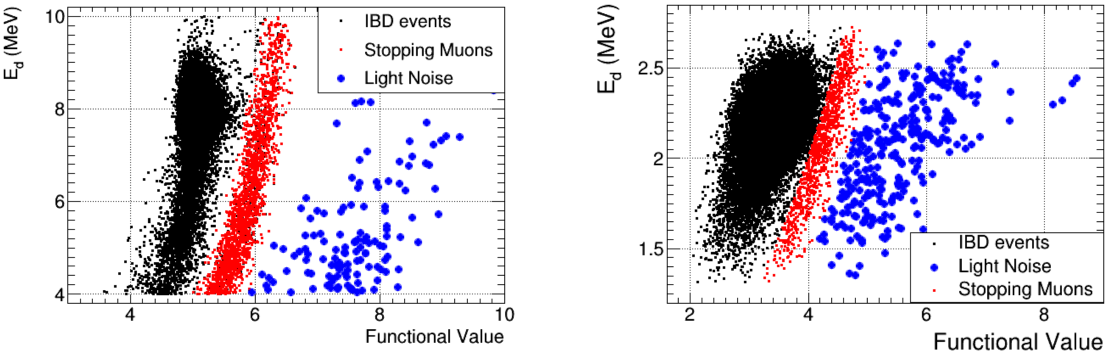
6.3.a Variable Presentation

Several variables have been investigated to separate Stopping Muons from IBD candidates. The Inner Veto can only reject muons crossing the Inner Veto, but not the ones entering through the Chimney.

As previously explained in Section 2.8.b, The vertex reconstruction algorithm called RecoBAMA is based on the hypotheses that the energy deposition in the detector is point-like, and that the subsequent light propagation is isotropic. A likelihood is built taking into account the expected charge in non-hit PMTs and the collected charge and hit time of hit PMTs:

$$\mathcal{L}(\vec{X}) = \prod_{q_i=0} f_q(0;\mu_i) \prod_{q_i>0} f_q(q_i;\mu_i) f_t(t_i;t_i^{pred},\mu_i)$$

$$FV(\mathcal{L}) = -\log\mathcal{L}(\vec{X}) = -\sum_i \ln f_q(q_i;\vec{X}) - \sum_{q_i>0} \ln f_t(t_i;\vec{X}) = FV_q(\vec{X}) + FV_t(\vec{X}) \tag{6.7}$$



(a)

(b)

Figure 6.6: The distribution of delayed energy versus Functional Value shows three populations: IBD events (black), Stopping Muon events (red), and remaining Light Noise events (blue) for both n-Gd (6.6(a)) and n-H (6.6(b)) captures.

Figure 6.6 shows the distribution of delayed energy against the delayed Functional Value (FV_d) for the DC-III analysis, n-Gd and n-H. For the n-H sample, the ANN have been applied by default to reduce the population of Accidentals. As the FV is correlated with the energy of the event, only the delayed trigger will be used so as not to bias the energy spectrum of the prompt, and the measure of $\sin^2(2\theta_{13})$. In each distribution, three populations can be identified:

- IBD-like events (black),
- Stopping Muons (red),
- Light Noise (blue).

Stopping Muons decay in the Chimney, so the Michel electron scintillates in the Chimney as well. An event occurring in the Chimney is outside the volume of the detector, and a pull term is used to lower it inside the Target, at the cost of lowering the likelihood. Moreover, because of Chimney shadow effects, the assumption that light propagates isotropically is not true any more. Those effects explain why Chimney Stopping Muons have such a high Functional Value.

Identically, Light Noise happens at the walls of the Buffer. A pull term in RecoBAMA penalizes reconstruction in the Buffer, pulling the events back in the center of the detector. Hence Light Noise has a high Functional Value as well.

It is therefore possible to separate IBD from Stopping Muons using this distribution, with a selection based on both the delayed energy (E_d) and the delayed Functional Value (FV_d), later on referred to as the FV-cut.

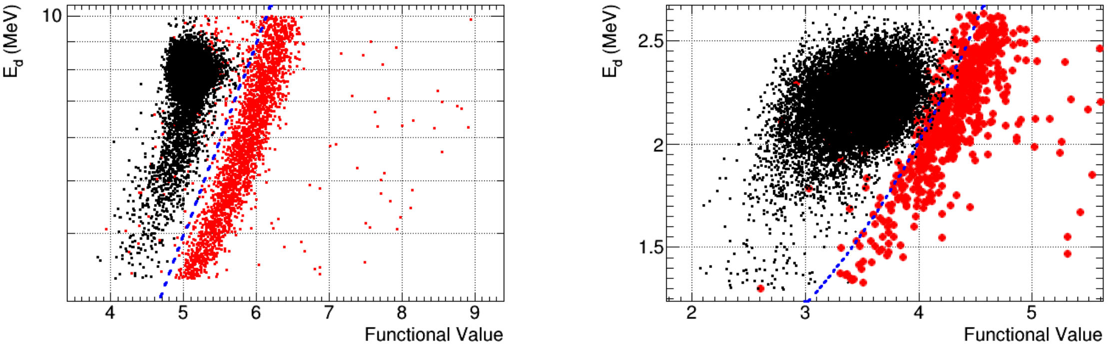
In an effort to reject background, the exact limit before Stopping Muon and Light Noise does not need to be precisely identified as both need to be rejected. However, it may be difficult to separate Light Noise from IBD when estimating the efficiency of the selection. Therefore, a looser FV-cut is applied to remove Light Noise, far away from the IBD region enough that it leads to no inefficiency on the IBD sample. The IBD/Stopping Muons limit is then tuned independently of Light Noise.

6.3.b Cut Parametrization

The Functional Value behaves logarithmically with respect to the energy as visible in Figure 6.7. To follow this behaviour, the parametrization of the cut is the exponential in FV:

$$E_d > \mathrm{A} \times e^{\mathrm{FV}_d/\mathrm{B}} \tag{6.8}$$

Events with lower delayed energy, i.e. below the line, are rejected as Stopping Muons or Light Noise. This 2-parameter optimization can be brought down to one parameter when fixing the slope of the exponential to follow that of the IBD candidates (1.23 for n-Gd, and 2.0125 for n-H).The only remaining parameter is a normalization parameter (or offset of the straight line in logarithmic scale).



(a)

(b)

Figure 6.7: Parametrization of the Functional Value cut to follow the exponential behaviour of the log-likelihood with energy. Black distributions correspond to IBD-like events and red distributions to Stopping Muons or Light Noise, for n-Gd (6.7(a)) and n-H (6.7(b)).

6.3.c Criteria for Cut Optimization

The $\sin^2(2\theta_{13})$ analysis requires that the number of IBD candidate is well known, it is thus important to estimate the inefficiency of the FV cut, i.e. how many IBD candidate are misinterpreted as Stopping Muons and rejected (error of type I).

An ideal cut would reject as many muons as possible while rejecting the fewest possible IBD candidates. A possible estimator of the goodness of the cut is therefore the ratio of the inefficiency over the number of rejected Stopping Muons: $\bar{\epsilon}/N_{\mu}^{\mathrm{rej}}$.

The inefficiency of the FV cut is estimated after all other vetoes and selection are applied such that, as much as possible, only IBD and Stopping Muons are present in the sample. Irreducible contaminations of Accidentals and Fast Neutrons are always present, and have to be estimated.

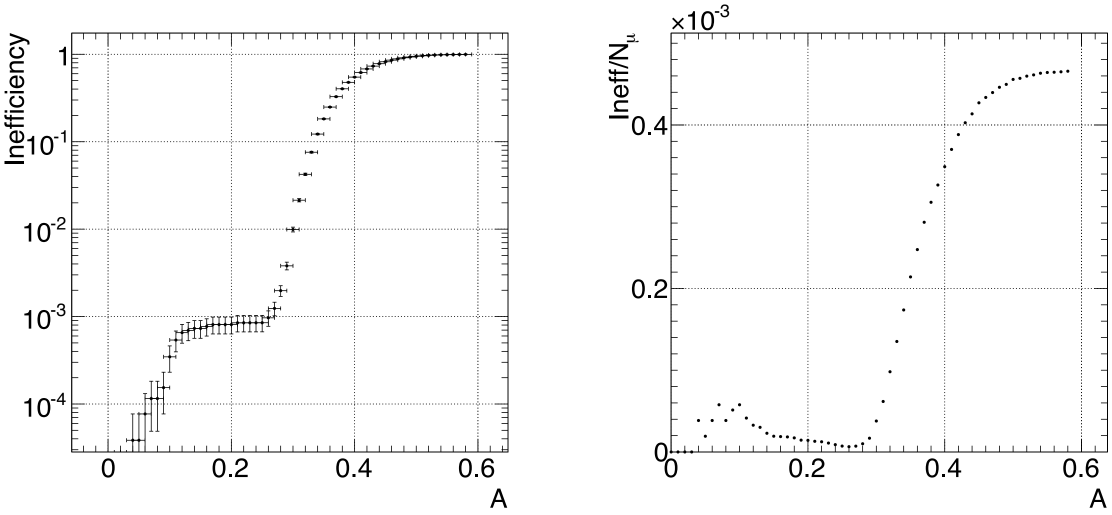
In order to reduce contamination from Accidentals, events with prompt energy $E_p < 2\,\mathrm{MeV}$ are rejected.
FV-rejected events with $\Delta T < 10\,\mathrm{\mu s}$ are counted as rejected Stopping Muons N_{μ}^{rej}.
FV-rejected events with $\Delta T > 10\,\mathrm{\mu s}$ and $E_p > 10\,\mathrm{MeV}$ are counted as high energy Fast neutrons $N_{FN,HE}^{\mathrm{rej}}$ and scaled from $[10-20]$MeV to $[2-10]$ MeV assuming a flat FN spectrum: $N_{FN}^{\mathrm{rej}} = N_{FN,HE} * (10-2)/(20-10)$.
Finally, the number of rejected IBD $N_{\mathrm{IBD}}^{\mathrm{rej}}$ is the number of rejected events with $\Delta T > 10\,\mathrm{\mu s}$ and $E_p < 10$ minus N_{FN}.

$$\begin{aligned}
\frac{\bar{\epsilon}}{N_{\mu}^{\mathrm{rej}}} &= \frac{N_{\mathrm{IBD}}^{\mathrm{rej}}/N_{\mathrm{IBD}}^{\mathrm{tot}}}{N_{\mu}^{\mathrm{rej}}} \qquad (6.9)\\
N_{\mu}^{\mathrm{rej}} &= N((\cancel{\mathrm{FV}}) \cap (2\,\mathrm{MeV} < E_p < 20\,\mathrm{MeV}) \cap (\Delta T < 10\,\mathrm{\mu s}))\\
N_{FN}^{\mathrm{rej}} &= N((\cancel{\mathrm{FV}}) \cap (10\,\mathrm{MeV} < E_p < 20\,\mathrm{MeV}) \cap (\Delta T > 10\,\mathrm{\mu s})) \times \frac{10-2}{20-10}\\
N_{FN}^{\mathrm{acc}} &= N((\mathrm{FV}) \cap (10\,\mathrm{MeV} < E_p < 20\,\mathrm{MeV}) \cap (\Delta T > 10\,\mathrm{\mu s})) \times \frac{10-2}{20-10}\\
N_{\mathrm{IBD}}^{\mathrm{rej}} &= N((\cancel{\mathrm{FV}}) \cap (2\,\mathrm{MeV} < E_p < 10\,\mathrm{MeV}) \cap (\Delta T > 10\,\mathrm{\mu s})) - N_{\mathrm{FN}}^{\mathrm{rej}}\\
N_{\mathrm{IBD}}^{\mathrm{tot}} &= N(\mathrm{FV}) \cap (2\,\mathrm{MeV} < E_p < 10\,\mathrm{MeV}) \cap (\Delta T > 10\,\mathrm{\mu s})) - N_{FN}^{\mathrm{acc}} + N_{\mathrm{IBD}}^{\mathrm{rej}}
\end{aligned}$$

Figure 6.8(a) shows the evolution of the inefficiency of the FV cut with the parameter A of the cut. As the parameter A increases the cut moves further towards the IBD region. The inefficiency shows a plateau up to 0.28 as it only rejects Stopping Muons, and then increases very quickly because of the rejection of the IBD candidates. The rise of the inefficiency at low values is due to Light Noise events misinterpreted as IBD.

Figure 6.8(b) shows the evolution of the ration $\bar{\epsilon}/N_{\mu}^{\mathrm{rej}}$ with the parameter A. The ratio first decreases as the number of IBD events does not increase over the range of the efficiency plateau, while the number of rejected Stopping Muons increases. Then the ratio increases as the number of rejected Stopping Muons does not evolves but the number of rejected IBD increases.

The chosen value for the cut parameter A corresponds to the position of the end of the inefficiency plateau and the minimum of the ratio $\bar{\epsilon}/N_{\mu}^{\mathrm{rej}}$. The cuts for DC-III are expressed as:



(a) (b)

Figure 6.8: FV cut optimization for DC-III n-H analysis. The chosen value for parameter A is the position of the minimum of the ratio $\bar{\epsilon}/N_\mu^{\text{rej}}$: 0.2755.

$$\text{n-Gd} : E_d > 0.068 \times e^{\text{FV}_d/1.23} \tag{6.10}$$

$$\text{n-H} \; : E_d > 0.2755 \times e^{\text{FV}_d/2.0125} \tag{6.11}$$

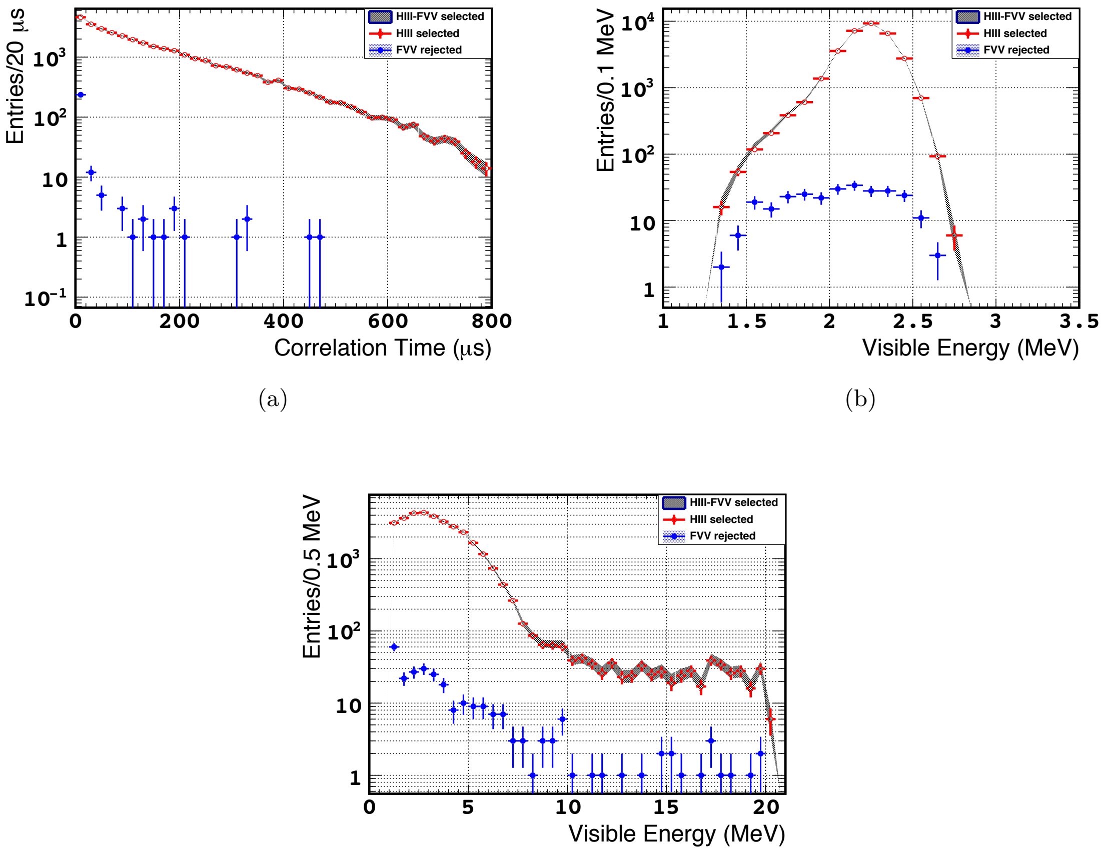
In order to display the evolution of the efficiency and the Inefficiency/N_μ ratio, the Light Noise events were kept. To compute the true inefficiency, they must be taken out, which lowers the inefficiency by a constant value. The inefficiencies of the FV cut on the IBD sample for n-Gd and n-H samples of the DC-III analysis are:

$$\text{n-Gd} : \bar{\epsilon} = (5.0 \pm 1.5) \cdot 10^{-4}$$

$$\text{n-H} \; : \bar{\epsilon} = (4.64 \pm 1.4) \cdot 10^{-4}$$

Figure 6.9 shows the "N-1" distributions for the FV-cut on the DC-III n-H analysis, i.e. the modification operated by the FV-cut on the distributions when all other cuts are applied. Figure 6.9(a) shows that more than 90% of the rejected events are piling-up in the first bin, corresponding to the Stopping Muon events. Indeed, Figure 6.9(b) shows that the rejected sample is flat at the energy of the n-H capture peak, furthering our confidence that the rejected events are not IBD candidates.

Figure 6.9(c) shows an unexpected behaviour of the Stopping Muons spectrum. Indeed, the Stopping Muon spectrum was found to be flat in the n-Gd sample [61], while it is not flat any more in the n-H sample. This trend also appeared in the Fast Neutron spectrum for which, as explained in Section 4.4.c, the flat spectrum approximation could no longer be used.



(a)

(b)

(c)

Figure 6.9: Impact of the FV cut on HIII distributions. For each distribution, all cuts other than FV are applied (black shaded area), the red dots shows the final selected sample, selected by the FV cut, and the blue dots correspond to the FV-rejected events. Distributions shown are the ΔT distribution (6.9(a)), the delayed energy distribution (6.9(b)) and the prompt energy distribution (6.9(c)).

6.3.d Estimation of the Remaining Contamination

A profile of the E_d vs FV_d distribution in Figure 6.7(b) along the parameter A variable can be achieved by counting the number of events between two positions of the cut line (i.e between $E_d = A \cdot e^{FV_d/B}$ and $E_d = (A + \delta A) \cdot e^{FV_d/B}$). A higher value for A means a more stringent cut, and thus, cuts more into the IBD region.

The remaining contaminations can be estimated from the profile shown in Figure 6.10. The rising edge of the IBD region is modelled by a Gaussian (in blue) as well as the SM distribution (in green). Two models of Light Noise (in grey) have been tested : a flat distribution (left) and a linear distribution (right).

The remaining Stopping Muon contamination after application of the FV cut can be esti-

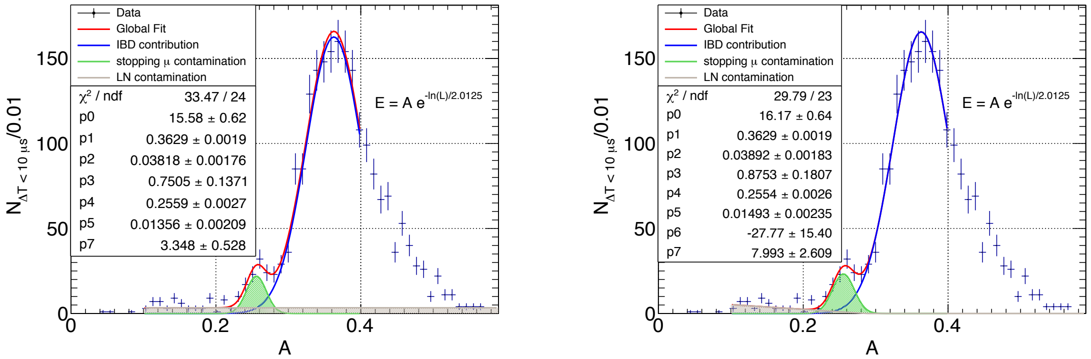


Figure 6.10: Profile of the E_p vs FV_d distribution along the moving direction of the cut. Two models of LN have been considered, one with a flat distribution (left) and one with a linear distribution (right).

mated by integrating the Stopping Muon Gaussian for $A > 0.2755$:

$$\begin{aligned} \text{flat LN} &: 10.4^{+4.0}_{-3.6}\ \text{SM} \\ \text{linear LN} &: 7.6^{+3.4}_{-3.0}\ \text{SM} \\ \text{combination} &: 9.0^{+3.7}_{-3.2}(\text{stat.}) \pm 1.4(\text{syst.}) \quad \text{SM} \end{aligned} \tag{6.12}$$

Over 467.903 live days of data taking, the remaining SM contamination is

$$(1.9^{+0.8}_{-0.7} \pm 0.7) \cdot 10^{-2}\ (\text{day})^{-1}. \tag{6.13}$$

6.4 DC-IV Stopping Muon Rejection Strategy

The DC-IV analysis bears on the data of both Near and Far Detectors, with FD-I and FD-II configurations of the Far Detector. Both n-Gd and Gd++ analyses were performed, as it was demonstrated in DC-III that the best result was obtained with a combined analysis, and not separating n-H and n-Gd cancels the spill effects at the Target/γ-Catcher threshold.

6.4.a Scintillator Leak in the Near Detector : Buffer Events

During the commissioning of the Near Detector, tests were performed to evaluate the amount of Light Noise. During these tests, it was found that some events at the top of the detector were reconstructed outside of the γ-Catcher, in the Buffer. These events show patterns that correspond to the position of the PMTs, as shown in Figure 6.11, thus ruling out the possibility of them being simply mis-reconstructed.
Upon more dedicated testing, it was found that the top of the buffer was more opaque to the calibration LED light than expected.

This behaviour is explained by a leak of γ-Catcher scintillator into the buffer. The contamination is small, such that enough light to pass the trigger thresholds is being collected only for events close to a PMT, explaining why these events only appear close to PMTs.

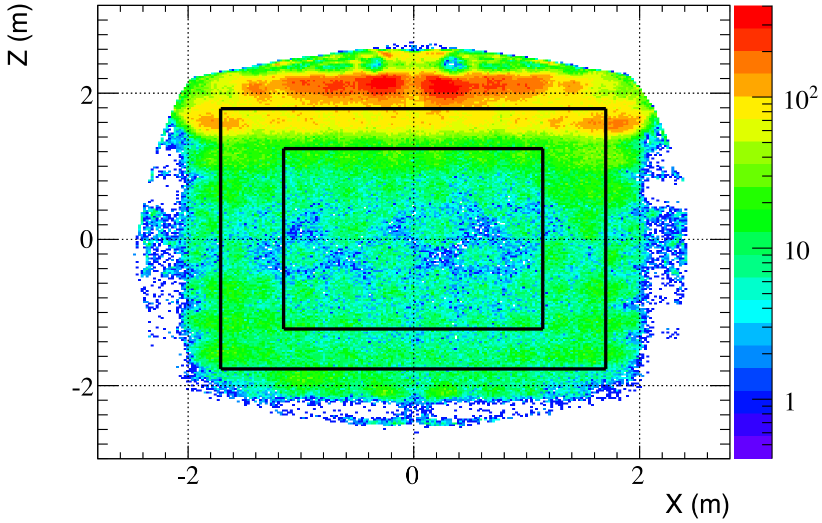


Figure 6.11: Some events are reconstructed largely outside of the active volumes (the Target and γ-Catcher are shown in black lines). They show a pattern corresponding to the position of the PMT in the Buffer, which means they happened close to the PMT, and do not result in a mis-reconstruction.

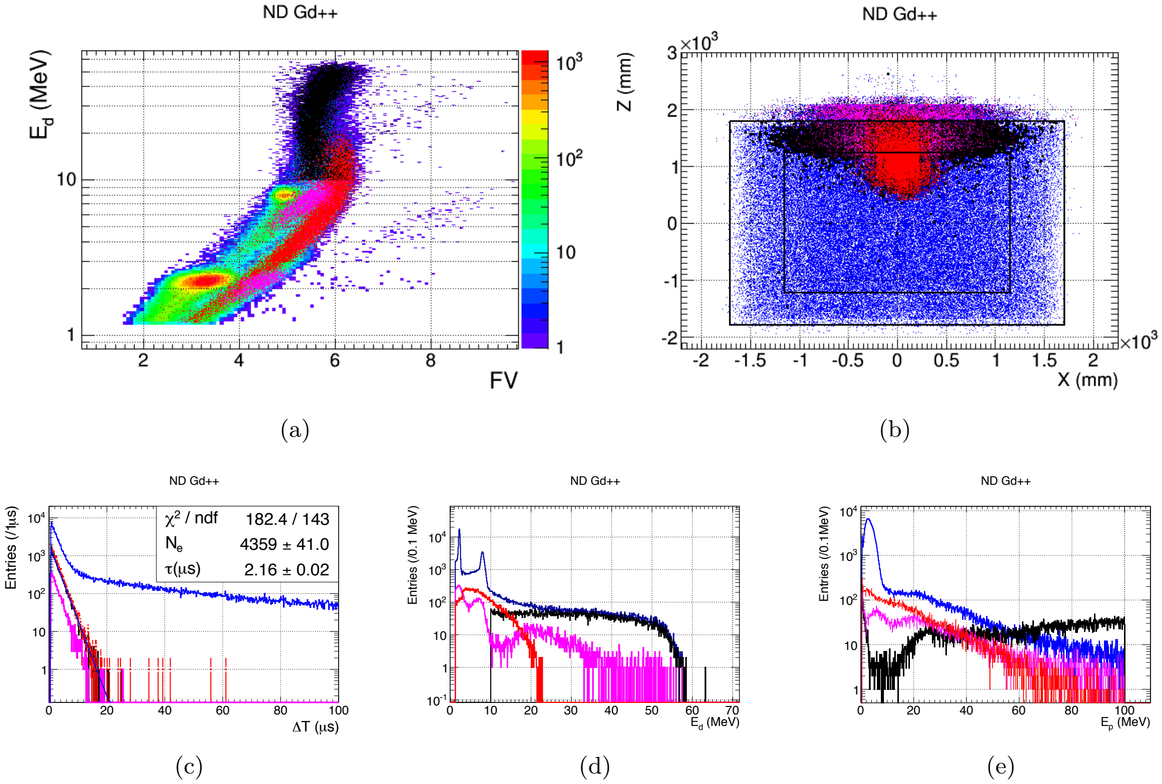
Due to a smaller overburden than the Far Detector (120 m.w.e. instead of 300 m.w.e), the Near Detector has a higher muon rate than the Far Detector by a factor of ~ 5. The scintillator leak into the Buffer increases the sensitivity to outside events, i.e. Fast Neutrons and Stopping Muons dominantly, increasing even further the magnitude of the Stopping Muons contamination in the Near Detector. Such Stopping Muons are called "Buffer Stopping Muons" (Buffer SM).

6.4.b DC-IV Functional Value

Figure 6.12 shows Distributions of Stopping Muons in the Near Detector Gd++ sample, for the Functional Value distribution (6.12(a)), the spatial distribution of the delayed events. (6.12(b)), the ΔT distribution (6.12(c)), the delayed and prompt energy (6.12(e) and 6.12(d)). Superimposed on the candidate distributions (in blue) are shown the various contributions of Stopping Muons:

- The Chimney Stopping Muons, the ones present in the DC-III analysis, shown in red. They are spatially mis-reconstructed and appear below the chimney.
- The Buffer SM (in pink) appear because of the leak of scintillator in the buffer. They are reconstructed right above the γ-Catcher.
- Opening the prompt and delayed energy window up to 100 MeV allowed for a better handle on Fast Neutrons. However, it also enabled Stopping Muons to travel further before decaying, while still not triggering the Muon Veto. Such muon, and the subsequent Michel electron are no longer restrained to the Chimney, and can be located in the γ-Catcher, therefore, their Functional Value is lower than that of a usual Stopping Muon.

This population makes up the black distribution in the γ-Catcher, and will be referred to as GC-SM.



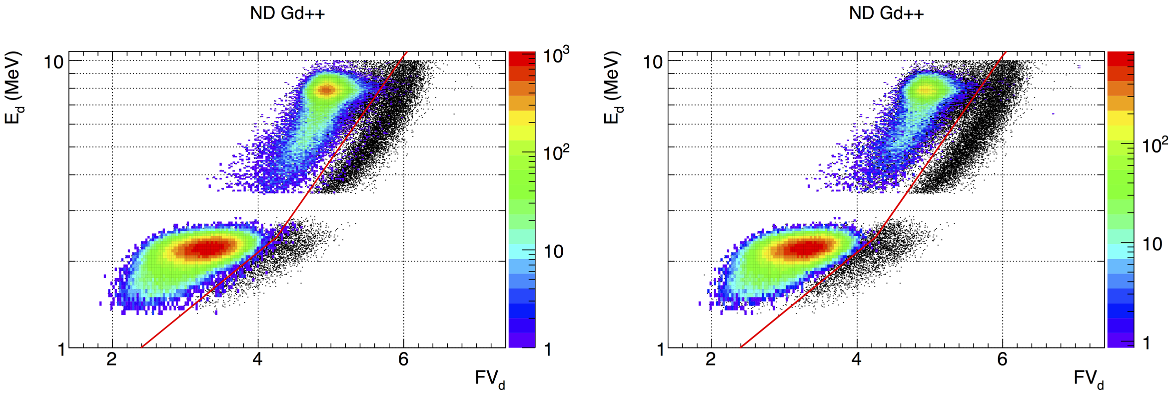
(a) (b)

(c) (d) (e)

Figure 6.12: Description of the different populations of Stopping Muons in the Near Detector.
6.12(a): E_d vs FV_d distribution. All candidates are shown (color), GC-SM (black), Chimney SM (red), and Buffer SM(pink).
6.12(b): Delayed vertex position: The blue points correspond to all the candidates and the coloured distributions are each SM contributions.
6.12(c): ΔT distributions: The blue line corresponds to all the candidates and the coloured distributions are each SM contributions. SM contributions show a 2.2 µs life-time fully compatible with Stopping Muons.
6.12(d): Delayed energy distribution for all candidates (blue) and superimposed SM contributions.
6.12(e): Prompt energy distribution for DC-IV selected candidates (blue), and raw SM contributions.

Due to multiple Stopping Muons contributions, and the fact that the Buffer SM distribution overlaps the n-Gd peak, it is not possible to reduce fully the SM with a simple FV cut as in DC-III. While a looser FV-cut is kept, as shown in Figure 6.13, other variables have been investigated in order to improve the SM rejection power.

Figure 6.13 shows the FV cut for the ND, Gd++. The black dots correspond to the default sample, after application of all non-SM vetoes. In Figure 6.13(a) is superimposed the "N-1" sample, after application of all other SM-vetoes except the FV cut (cuts on the other variables investigated).



(a)

(b)

Figure 6.13: Definition of the FV cut for ND Gd++. The black dots correspond to the E_d vs FV_d distribution for all candidates after application of all non-SM vetoes. Superimposed is the N-1 sample (6.13(a)), and the control sample with $\Delta T > 30\,\mu\text{s}$ (6.13(b)).

A SM-free sample can be defined by a simple cut on ΔT. Because SM have a very short correlation time, imposing $\Delta T > 30\,\mu\text{s}$ reduces completely the SM contamination. This cut is not used in the θ_{13} analysis because of its very high inefficiency. Indeed, the maximum of the n-Gd peak is before $30\,\mu\text{s}$, leading to a $\sim 60\%$ inefficiency in the Target, and a $\sim 20\%$ in the γ-Catcher.
However, this cut is very useful in providing a SM-free sample to estimate the inefficiency of a cut.
Figure 6.13(b) shows the default sample in black dots, and the control sample is superimposed in color. This allows to visualize the sanity of the cut position with respect to the pure IBD-like sample.

The FV cut consists in two exponentials, corresponding to each neutron capture. In the case of Gd-only, the same cut is applied for consistency, but the n-H part becomes irrelevant. The cut is defined as:

$$E_d > \max\left(A_{\rm H}\cdot e^{FV_d/B_{\rm H}}, A_{\rm Gd}\cdot e^{FV_d/B_{\rm Gd}}\right) \tag{6.14}$$

Table 6.1: Summary of FV-cuts for the various configurations of DC-IV.

Detector	FV-Cut
FD-I	$E_d > \max\left(0.36\cdot e^{FV_d/2.4}, 0.06\cdot e^{FV_d/1.2}\right)$
FD-II	$E_d > \max\left(0.20\cdot e^{FV_d/1.8}, 0.05\cdot e^{FV_d/1.2}\right)$
ND	$E_d > \max\left(0.32\cdot e^{FV_d/2.1}, 0.07\cdot e^{FV_d/1.2}\right)$

6.4.c A New Functional Value: FV*

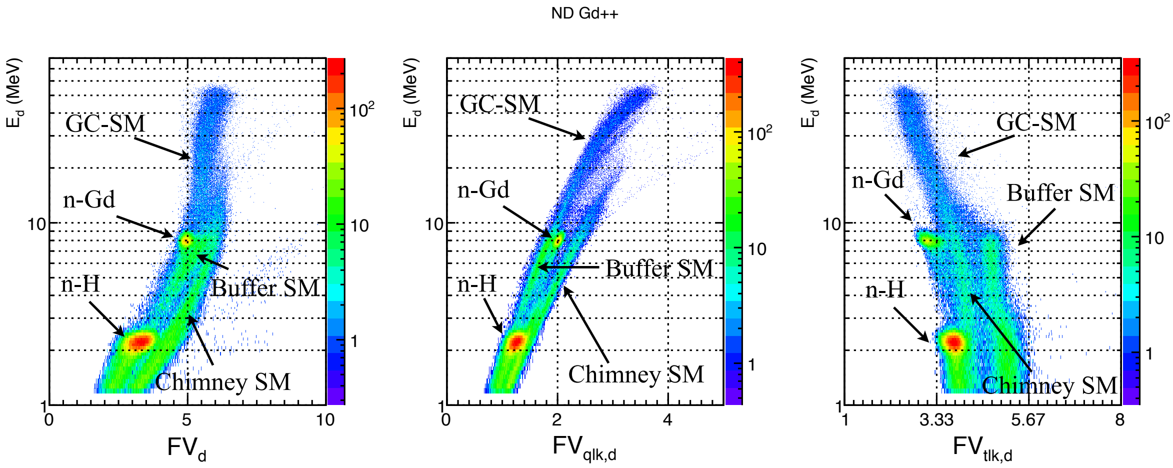


Figure 6.14: Break down of the Functional Value (left) into its charge-only (center) and time-only (right) components. The various Stopping Muons contributions are better separated in the $\mathrm{FV}_{\mathrm{tlk,d}}$. The left-most population is the Buffer SM population.

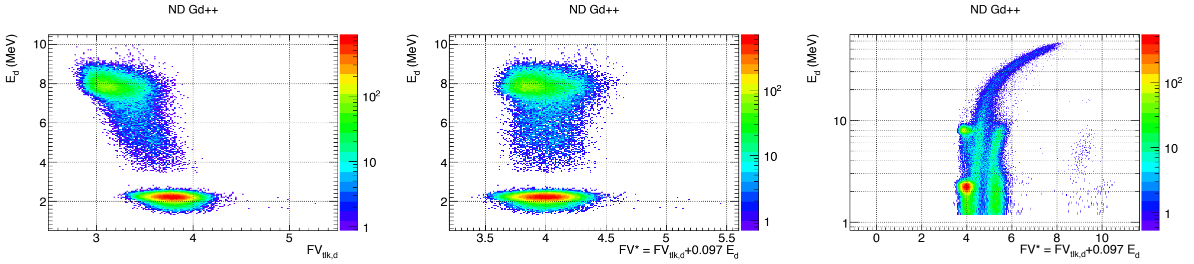
The Functional Value can be broken down in two separate quantities : $\mathrm{FV}_{\mathrm{tlk}}$, standing for Functional Value Time Likelihood, using only the hit time information in RecoBAMA, and $\mathrm{FV}_{\mathrm{qlk}}$, for Functional Value Charge Likelihood, using only the charge information. The three variables are shown in Figure 6.14. The time-only component ($\mathrm{FV}_{\mathrm{tlk}}$) seems to be the best choice to separate Buffer Stopping Muons.

$\mathrm{FV}_{\mathrm{tlk}}$ is not independent from energy. The more energy an event deposits, the more photons are emitted, and the more PMTs are hit, which in turns increases the reconstruction and lowers the negative log-likelihood. Therefore, two separate conditions are required for the n-H and n-Gd captures, which is not satisfactory. To correct for this dependency with energy, a new variable, called FV*, is defined as a linear combination of $\mathrm{FV}_{\mathrm{tlk,d}}$ and E_d. The definitions of FV* for each detector configuration are summarized in Table 6.2.

Table 6.2: Summary of FV* definition for the various configurations of DC-IV.

Detector	FV*
FD-I	$\mathrm{FV}_{\mathrm{tlk,d}} + 0.110 \cdot E_d$
FD-II	$\mathrm{FV}_{\mathrm{tlk,d}} + 0.120 \cdot E_d$
ND	$\mathrm{FV}_{\mathrm{tlk,d}} + 0.097 \cdot E_d$

The E_d vs $\mathrm{FV}_{\mathrm{tlk,d}}$, and the corrected E_d vs FV* distributions are shown in Figures 6.15(a) and 6.15(b) respectively. Figure 6.15(c) shows the distribution with all candidates, where a straight cut in FV* could be envisioned, but with poor rejection power on the Chimney SM. However, Buffer SM and GC-SM are very well separated from the IBD sample.



(a) (b) (c)

Figure 6.15: 6.15(a) and 6.15(b): Defintion of $\mathrm{FV}^\star$ as a linear combination of $\mathrm{FV}_{\mathrm{tlk,d}}$ and E_d. to correct for the dependency of $\mathrm{FV}_{\mathrm{tlk,d}}$ on E_d in the SM-free sample.
6.15(c): E_d vs $\mathrm{FV}^\star$ for all candidates.

6.4.d Discrepancies in the Vertex Reconstruction

In parallel to RecoBAMA, another vertex reconstruction has been developed by the Niigata group, called RecoJapan. Figures 6.16(a) and 6.16(b) show comparisons of RecoBAMA and RecoJapan for the vertical and radial reconstructions. The RecoJapan reconstruction uses different pull term strategies and does not force vertices inside the detector, hence some differences arise with respect to RecoBAMA.

Figure 6.16(c) shows the evolution of the absolute difference $\Delta\mathrm{BJ}_{z,d} = Z_{Japan} - Z_{BAMA}$ with the position of the vertex as reconstructed by RecoBAMA, here for delayed triggers. In the majority of cases, the two algorithms agree within ±0.5 m, however, populations of events at the top of the detector, that are not present in the SM-free events (in color), have bigger discrepancies. They correspond to Chimney Stopping Muons, for Z between ~ 0.5 m and ~ 1.5 m, and to Buffer Stopping Muons, for Z higher than ~ 1.75 m. A cut in both $\Delta\mathrm{BJ}_{z,d}$ and Z would help reducing these contribution, however, because of the relatively poor spatial resolution of RecoBAMA, it was decided in the collaboration not to use any cut based on the absolute position of the events.

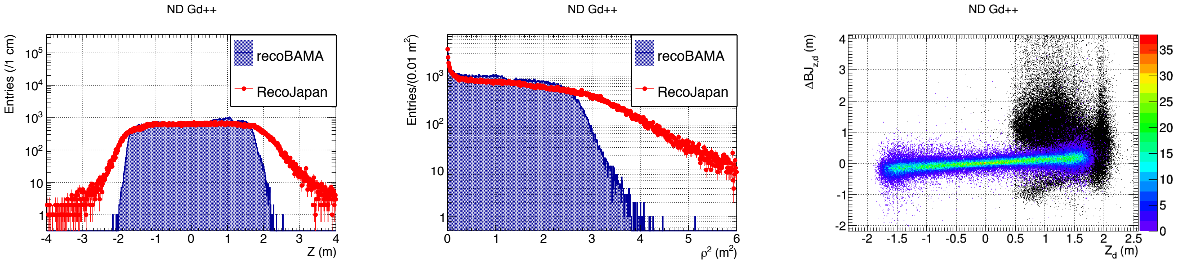
$\Delta\mathrm{BJ}_{z,d}$ is a powerful way to separate Stopping Muons from IBD, however, $\Delta\mathrm{BJ}_{z,d}$ alone is not sufficient to remove all Stopping Muons.

6.4.e Combining ΔBJ and FV⋆

As previously explained, neither $\mathrm{FV}^\star$ nor $\Delta\mathrm{BJ}_{z,d}$ are sufficient to reject Stopping Muons on their own. However, both variables allow a good degree of separation between IBD and different types of Stopping Muons. $\mathrm{FV}^\star$ is good at separating Buffer SM and GC-SM while $\Delta\mathrm{BJ}_{z,d}$ is good at rejecting Chimney SM.

Combining both variables, as shown in Figure 6.17(a), allows a clearer separations of each population. Figure 6.17(b) shows the individual contributions of each population:

- Chimney SM (red)
- Buffer SM (pink)



(a) (b) (c)

Figure 6.16: 6.16(a) and 6.16(b): Z and ρ^2 vertex distributions as reconstructed by RecoBAMA (blue) and RecoJapan (red).
6.16(c): Evolution of the difference between RecoJapan and RecoBAMA Z-axis reconstructions of delayed events ($\Delta\mathrm{BJ}_{z,d}$) with the Z position of the vertices (by RecoBAMA). The black dots correspond to all candidates, and the color histogram shows the SM-free sample with $\Delta T > 30\,\mu\mathrm{s}$.

- GC-SM (black)
- SM-free sample with $\Delta T > 30\,\mu\mathrm{s}$ (blue)

The Chimney SM population separation from FV* is increased by the breaking of degeneracy from $\Delta\mathrm{BJ}_{z,d}$, and FV* drives the separation of Buffer SM and GC-SM.

Figure 6.17(c) shows the N-1 distributions, when applying all non-SM vetoes (black dots), and when applying all vetoes except the one in the $\Delta\mathrm{BJ}_{z,d}$vsFV* (color), wich is materialized by the red line. It is to be noted that the Chimney SM disappear already when applying all other vetoes, leaving the Buffer SM to be rejected. Figure 6.17(d) shows the SM-free sample, and the red line materializes the cut in the $\Delta\mathrm{BJ}_{z,d}$ vs FV* space.

The cut used here, called later on ΔBJ-cut, is an ellipse fitted onto top right edge of the IBD region, i.e. the border between IBD and Chimney SM. Events outside the ellipse are rejected. As seen in Figure 6.17(d), the IBD distribution leaks at high absolute values of $\Delta\mathrm{BJ}_{z,d}$. To account for that behaviour, a straight line is added, to protect this population of leaking IBDs. Finally, all events within the ellipse or at the left side of the line are kept as IBD candidates.

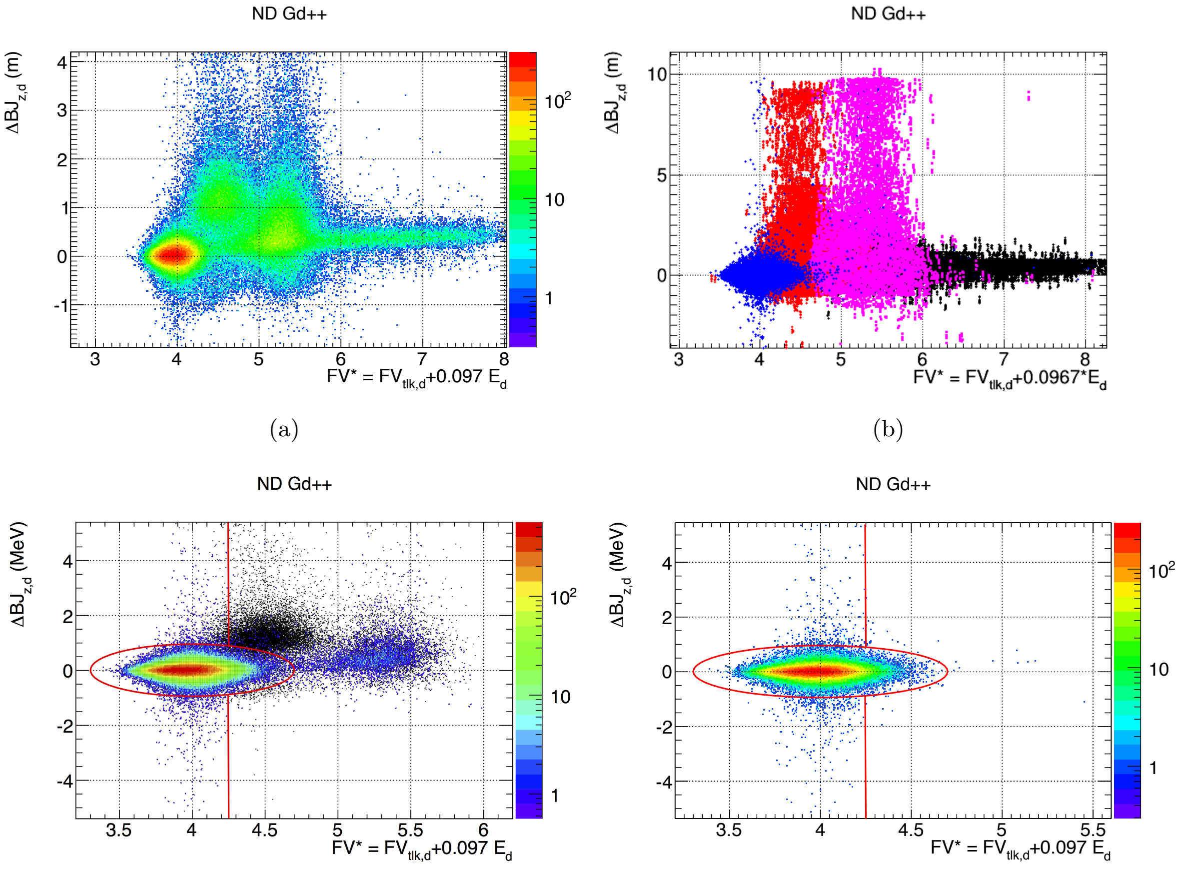
The description of the cut needs five parameters for the ellipse, and one for the vertical protection:

$$\left(\frac{\mathrm{FV}^\star - \mathrm{FV}^\star_0}{\sigma_{\mathrm{FV}^\star}}\right)^2 + \left(\frac{\Delta\mathrm{BJ}_{z,d} - \Delta\mathrm{BJ}_{z,0}}{\sigma_{\Delta\mathrm{BJ}_{z,d}}}\right)^2 < R^2 \parallel \mathrm{FV}^\star < \mathrm{FV}^\star_{\mathrm{protect}} \tag{6.15}$$

The parameters for the analyses of each Detector and capture configuration are summarized in Table 6.3.

6.4.f Chimney Pulse-Shape

A SM discrimination can be achieved on pulse-shape criteria. First the waveforms from all the PMT for a given event are shifted to correct for the time of flight assuming the vertex



(a) (b)

(c) (d)

Figure 6.17: $\Delta \mathrm{BJ}_{z,d}$ vs FV* distributions:
6.17(a): $\Delta \mathrm{BJ}_{z,d}$vsFV* distribution for all candidates.
6.17(b): Contributions of each SM populations: Chimney SM (red), Buffer SM (pink) and GC-SM (black). The SM-free sample is shown in blue.
6.17(c): N-1 distribution, when applying all non-SM related vetoes (black dots), and when applying all cuts except the cut in $\Delta \mathrm{BJ}_{z,d}$ vs FV*. The red line materializes the $\Delta \mathrm{BJ}_{z,d}$ vs FV* cut.
6.17(d): SM-free sample with $\Delta T > 30\,\mu\mathrm{s}$. The red line materializes the $\Delta \mathrm{BJ}_{z,d}$ vs FV* cut.

position reconstructed by RecoBAMA: $\vec{r}_{\mathrm{BAMA}}$. A likelihood $\mathcal{L}_{\mathrm{vtx}}$ is computed using the average pulse shape of a Co calibration run as reference. Then the waveforms are time-of-flight corrected assuming that the vertex is located at the entrance of the Chimney: $\vec{r}_{\mathrm{chm}} = (0, 0, 250\,\mathrm{cm})$ and a new likelihood $\mathcal{L}_{\mathrm{chm}}$ is computed. $\mathcal{L}_{\mathrm{vtx}}$ and $\mathcal{L}_{\mathrm{chm}}$ are defined as:

$$\mathcal{L}_{\mathrm{vtx}}(\vec{r} = \vec{r}_{\mathrm{BAMA}}) = \prod_{i=1}^{128} Co_i(\vec{r}_{\mathrm{BAMA}}) \times WF_i(\vec{r}_{\mathrm{BAMA}}) \tag{6.16}$$

$$\mathcal{L}_{\mathrm{chm}}(\vec{r} = \vec{r}_{\mathrm{chm}}) = \prod_{i=1}^{128} Co_i(\vec{r}_{\mathrm{chm}}) \times WF_i(\vec{r}_{\mathrm{chm}}) \tag{6.17}$$

Table 6.3: ΔBJ cut parameters summary.

Configuration		$\mathrm{FV}_0^\star$	$\Delta\mathrm{BJ}_{z,0}$	$\sigma_{\mathrm{FV}^\star}$	$\sigma_{\Delta\mathrm{BJ}_{z,d}}$	R	$\mathrm{FV}^\star_{\mathrm{protect}}$
Gd	FD-I	3.9	-0.03	0.16	0.14	6	4.3
	FD-II	3.9	-0.03	0.17	0.14	6	4.3
	ND	3.9	0.01	0.15	0.13	5	4.3
Gd++	FD-I	4.1	-0.03	0.14	0.21	4	4.3
	FD-II	4.1	0.01	0.15	0.21	5	4.3
	ND	4.0	0.01	0.14	0.19	5	4.3

Where $WF_j(\vec{r})$ is the value of the event pulse shape at the time sample j, assuming a vertex position $\vec{r}$, and $Co_j(\vec{r})$ corresponds to the value of the Co pulse shape at the time sample j assuming a position $\vec{r}$. The ratio $\ln(\mathcal{L}_{\mathrm{chm}})/\ln(\mathcal{L}_{\mathrm{vtx}})$ defines the CPS variable.

Figure 6.18(a) shows the CPS distribution as a function of the distance of the vertex (estimated by RecoBAMA) from the Chimney for all candidates (blue dots). Each individual SM contributions is shown: The Chimney Stopping Muons (red dots), the Buffer Stopping Muons (pink dots), and the GC-SM (black dots). It is interesting to note that the CPS is constructed to reject Chimney events, and as expected, is well separating the Chimney SM from the IBD candidates. However, it is not so powerful in rejecting Buffer SM and GC-SM as the delayed triggers of those events do not occur in the Chimney.

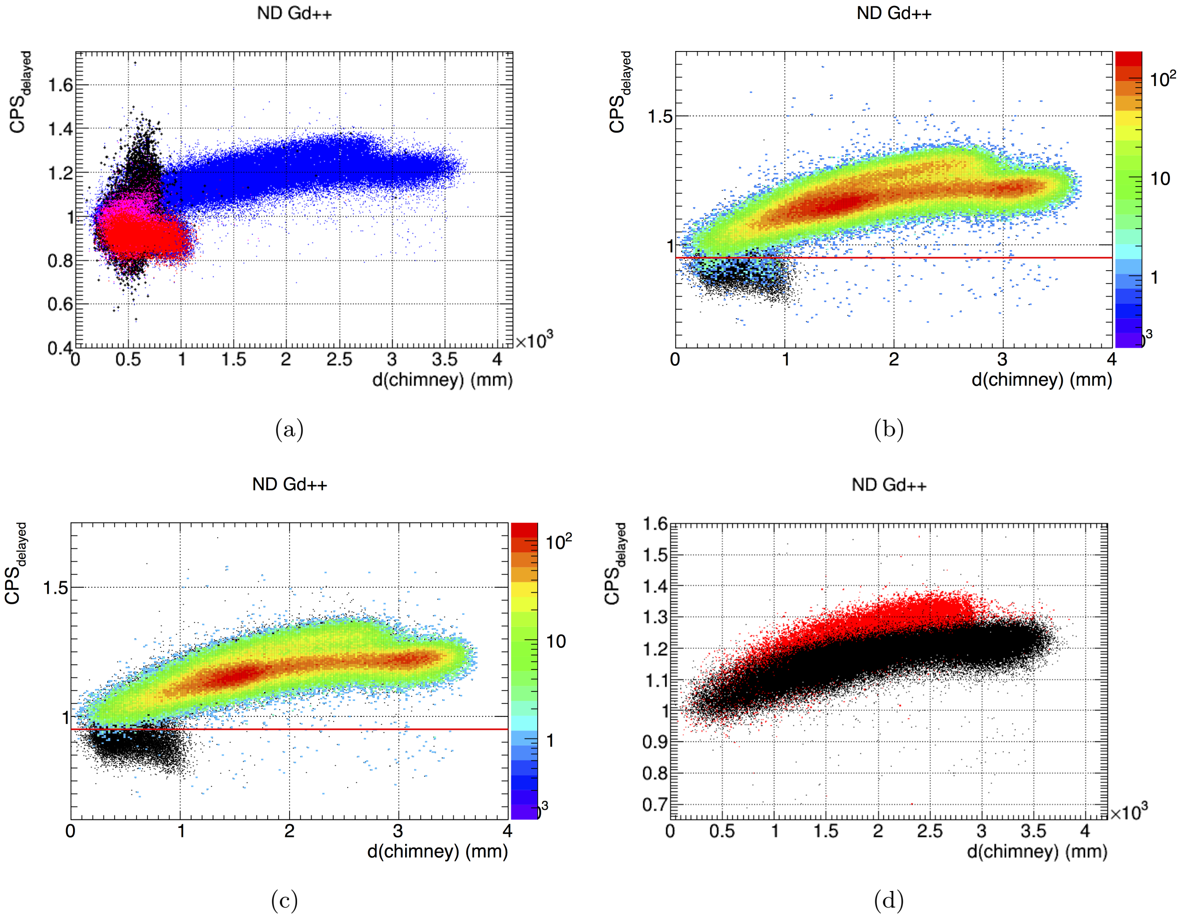
Figure 6.18(b) and 6.18(c) show the behaviour of the CPS cut, parametrized as a 1D cut on the CPS value of the delayed event, with respect to the "N-1" sample, i.e. the candidates on which all other cut except the CPS cut are applied, and the SM-free sample with $\Delta T > 30\,\mu\mathrm{s}$. The IBD-like distribution shows two lobes as shown in Figure 6.18(d), corresponding to the n-H (black) and n-Gd (red) captures.

The cut used here, called later on CPS-cut, is $\mathrm{CPS}_d > 0.95$ for all captures and configurations.

6.5 Performance Evaluation

6.5.a Stopping Muon Rejection Power

The Stopping Muon rejection power is the ratio $N_{\mathrm{SM}}^{\mathrm{rej}}/N_{\mathrm{SM}}^{\mathrm{tot}}$. Two methods are used to estimate this ratio.
The first method is to compare the number of rejected candidates in a Muon-enriched sample to the total number of candidates in this sample.



(a)

(b)

(c)

(d)

Figure 6.18: Definition of the CPS cut for ND Gd++.
6.18(a): Individual SM contributions overlaid on all candidates after application of the non-SM vetoes (blue dots). Chimney Stopping Muons (red), Buffer Stopping Muons (pink) and GC-SM (black).
6.18(b): Remaining candidates after application of all vetoes except the CPS cut (color). The black dots correspond to the candidates that passed the non-SM vetoes.
6.18(c): Candidates of the SM-free sample ($\Delta T > 30\,\mu\text{s}$) in color. The black dots correspond to the candidates that passed the non-SM vetoes.
6.18(d): CPS distributions for captures on n-Gd (red) and n-H (black).

The Muon-enriched sample is selected with the following cuts.

$$\begin{aligned}\text{SM-enriched:}\ & ((\text{CPS}_p + \text{CPS}_d < 1.9) \cap (\Delta T < 20\,\mu\text{s}) \cap (E_d < 0.16 \times \exp(\text{FV}_d/1.5)))\\ \text{rejection Power:}\ & \frac{N((\text{SM-enriched}) \cap \overline{(\text{SM-cuts})})}{N(\text{SM-enriched})} \qquad (6.18)\end{aligned}$$

The sum of the CPS variables for prompt and delayed triggers, shown in Figure 4.20, has been found to be very efficient to tag Stopping Muons and was used in the results presented at the Moriond 2016 conference. However, it was decided to avoid using it since it uses the information of the prompt trigger, which could bias the prompt energy spectrum.

The ΔT distribution of the Muon-enriched sample is shown in Figure 6.19(a). The black histogram corresponds to the full SM-enriched sample, the cut $\Delta T > 20\,\mu\text{s}$ is left out in the plot to allow an estimation of the negligible IBD-like contamination. The red histogram corresponds to the events from this sample rejected by the SM cuts, and the blue histogram corresponds to those that pass the SM-cuts.

The second method uses a fit of the correlation time model from Equation 6.6 to the full sample before and after application of the SM-cuts in order to estimate the reduction of the Stopping Muons. The best fit of the IBD candidates before and after applications of the SM cuts is shown in Figure 6.19(b). The black line corresponds to the IBD sample before SM-cuts, the best fit parameters are shown in the black box. The blue histogram corresponds to the selected sample, after application of the SM-cuts, the best fit parameters are shown in the blue box. Rejection powers estimated from these two methods for the Near Detector, Gd + + are:

$$\begin{aligned}\epsilon_{\text{count}} &= (99.73^{+0.27}_{-1.46})\% \\ \epsilon_{\text{Fit}} &= (96.83 \pm 1.86)\%\end{aligned}$$

The fit of the IBD selection before Stopping Muon rejection allows to estimate their original contamination of the IBD sample to $\sim 15\%$ in the Near Detector Gd++ configuration. Given such a level of contamination, it is indeed un-realistic to perform an analysis without Stopping Muons rejection.

Both methods give consistent results within 1σ. The fitting method only provides cross-check as it is not sensitive enough to the low contamination remaining in the sample, even more so for the Far Detector, where the Stopping Muon contamination is 20 times smaller, as shown in Figure 6.19(c), where the Stopping Muon contamination is too low to be precisely estimated.

The rejection powers estimated with the first method for each detector and capture configurations are summarized in Table 6.4. The $\Delta\text{BJ}_{z,d}$ cut is comparatively much more efficient for the Near Detector than for the Far Detector because of the presence of Buffer SM only in the Near Detector, that other cuts cannot address as efficiently. For the Far Detector, the FV-cut is the dominant cut.

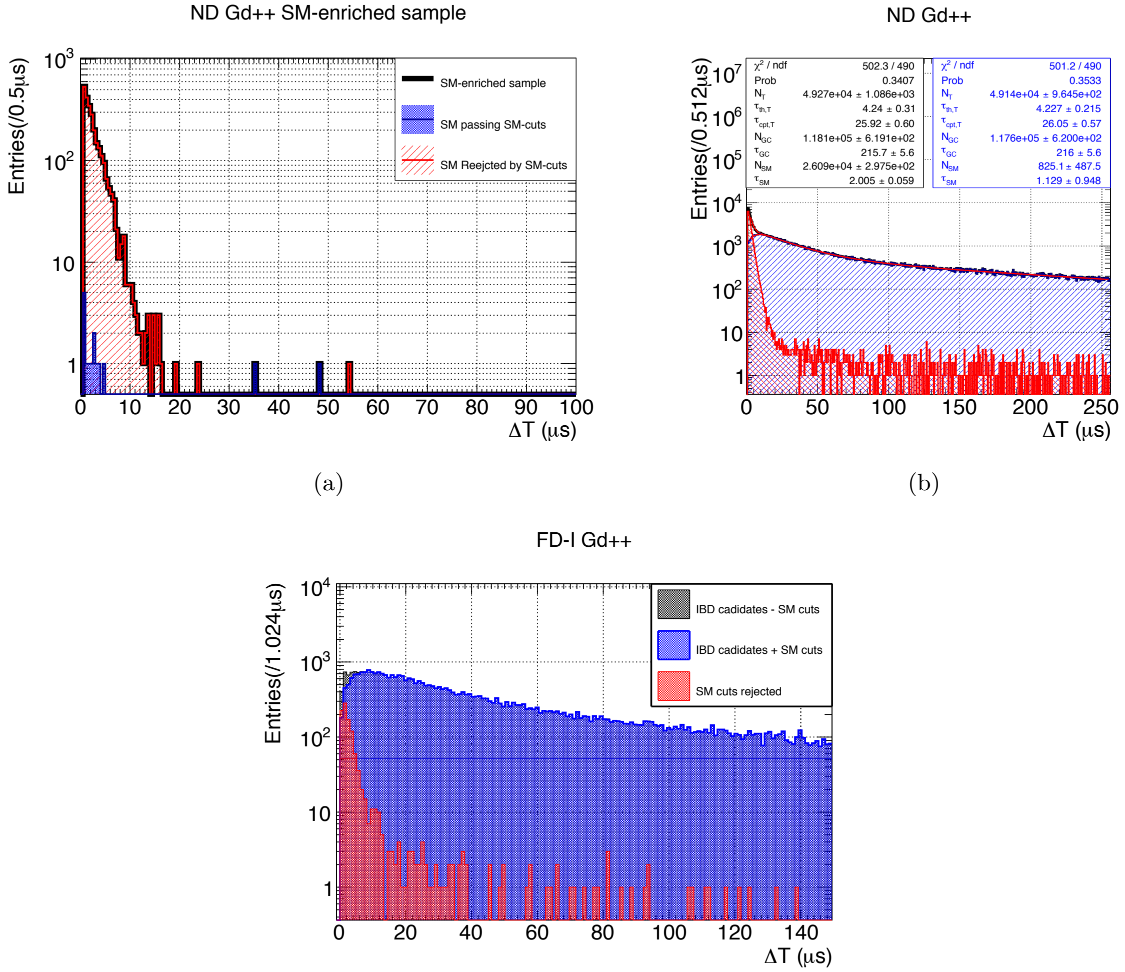
6.5.b IBD Inefficiency Estimation

The full SM cut is a concatenation of three cuts:

$$\text{SM cut} = \Delta\text{BJ}_{z,d}\text{-cut} \cap \text{FV-cut} \cap \text{CPS-cut} \tag{6.19}$$

For a given cut, the inefficiency of the cut is computed as the ratio of the estimated number of IBD rejected by the cut over the total estimated number of IBD in the original sample: $N^{\text{rej}}_{\text{IBD}}/N^{\text{tot}}_{\text{IBD}}$. As in the Functional Value analysis for DC-III presented in Section 6.3, an estimation of remaining backgrounds in the total and rejected sample must be carried out in order to obtain an unbiased inefficiency.

The contamination of Accidentals is reduced by applying the ANN-cut, the ^{9}Li is reduced by applying the cut on $\mathcal{L}_{^9\text{Li}}$ and the Fast Neutron contribution is reduced by applying the Inner Veto cuts. Moreover, the inefficiency is estimated on a SM-free sample, with $\Delta T > 30\,\mu\text{s}$, so as not to count SM rejection as IBD rejection. All events with $E_d > 10\,\text{MeV}$ are either Stopping



(a)

(b)

(c)

Figure 6.19: Two rejection power estimation methods have been developed:
6.19(a): Counting events before and after application of the SM cuts on a Stopping Muon enriched sample. Here the cut $\Delta T < 20\,\mu s$ have been left out to visualize the negligible IBD-like contamination of this sample.
6.19(b): Fitting the ΔT distribution of the full candidates with the correlation time model described in Equation 6.6, before (black line) and after application of the SM cuts. The selected sample is shown in blue, and the rejected sample in red. The red lines show the best fit result before and after application of the SM-cut. The parameters results in each case are shown in black (before cuts) and blue (after cut).
6.19(c): Limitations for the fit method arise for the FD where the Stopping Muon contamination is to low to be well estimated.

Muons or Fast Neutrons, and can therefore be rejected a priori. The sample of events can then be described as:

$$\begin{aligned}
\text{FN} &= (\text{ANN-cut} \cap (\mathcal{L}_{^9\text{Li}} < 0.4) \cap (\Delta T > 30\,\mu\text{s}) \cap \text{IV-Tag}) \\
\text{IBD}^{\text{tot}} &= (\text{ANN-cut} \cap (\mathcal{L}_{^9\text{Li}} < 0.4) \cap (\Delta T > 30\,\mu\text{s}) \cap \text{IV-cut} \cap (E_d < 10\,\text{MeV})) \qquad (6.20) \\
\text{IBD}^{\text{rej}} &= (\text{ANN-cut} \cap (\mathcal{L}_{^9\text{Li}} < 0.4) \cap (\Delta T > 30\,\mu\text{s}) \cap \text{IV-cut} \cap (E_d < 10\,\text{MeV}) \cap \cancel{\text{Cut}})
\end{aligned}$$

$$\begin{aligned}
\text{let}: \text{FN}^{\text{tot}} &= \text{FN scaled to IBD}^{\text{tot}} \text{ on } [15-100]\text{MeV} \\
\text{and}: \text{FN}^{\text{rej}} &= \text{FN scaled to IBD}^{\text{rej}} \text{ on } [15-100]\text{MeV} \\
\text{then}: \bar{\epsilon} &= \frac{\text{IBD}^{\text{rej}} - \text{FN}^{\text{rej}}}{\text{IBD}^{\text{tot}} - \text{FN}^{\text{tot}}} \qquad (6.21)
\end{aligned}$$

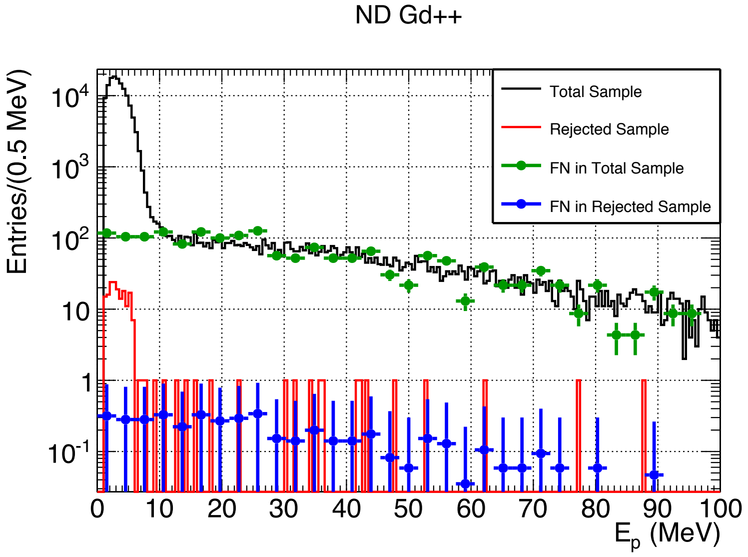


Figure 6.20: Prompt spectrum of events before applying SM-cuts (black), and those rejected by SM-cuts (red). In green and blue are shown the shape of Fast Neutrons, normalized over $[15-100]$MeV to respectively the total events, and the rejected ones. The integral of these Fast Neutrons spectra gives the Fast Neutron contamination in both samples.

Figure 6.20 shows the spectra of the IBD-like events that pass all default cuts as defined above in black, and the ones that pass all default cut except the considered cut in red (here, the full SM-cut). The green and blue histograms show the Fast Neutron contamination in each sample, obtained by scaling the IV-Tagged Fast Neutron shape on the $[15-100]$MeV range of each sample. The subtraction of the Fast Neutron histograms from the IBD-like histograms allows to estimate the pure IBD inefficiency.

Inefficiencies of each individual SM cut and of the full SM cut are computed as defined in Equation 6.21, and are summarized in Table 6.4.

Table 6.4: Inefficiency and Rejection power estimation for SM cuts for each detector and capture configuration.

FD-I Gd++

veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.13 \pm 0.02)\%$	$(90.6 \pm 4.5)\%$
FV	$(0.049 \pm 0.012)\%$	$(94.22 \pm 4.64)\%$
CPS	$(0.011 \pm 0.007)\%$	$(93.8 \pm 4.6)\%$
Total	$(0.166 \pm 0.024)\%$	$(99.88^{+0.12}_{-4.85})\%$
Remaining	$< 0.2\%$ of IBD	

FD-I Gd

veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.010 \pm 0.0094)\%$	$(97.0^{+3.0}_{-5.7})\%$
FV	$(0.056 \pm 0.023)\%$	$(99.6^{+0.4}_{5.8})\%$
CPS	$(0.007^{+0.01}_{-0.007})\%$	$(97.6^{+2.4}_{-5.8})\%$
Total	$(0.06^{+0.025}_{-0.028})\%$	$(100^{+0}_{-5.8})\%$
Remaining	$< 0.05\%$ of IBD	

FD-II Gd++

veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.104 \pm 0.020)\%$	$(83.3 \pm 4.8)\%$
FV	$(0.083 \pm 0.018)\%$	$(93.24 \pm 5.2)\%$
CPS	$(0.153 \pm 0.024)\%$	$(91.0 \pm 5.1)\%$
Total	$(0.32 \pm 0.035)\%$	$(99.25^{+0.75}_{-5.5})\%$
Remaining	$< 0.2\%$ of IBD	

FD-II Gd

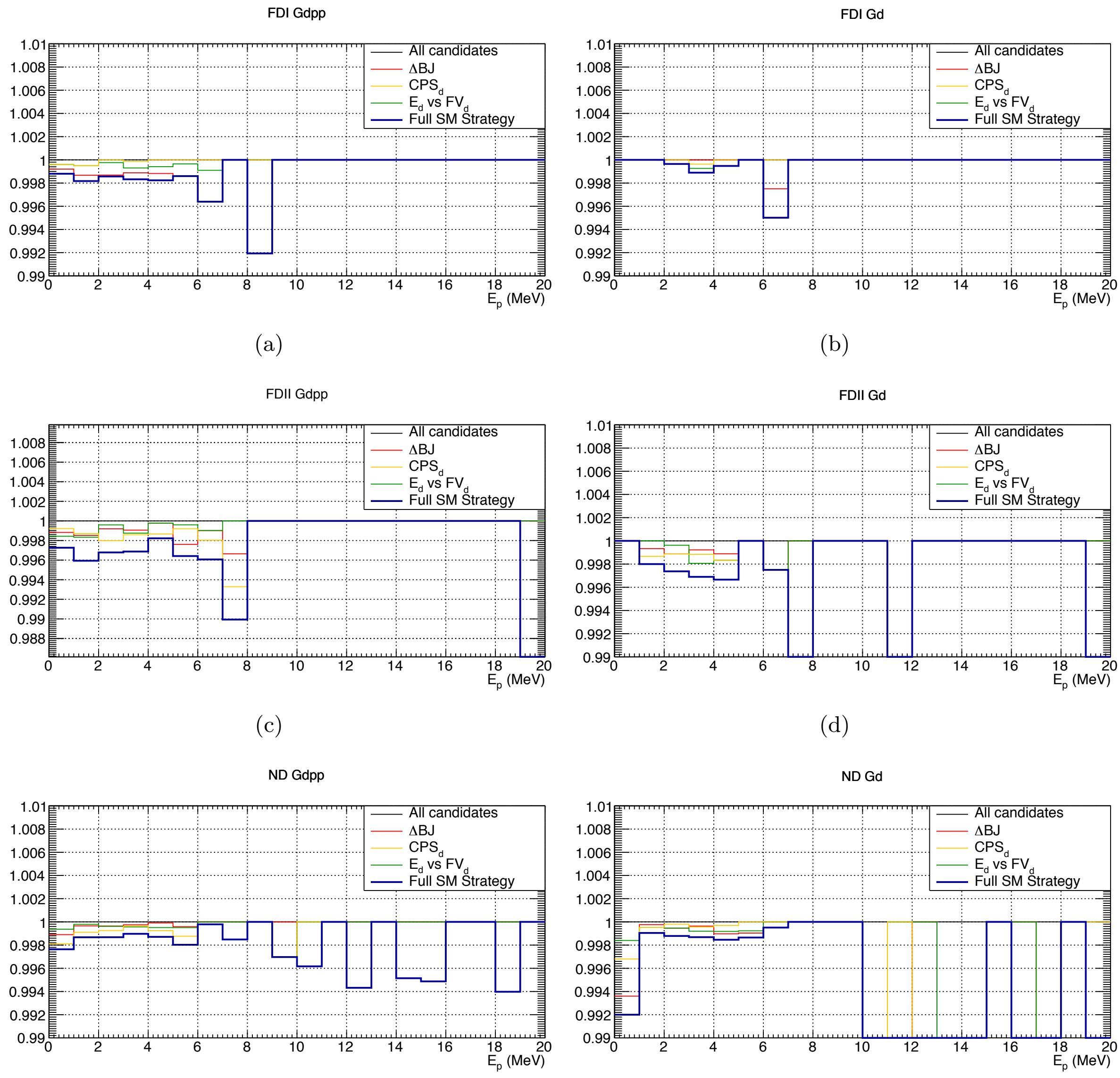
veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.089^{+0.029}_{-0.032})\%$	$(96.4 \pm 7.0)\%$
FV	$(0.10 \pm 0.03)\%$	$(95.89 \pm 7.0)\%$
CPS	$(0.14^{+0.037}_{-0.039})\%$	$(96.14 \pm 7.0)\%$
Total	$(0.26 \pm 0.05)\%$	$(100^{+0}_{-7.0})\%$
Remaining	$< 0.05\%$ of IBD	

ND Gd++

veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.028 \pm 0.005)\%$	$(92.44 \pm 1.77)\%$
FV	$(0.037 \pm 0.006)\%$	$(58.74 \pm 1.67)\%$
CPS	$(0.070 \pm 0.008)\%$	$(49.56 \pm 1.67)\%$
Total	$(0.125 \pm 0.010)\%$	$(99.73^{+0.27}_{-1.46})\%$
Remaining	$< 0.05\%$ of IBD	

ND Gd

veto	Inefficiency	Rejection Power
$\Delta \mathrm{BJ}_{z,d}$	$(0.062 \pm 0.011)\%$	$(99.6^{+0.4}_{-2.9})\%$
FV	$(0.068 \pm 0.012)\%$	$(87.1 \pm 2.8)\%$
CPS	$(0.035 \pm 0.009)\%$	$(83.1 \pm 2.8)\%$
Total	$(0.129 \pm 0.016)\%$	$(100^{+0}_{2.8})\%$
Remaining	$< 0.05\%$ of IBD	



(a)

(b)

(c)

(d)

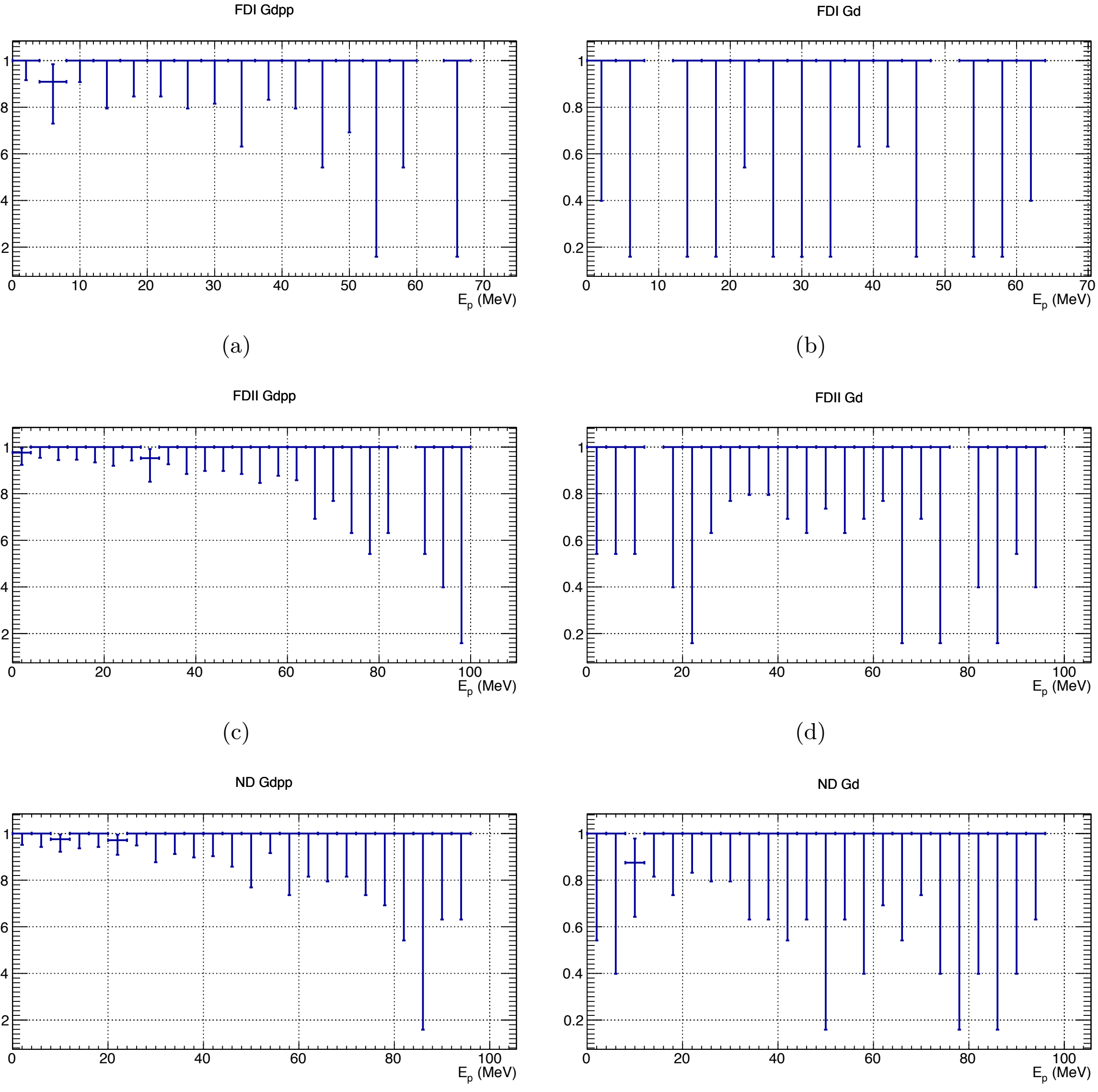
(e)

(f)

Figure 6.21: Impact of each of the Stopping Muons cuts on the prompt energy spectrum of the SM-free sample ($\Delta T > 30\,\mu s$) for each of the detector and capture configurations. Deviations from the unity show the impact of the cut inefficiency.

The efficiency computed in this way is an integral effect, averaged over the full energy range. However, in Double Chooz, the shape of the prompt energy spectrum is used to estimate $\sin^2(2\theta_{13})$, it is therefore important to verify that the cut does not impact the shape of the spectrum.
Figure 6.21 shows the ratio of the SM-free sample ($\Delta T > 30\,\mu s$) spectrum after application of the SM-cuts divided by the spectrum before applying the SM-cuts, here zoomed in on $[0; 20]$MeV for a better visualization of the IBD interest region of $[1; 10]$MeV. Deviations from the unity show the impact of the inefficiency of the cut.

An inefficiency small over the full spectrum, but more important at low energies could be dangerous for the Rate+Shape analysis as it would impact the shape of the spectrum. The shape of the prompt energy spectrum is particularly sensitive to $\sin^2(2\theta_{13})$ in the $[1-3]$MeV range; in this range, for all six configurations, the average inefficiency of the full cut does not exceed $\sim 0.4\%$, making it negligible compared to the $\sim 4\%$ systematic uncertainties of DC-III in that particular energy range.



(a)

(b)

(c)

(d)

(e)

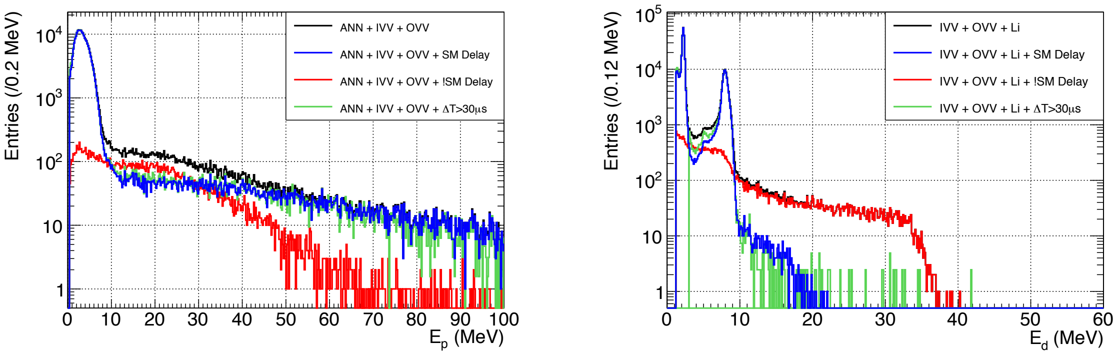
(f)

Figure 6.22: Ratio of the Fast Neutron spectrum before and after application of the SM-cut. No noticeable impact is visible here, hence the SM-cut does not bias the estimation of the Fast Neutron shape used to fit the prompt energy spectrum in the Rate+Shape analysis.

An impact on the SM-free spectrum at higher energies could mean that the SM-cut biases the shape of the Fast Neutron spectrum, which would be problematic since the shape of the Fast Neutrons is on input of the final Rate+Shape fit for the $\sin^2(2\theta)$ measurement. Figure 6.22 shows the ratio for each bin of the Fast Neutron spectrum with and without application of the SM-cut. The pure Fast Neutron spectrum is obtained by selecting the IV-tagged events and applying a $\Delta T > 30\,\mu\text{s}$ cut in order to reject all Stopping Muons tagged by the IV-tag. Deviation from the unity shows the bias on the Fast Neutron spectrum induced by the SM-cut. No impact on the Fast Neutron spectrum was found, allowing the shape of Fast Neutrons to be used in the Rate+Shape fit with no SM-induced bias on the θ_{13} measurement.

6.6 Impact of the SM-cuts on the IBD selection

Hereafter, in Figures 6.23 to 6.32 are presented the effects of the SM cut on "canonical" distributions such as, in usual reading order, the prompt energy spectrum, the delayed energy spectrum, the ΔT distribution, the vertex distribution along the Z-axis, and the selected and rejected vertex distributions projected on the (X,Z) and (X,Y) planes, where the black lines correspond to the vessels of the Target (innermost) and γ-Catcher (outermost). The plots corresponding to the ND Gd++ are broken down hereafter.



(a) Prompt energy spectrum

(b) Delayed energy spectrum

Figure 6.23: Prompt and delayed energy spectra, the black line corresponds to the IBD candidates selected by the full selection except SM-cuts, the blue line corresponds to the fully selected IBD candidates, the red line corresponds to the events rejected by the SM-cuts. The green line corresponds to the SM-free sample, with the cut $\Delta T > 30\,\mu\text{s}$, scaled to the blue line on the IBD energy region of interest. The good agreement between the blue and the green line strengthens the confidence that the SM-cut is removing all SM without impact on the IBD.

Figure 6.23 shows the energy spectra for the prompt and delayed events. The black line corresponds to the IBD sample selected by all vetoes except the SM-cut. The blue line corresponds to the candidates that pass all vetoes including the SM-cut. The red line corresponds to events that pass all other cuts but are rejected by the SM-cut. Finally, the green line corresponds to the SM-free sample, where SM are rejected by the $\Delta T > 30\,\mu\text{s}$, and scaled back to the black histogram on the IBD region of interest. This alternative method does not alter the IBD distributions, and provides a good alternative method to cross-check the sanity of the SM-cut. As previously stated, a 30 µs cut leads to different inefficiencies in the Target and γ-Catcher ($\sim 60\%$ in the Target and $\sim 20\%$ in the γ-Catcher) because of the order of magnitude difference between the capture times of each volume. Therefore, the n-H and n-Gd peak of the SM-free sample in the delayed energy spectrum need to be scaled separately to fit the black curve. The scaling is performed on the integral on the $[2; 2.4]$MeV and $[7.5; 8.5]$MeV. The good agreement between the blue and green lines over the IBD regions visually confirms the 0.129% inefficiency previously estimated, and the disappearance of the SM populations. This is especially visible for the GC-SM, in the high delayed energy region, which confirms the 99.73% rejection power.

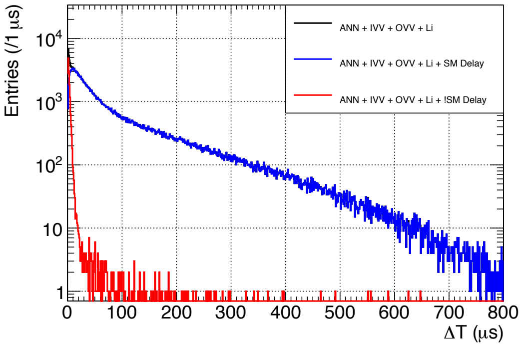
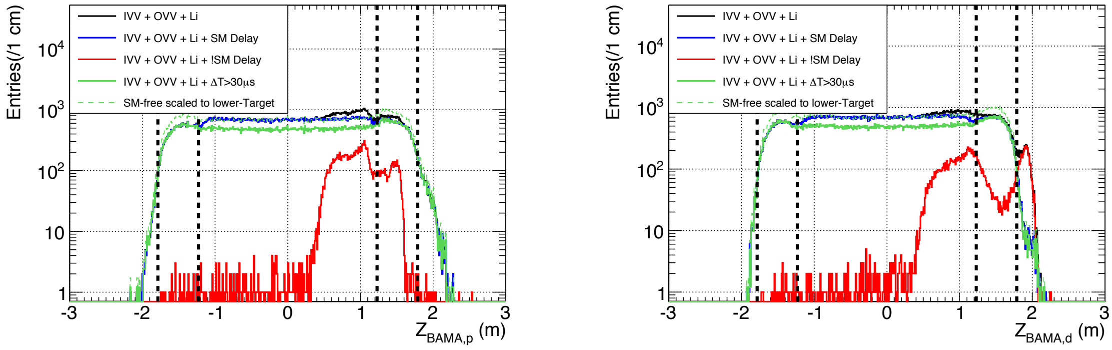


Figure 6.24: ΔT distribution. The black line corresponds to the full IBD candidates selected by all cuts except SM-cuts, the blue line corresponds to the full selection, including SM-cuts, and the red line correspond to the events rejected only by the SM-cuts.

Figure 6.24 shows the ΔT distribution. The black line correspond to events selected by all cuts except SM-cuts, the blue and red lines correspond to the events that pass all other cuts, and are respectively selected and rejected by the SM-Cut. The rejected sample shows a characteristic time of ~ 2 µs, compatible with Stopping Muon decays, and a very low contribution of IBD or Accidentals, with higher correlation times.

Figure 6.25 shows the candidate vertices projected on the Z-axis for the prompt (6.25(a)), and delayed (6.25(b)) triggers. The dotted lines show the position of the Target (two innermost lines) and the γ-Catcher (two outermost lines) acrylic vessels. In both cases, the black histogram, corresponding to candidates passing all non-SM cuts, shows a clear excess of events at the top of Target and γ-Catcher, corresponding to Stopping Muons. This excess disappears after application of the SM-cut. Rejected events, in red, correspond exclusively to this topological excess, with close to no rejected events in the lower part of the detector. The Buffer SM are particularly visible in the delayed trigger distribution, for which the Michel electron is reconstructed outside the γ-Catcher whereas the muon of such events (prompt trigger) is mis-reconstructed in the γ-Catcher. The SM-free sample, in green is here again scaled on the γ-Catcher (solid green line) and the Target (dashed green line) separately because of the different efficiencies. In order not to be biased by the SM rejection, the scaling on each volume is performed in the bottom part only ($[-2, -1.3]$m for the γ-Catcher, and $[-0.8, -0.4]$m for the Target). The agreement between the blue and green lines at positive values means that no excess of low-correlation-time events are left after the SM-cut is applied, and that no IBD-like



(a)

(b)

Figure 6.25: Z-axis vertex distribution for the prompt (6.25(a)), and for the delayed (6.25(b)) triggers. The black, blue and red lines are the same selection as previously described, the green solid line corresponds to the SM-free sample, scaled to the black line on the bottom part of the γ-Catcher, and the dashed green line corresponds to the same sample scaled to the bottom of the Target. The green lines are scaled on the bottom of the γ-Catcher and Target, where no SM are expected.

events are over-subtracted. Concerns were raised because of the excess of events at the top of the distribution compared to the bottom, even after SM reduction. However, the agreement with the SM-free sample on the upper regions demonstrates that this excess is not related to SM, and corresponds to remaining Fast Neutrons, with higher time correlation.

Figure 6.26 shows the delayed vertex distribution projected on the vertical (top) and horizontal planes (bottom), for rejected (left) and accepted events without applying the ANN veto. Rejected events pile up below the Chimney for mis-reconstructed Chimney SM, in the Buffer for Buffer SM and in the γ-Catcher for the GC-SM. No more excess corresponding to the rejected SM is visible in the accepted distributions. The remaining excess correspond to Accidentals and Fast Neutrons and disappear after application of the ANN and are also more abundant below the veto-acceptance hole that is the Chimney. Figure 6.27 shows the same distributions for the prompt vertices. Here Muon vertices are mis-reconstructed and pulled toward the center of the detector.

Figures 6.28 to 6.32 show the previously presented distributions for the ND Gd , FD-I Gd++, FD-I Gd, FD-II Gd++ and FD-II Gd analyses. For simplicity, only delayed event vertices are shown as the SM cut only applies on the delayed trigger. It is particularly notable that the Buffer SM and GC-SM disappear in the Far Detector. The Buffer SM are no longer present because there is no scintillator leak in the buffer of the Far Detector, and the GC-SM disappear because the Muon spectrum is harder, due to more shielding, and the fraction of Muons that could produce such GC-SM is much smaller.

In all configurations, the good agreement between the distribution of selected events (blue) and the SM-free distributions (green) provides a confirmation of the good rejection powers, and low inefficiencies achieved.

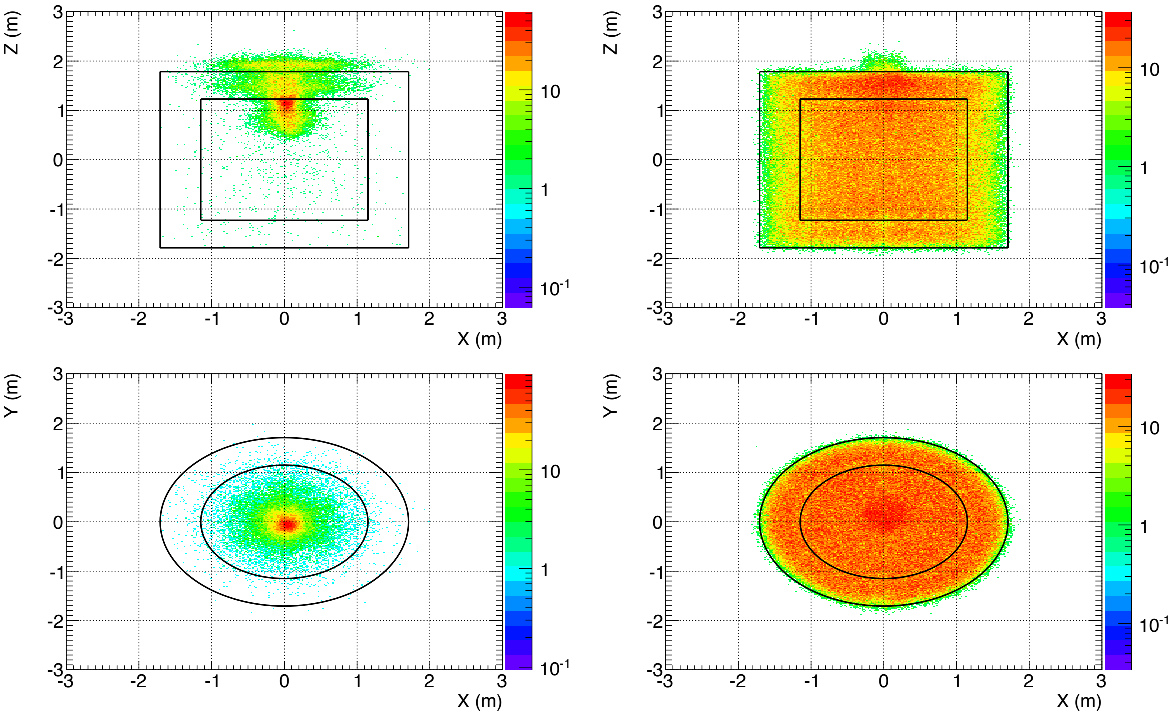


Figure 6.26: 2D projections of the delayed vertices of rejected (left) and accepted events (right), for projections on the vertical (top) and horizontal plane (bottom).

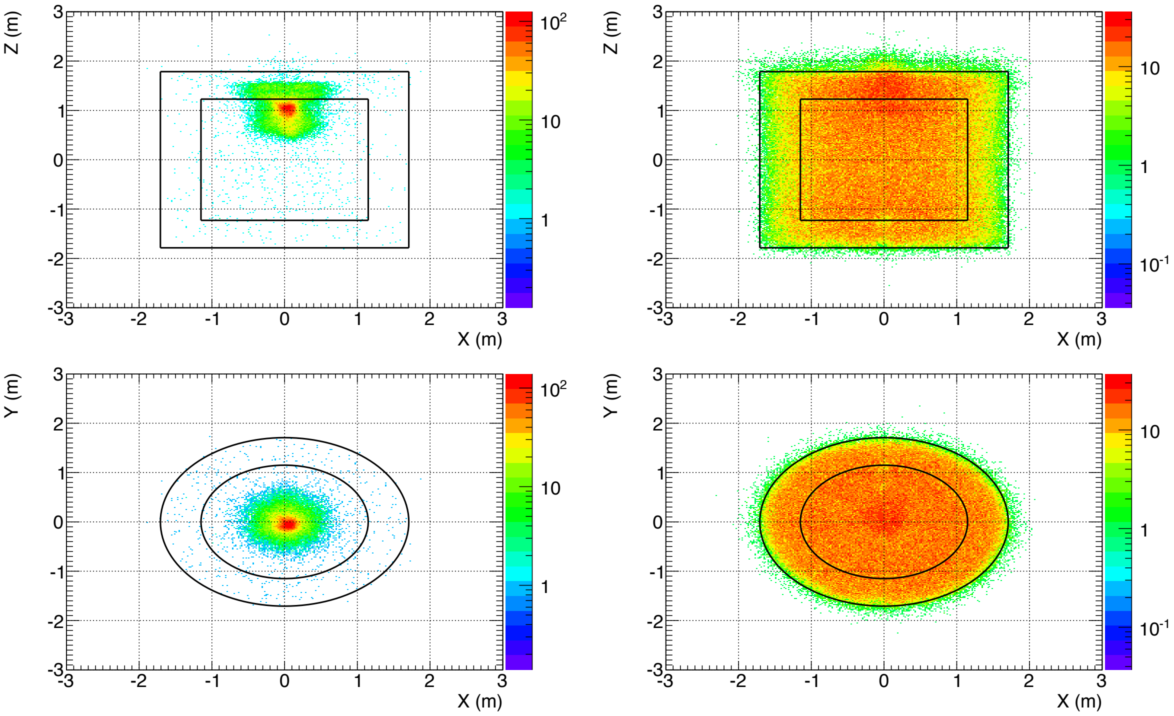
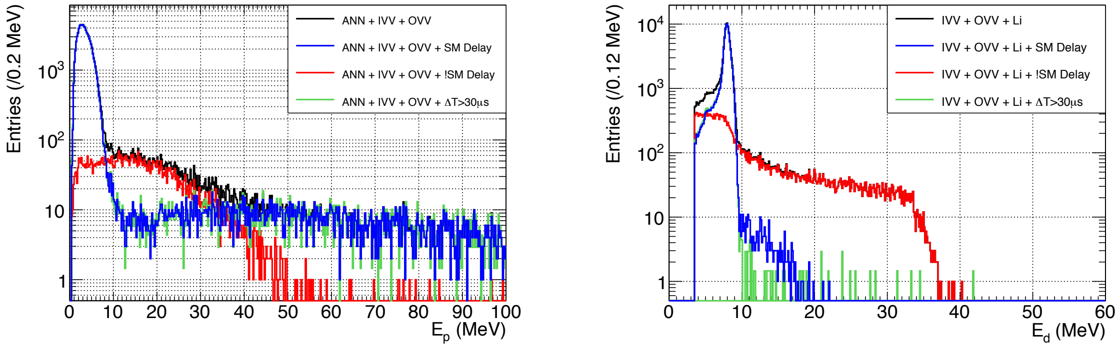
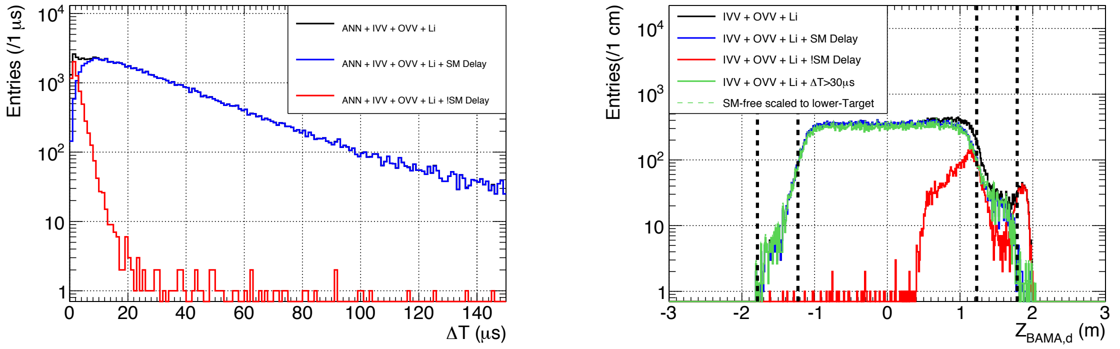


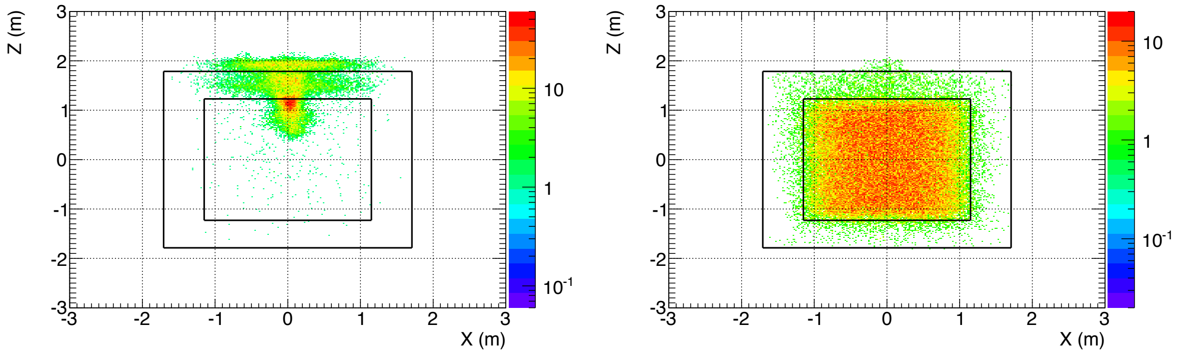
Figure 6.27: 2D projections of the prompt vertices of rejected (left) and accepted events (right), for projections on the vertical (top) and horizontal plane (bottom).



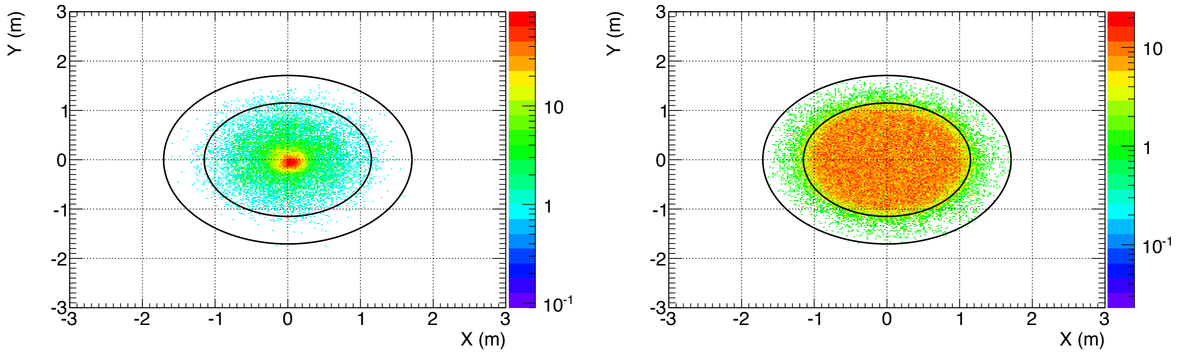
(a) Prompt energy spectrum (b) Delayed energy spectrum



(c) Time correlation (d) Z-axis delayed vertex distribution

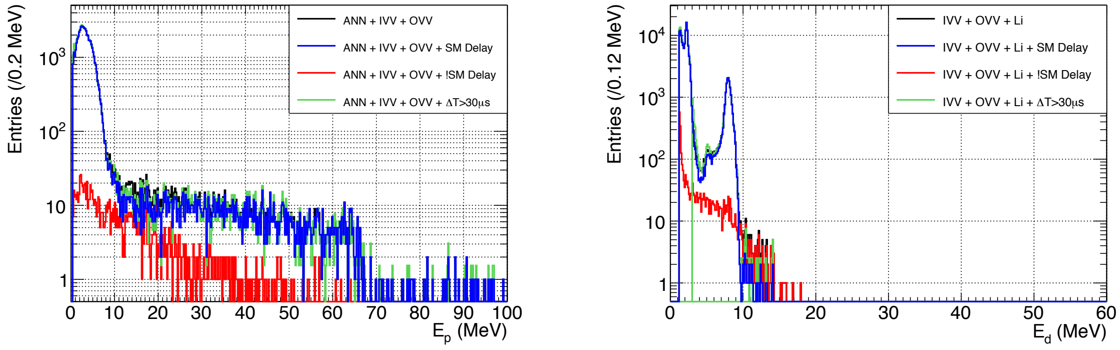


(e) (X,Z) rejected delayed SM vertex distribution (f) (X,Z) accepted delayed IBD vertex distribution

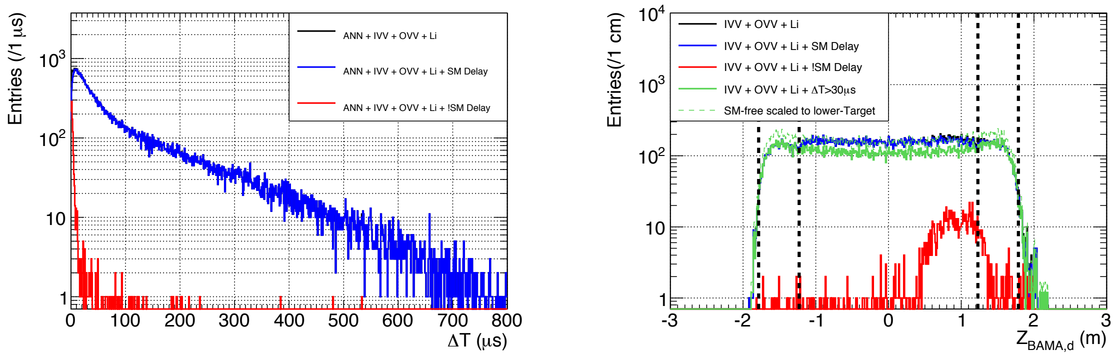


(g) (X,Y) rejected delayed SM vertex distribution (h) (X,Y) accepted delayed IBD vertex distribution

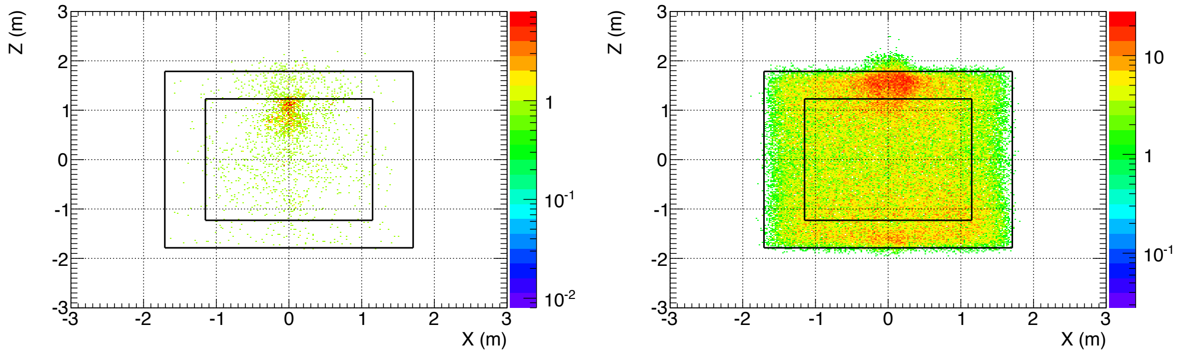
Figure 6.28: Effects of the SM-cut on the **ND Gd** IBD selection.



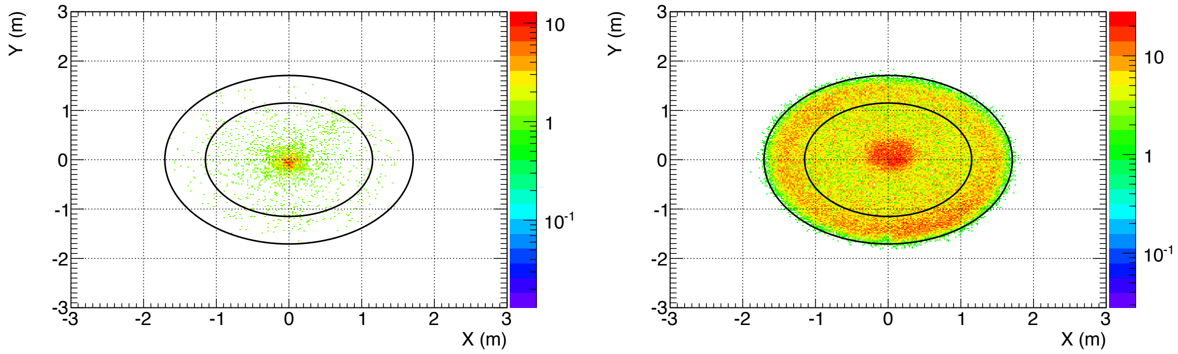
(a) Prompt energy spectrum (b) Delayed energy spectrum



(c) Time correlation (d) Z-axis delayed vertex distribution

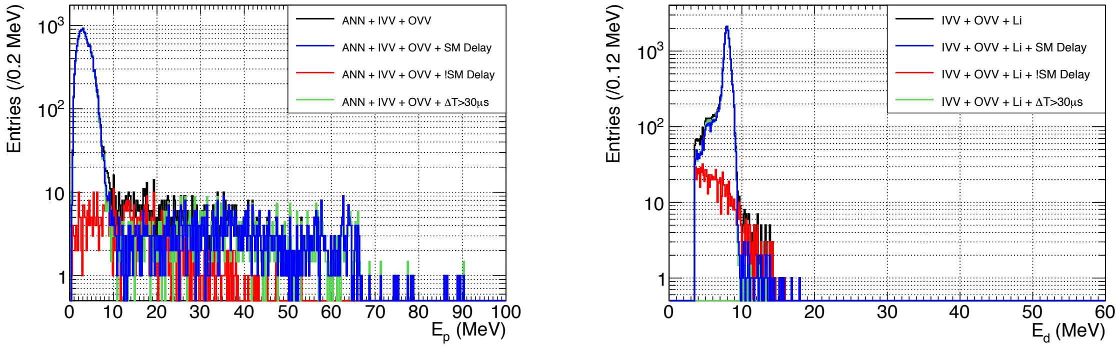


(e) (X,Z) rejected delayed SM vertex distribution (f) (X,Z) accepted delayed IBD vertex distribution



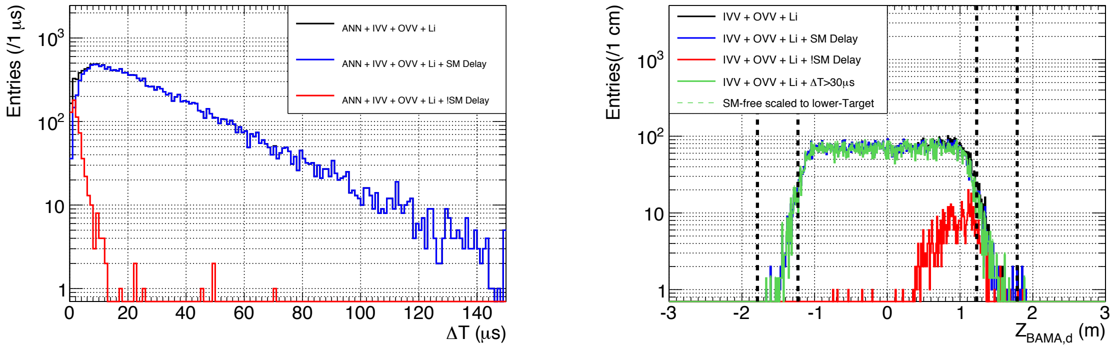
(g) (X,Y) rejected delayed SM vertex distribution (h) (X,Y) accepted delayed IBD vertex distribution

Figure 6.29: Effects of the SM-cut on the **FD-I Gd++** IBD selection.



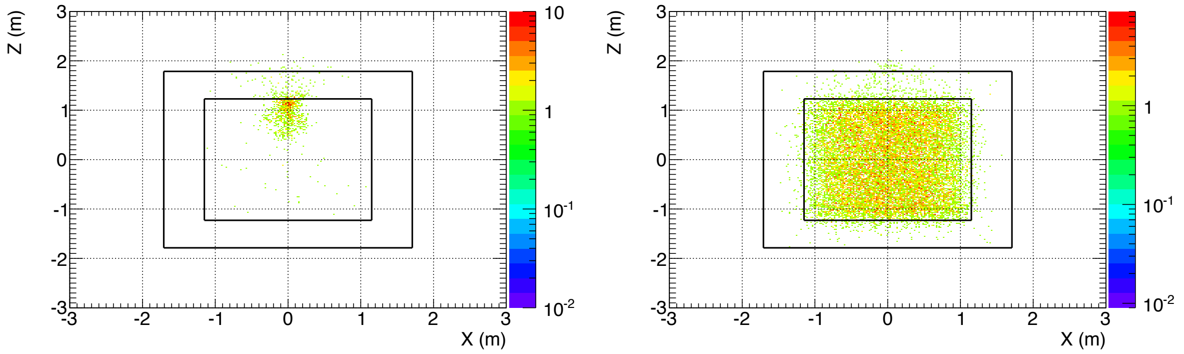
(a) Prompt energy spectrum

(b) Delayed energy spectrum



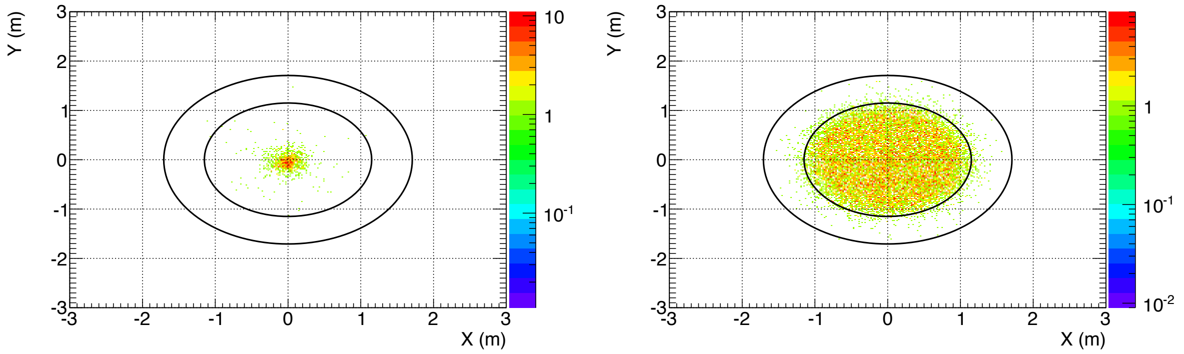
(c) Time correlation

(d) Z-axis delayed vertex distribution



(e) (X,Z) rejected delayed SM vertex distribution

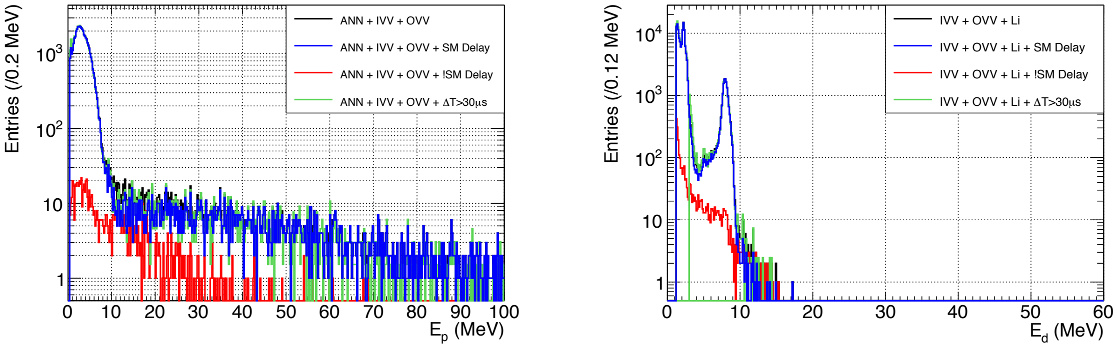
(f) (X,Z) accepted delayed IBD vertex distribution



(g) (X,Y) rejected delayed SM vertex distribution

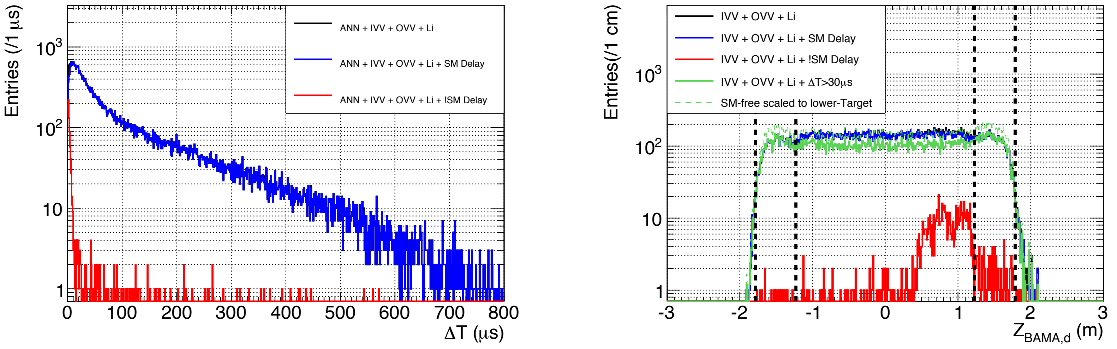
(h) (X,Y) accepted delayed IBD vertex distribution

Figure 6.30: Effects of the SM-cut on the **FD-I Gd** IBD selection.



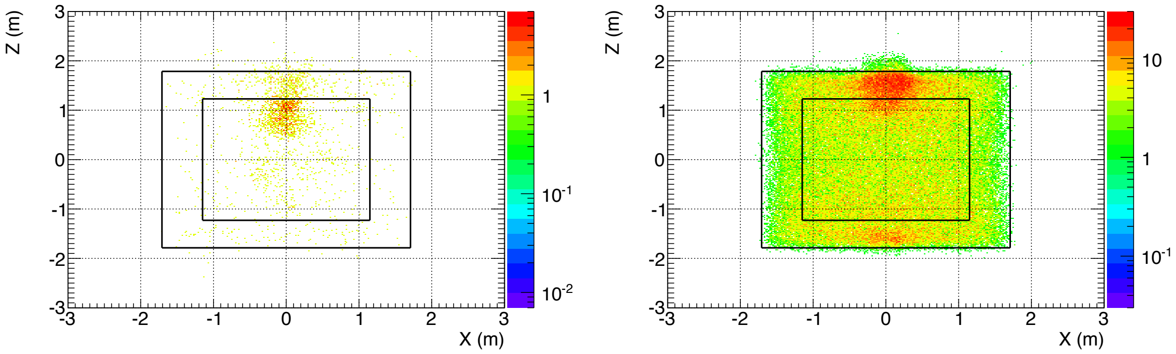
(a) Prompt energy spectrum

(b) Delayed energy spectrum



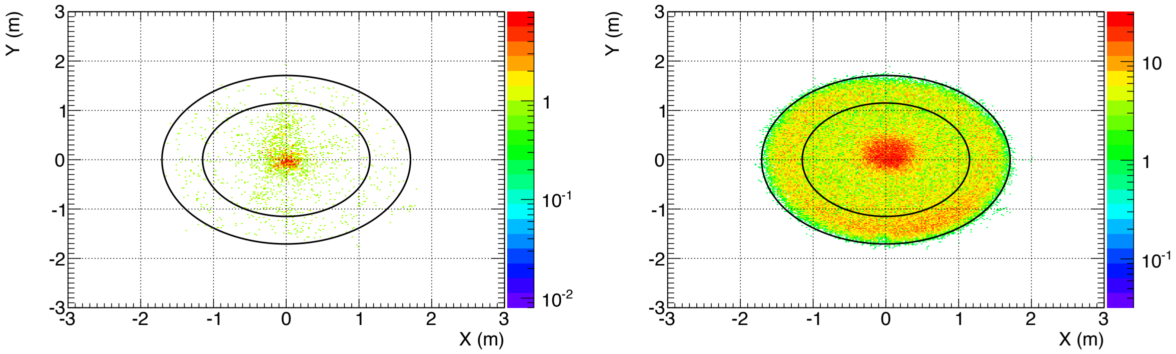
(c) Time correlation

(d) Z-axis delayed vertex distribution



(e) (X,Z) rejected delayed SM vertex distribution

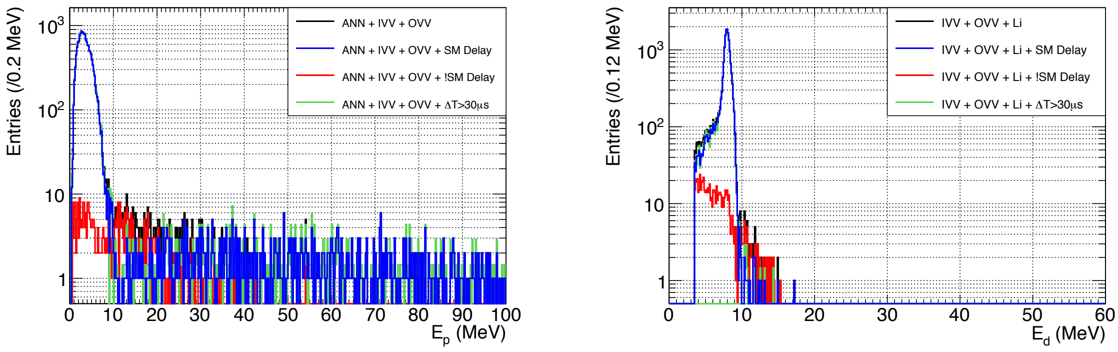
(f) (X,Z) accepted delayed IBD vertex distribution



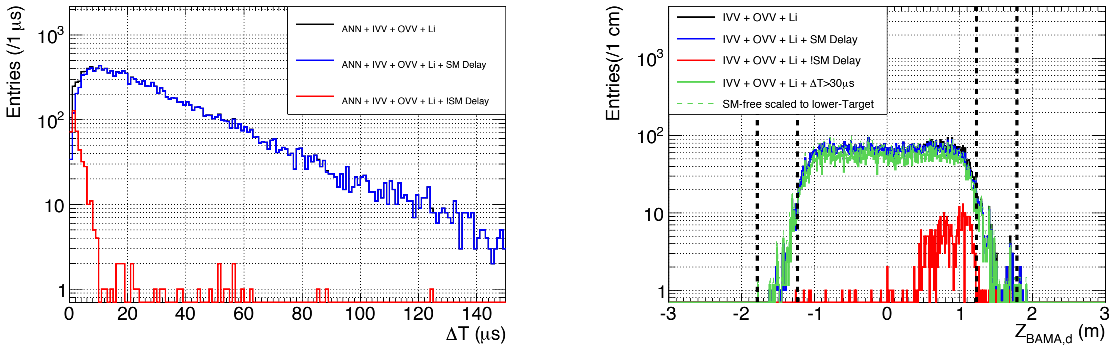
(g) (X,Y) rejected delayed SM vertex distribution

(h) (X,Y) accepted delayed IBD vertex distribution

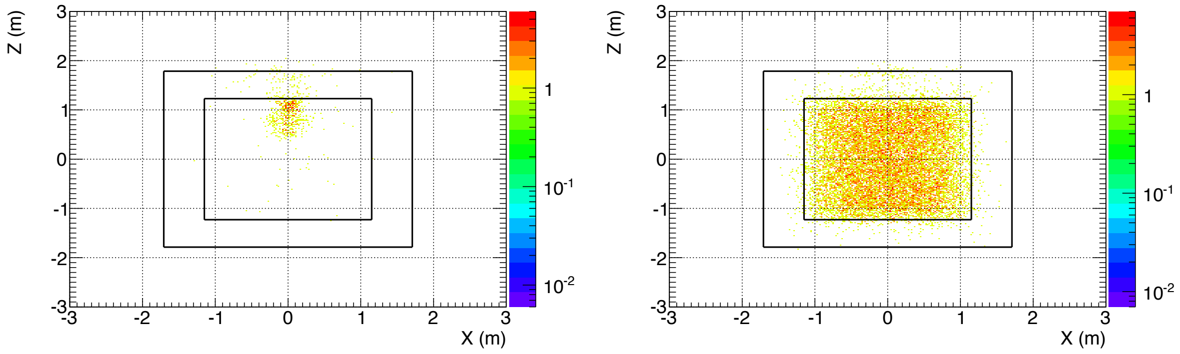
Figure 6.31: Effects of the SM-cut on the **FD-II Gd++** IBD selection.



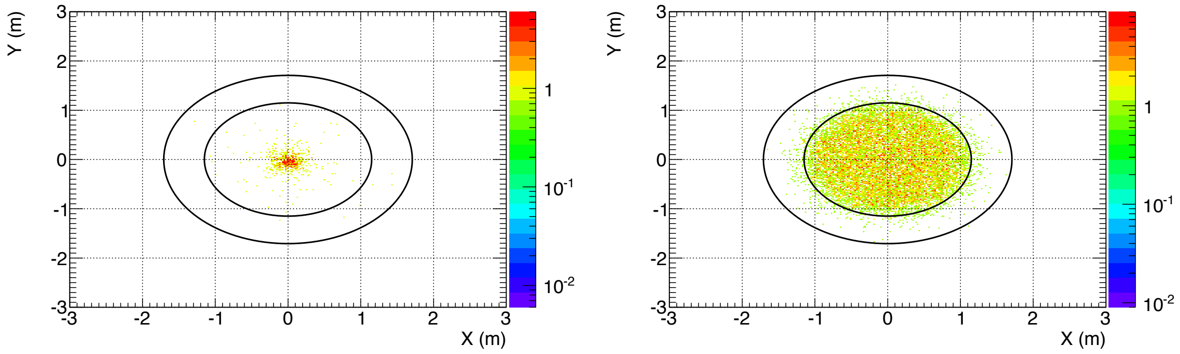
(a) Prompt energy spectrum (b) Delayed energy spectrum



(c) Time correlation (d) Z-axis delayed vertex distribution



(e) (X,Z) rejected delayed SM vertex distribution (f) (X,Z) accepted delayedIBD vertex distribution



(g) (X,Y) rejected delayed SM vertex distribution (h) (X,Y) accepted delayed IBD vertex distribution

Figure 6.32: Effects of the SM-cut on the **FD-II Gd** IBD selection.

6.7 Conclusion

Stopping Muons are a dangerous background of Double Chooz, especially in the Near Detector, where their rate is ~ 20 times higher than in the Far Detector, and reaches 15% of the IBD sample. The method developed during my thesis to reduce Stopping Muons allowed a near to total rejection of the Stopping Muons, bringing their contamination down to less than 0.2% of the IBD sample in the worst configuration. Moreover, the impact of the rejection on the IBD sample itself is kept to an inefficiency of a maximum of 0.32% in the FD-II configuration. The uncertainty of this estimation is limited by the statistics in the Far Detector and is expected to improve rapidly as FD-II accumulates more statistics in the future analyses.

Conclusion

Since their first detection by F. Reines and C. Cowan in 1956, sixty years ago, the neutrinos brought new light in our understanding of the physics, from high precision testing of the standard model to the explosion of supernovae. They can be used as messenger to probe the activity of nuclear reactors, the composition of earth's mantle, but also as astrophysical messenger as they are not deflected by magnetic fields nor absorbed by gas clouds. The discovery and understanding processes of the neutrino properties are briefly presented, as well as the inclusion of the neutrino in the Standard Model. I also describe the oscillation formalism and the historical measurements that lead to the understanding of these oscillations and to the measurement of the oscillation parameters.

I first present concept of the Double Chooz experiment aiming to measure the mixing parameter θ_{13} in the oscillation of neutrinos by quantifying the disappearance of the $\bar{\nu}_e$ from the two nuclear reactors of the Chooz nuclear power plant. Double Chooz proposed the multi-detector method to help constrain the reactor flux normalization independently from previous measurements, and from the flux prediction, in disagreement with existing data. This method have since then spread to other reactor θ_{13} experiments such as Daya Bay and RENO. The detector design has been presented, the three layers of the Inner detector allow a control of the detection systematics linked to the knowledge of the volume and the containment of the deposited energy by the Target and the γ-Catcher, but also helps reduce the Accidental background by introducing a Buffer volume to shield from radioactivity. The Inner Veto surrounding the detector helps tag and actively reject muons and muon-induced background. The data acquisition chain and event processing are also described.

Then are presented the analyses performed by Double Chooz in order to measure the θ_{13} parameter. First, the background rejection is presented. The agreement between the background model, i.e. the summed rates and shapes of the estimated remaining background contaminations, and the event rate observed in the reactor-off periods strengthens the confidence in the model. Then, the Rate+Shape and RRM analyses for DC-III, in the single-detector configuration, are presented. Finally, the first results of the DC-IV analysis, in the two-detector configuration, are presented. These results are still preliminary, and therefore the method is described instead of the actual results.

A first part of my work consisted in performing tests during the commissioning of the Near Detector to evaluate the linearity of the response of the Flash-ADC digitizers. The tests performed showed that the linearity of the Flash-ADCs was well within the manufacturer's limits, and that they could be installed in the acquisition chain of the Near Detector. From these tests, a better understanding of the Flash-ADC digitization was possible, especially on the inherent limitations on the baseline determination due to the flooring performed by the Flash-ADC and an analog noise small compared to the digitization step size.

The understanding of the Flash-ADC behaviour, and of the shape of a single PE waveform, elementary constituent of the PMT waveforms, allowed to develop an alternative charge reconstruction for the energy reconstruction of the events in Double Chooz. This new reconstruction has proven much more linear that the one currently used, however due to time and priority constraints, it has not been implemented in the Double Chooz framework and has only been applied to MC events. Nevertheless, a lot of attention was focused on understanding and quantifying every aspect of the reconstructions, guided by our understanding of the non-linearity sources in the Flash-ADC digitization.

Neutrino detection in Double Chooz relies on the time and space coincidence of two signals, one of which having an energy consistent with a neutron capture on Hydrogen or Gadolinium. As Double Chooz is a shallow experiment, with overburdens of 120 m w e and 300 m w e for respectively the Near and Far laboratories, Fast Neutrons showers from Muon interactions in the surrounding rocks create one of the major backgrounds in Double Chooz. For deeper experiments, or not relying on a prompt/delayed coincidence such as Dark Matter experiments, single Fast Neutron emission from (α, n) reactions in the laboratory rocks constitute the major source of background. It is then crucial to be able to obtain information on the ambient neutron flux and spectrum. A gaseous neutron TPC has been developed and installed in both Double Chooz laboratories. I took part in the installation, running, energy reconstruction, and analysis of the data from this TPC. Neutron scatterings on Helium nuclei produce a detectable ionization track that can be reconstructed in 3D to obtain information on energy and directionality. While the reconstruction of the neutron rate and spectrum still requires work, mostly due to the absence of a neutron simulation in Double Chooz, the recoil energy reconstruction achieved a resolution of $\sim 2\%$ for a nuclear recoil of 5 MeV.

A major part of my work was focused on developing rejection techniques for reducing the Stopping Muon background. I used variables based on the vertex reconstruction and on the delayed trigger energy to tag the Michel electron in the muon decay process. In order not to be tagged as Muons by other vetoes, a muon can only enter the Inner Detector through the Chimney and deposit less than 100 MeV in the Inner Detector. Three kinds of Muons decay topologies have been identified: decays located in the Chimney, in the Buffer or in the γ-Catcher.

- Muon decays located in the Chimney suffer from a poor energy and vertex reconstruction due to shadow effects; as a result the Michel electron energy spectrum has none of the characteristics of the Michel spectrum, and the vertex reconstruction likelihood is low.

- Muon decays located in the Buffer are only visible in the Near Detector, because of a, Inner Veto scintillator contamination into the Buffer. Since the contamination of scintillator is small, such events are only visible near the PMTs, in order for enough light to be collected.

- Muon decays in the γ-Catcher have a high delayed energy and are greatly reduced when limiting the delayed energy range to below 10 MeV as was the case for the n-Gd in DC-III. Those decays happen in a volume where the energy is well contained and the delayed energy spectrum exhibits a sharp cut-off at $\sim 50\,\mathrm{MeV}$ as expected from a Michel electron spectrum.

The techniques developed during my thesis have allowed for a reduction of Stopping Muons consistent with 100% in all Detector and capture configurations. The remaining contamination was estimated to be lower than 0.2% of the total IBD sample at 3σ C.L. in the Far Detector Gd++, and lower than 0.05% of the total IBD sample at 3σ C.L. in the other configurations. This is particularly important for the Near Detector in which the initial Stopping Muon contamination represents 15% of the total IBD candidates. The remaining contamination was deemed small enough to be neglected in the list of backgrounds included in the background model used to perform the $\sin^2(2\theta_{13})$ measurement with the Rate+Shape fit.

The oscillation parameter θ_{13} is now the best measured mixing angle. However, the neutrino mystery is far from solved, with many investigations still under way to measure the absolute mass of the neutrino, to determine the CP-violating phase δ_{CP}, or even to figure out whether one or several sterile neutrino exist. We merely stand at on the threshold, contemplating the path ahead to fully understanding this most evasive particle that is the Neutrino.

List of Figures

List of Tables

Glossary

IBD Inverse Beta Decay, refers to both the process through which the $\bar{\nu}_e$ are detected, and the candidates we consider as $\bar{\nu}_e$ interactions. 16

PMT Photo-Multiplier Tube. 16

$\mathcal{SM}$ Standard Model of particle physics, a model that describes the behaviour of elementary particles.. 19

SSM Solar Standard Model, a description of the production of solar energy and its thermodynamics. 23

MC Monte Carlo simulation. 45

ID-PMT Inner Detector PMT. 55

IV-PMT Inner Veto PMT. 55

OV Outer Veto. 55

ADC Analogue to Digital Converter, electronics that convert an analog signal into a digital value in ADC code. 57

DUQ Elementary charge unit corresponding to a 1 ADC code integrated over a time sample of 2 ns in the PMT waveform, knowing the effective impedance of the PMT $R = 500\,\Omega$. 61

PE Photo-Electron: can refer to either the primary electron knocked out of the photocathode of a PMT by a visible photon, or the electric signal output by the PMT corresponding to an amplified primary electron. 69

FEE Card of the read-out system that perform an amplification of the PMT signal for future digitization, and the summation of groups of non-amplified waveforms for trigger decision. 71

LSB Least Significant Bin: voltage step in a digitizer. It corresponds to the analog voltage equivalent of one ADC code. 72

ANN Artificial Neural Network designed to reject Accidentals based on the time and space correlation of IBD and their well defined delayed energy range, as opposed to Accidental coincidence being, by definition, uncorrelated. 110

TPC Time Projection Chamber. 147

SM Stopping Muon: Muon that stop in the detector, with a visible energy in the Inner Detector lower than the threshold $E_{p,\mathrm{max}}$. Stopping muons can either decay or capture on a nucleus. 173

FV Functional Value: negative logarithmic likelihood of RecoBAMA. 180

FV* Energy-corrected Functional Value (time only) of the delayed trigger.. 189

$\Delta\mathrm{BJ}_{z,d}$ Difference between the RecoJapan and RecoBAMA vertical vertex reconstruction. 190

CPS Ratio of Pulse Shape likelihoods assuming the event is in the Chimney over the likelihood for the reconstructed vertex. 193

Bibliography

[1] W. Pauli. Dear radioactive ladies and gentlemen. *Phys.Today*, 31N9:27, 1978.

[2] F. Reines et al. The Neutrino. *Nature*, 178:446, 1956.

[3] H.A. Tanaka. T2K: Latest Results. *Neutrino 2016 conference*, 2016.

[4] A. Bellerive et al. The Sudbury Neutrino Observatory. 2016.

[5] Y. Fukuda et al. (Super-Kamiokande Collaboration). Evidence for oscillation of atmospheric neutrinos. *Phys.Rev.Lett.*, 81:1562–1567, 1988.

[6] F. Ardellier et al. Letter of Intent for Double-CHOOZ: a Search for the Mixing Angle θ_{13}. arXiv:hep-ex/0405032v1, 14 May 2004.

[7] Y. Abe et al. Indication of Reactor $\overline{\nu}_e$ Disappearance in the Double Chooz Experiment. *Phys. Rev. Lett.*, 108:131801, Mar 2012.

[8] J. Chadwick. The intensty distribution in the magnetic spectrum of beta particles from radium (B + C). *Verh.Phys.Gesell.*, 16:383–391, 1914.

[9] N. Bohr *et al.* The Quantum Theory of Radiation. *Philosophical Magazine*, 47:785–802, 1924.

[10] J. Chadwick. Possible Existence of a Neutron. *Nature*, 129:312, 1932.

[11] E. Fermi. An attempt of a theory of beta radiation. *Z.Phys.*, 88:161–177, 1934.

[12] B. Pontecovro. Nuclear Capture of Mesons and the Meson Decay. *Phys. Rev.*, 72:246, 1947.

[13] C. L. Cowan *et al.* Detection of the free neutrino: A Confirmation. *Science*, 124:103–104, 1956.

[14] F. Reines et al. Detection of the Free Antineutrino. *Phys. Rev.*, 117:159–173, Jan 1960.

[15] T. D. Lee and C. N. Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[16] C. S. Wu et al. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[17] M. Goldhaber et al. Helicity of Neutrinos. *Phys. Rev.*, 109:1015–1017, Feb 1958.

[18] G Danby *et al.* Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys.Rev.Lett*, 9:36–44, Jul 1962.

[19] ALPEH DELPHI L3 OPAL SLD LEP. Precision electroweak measurements on the z resonance. *Physics Reports*, 427(5-6):257 – 454, 2006.

[20] DONUT collaboration. Observation of tau neutrino interactions. *Physics letters B*, 504:218–224, April 2001.

[21] John N Bahcall et al. New solar opacities, abundances, helioseismology, and neutrino fluxes. *The Astrophysical Journal Letters*, 621(1):L85, 2005.

[22] Yoichiro Suzuki. Neutrino 94 Kamiokande solar neutrino results. *Nuclear Physics B - Proceedings Supplements*, 38(1):54 – 59, 1995.

[23] Edward Kearns. Experimental measurements of atmospheric neutrinos. *Nuclear Physics B-Proceedings Supplements*, 70(1):315–323, 1999.

[24] M Ishitsuka et al. Super Kamiokande results: atmospheric and solar neutrinos. *arXiv preprint hep-ex/0406076*, 2004.

[25] Z. Maki et al. Remarks on the unified model of elementary particles. *Progress of Theoretical Physics*, 28(5):870–880, 1962.

[26] M. Apollonio et al. Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station. *Eur. Phys. J.*, C27:331–374, 2003.

[27] S. Abe et al. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. *Phys. Rev. Lett.*, 100:221803, 2008.

[28] P. Adamson et al. Combined Analysis of ν_μ Disappearance and $\nu_\mu \to \nu_e$ Appearance in MINOS Using Accelerator and Atmospheric Neutrinos. *Phys. Rev. Lett.*, 112:191801, May 2014.

[29] R. B. Patterson. Prospects for Measurement of the Neutrino Mass Hierarchy. *Ann. Rev. Nucl. Part. Sci.*, 65:177–192, 2015.

[30] A Yu Smirnov. The MSW effect and solar neutrinos. *arXiv preprint hep-ph/0305106*, 2003.

[31] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.

[32] K. Abe et al. Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6× 1 0 20 protons on target. *Physical Review D*, 91(7):072010, 2015.

[33] F.P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Physical Review Letters*, 108(17):171803, 2012.

[34] J. K. Ahn et al. Observation of reactor electron antineutrinos disappearance in the RENO experiment. *Physical Review Letters*, 108(19):191802, 2012.

[35] K Abe et al. Indication of electron neutrino appearance from an accelerator-produced off-axis muon neutrino beam. *Physical Review Letters*, 107(4):041801, 2011.

[36] K. Abe et al. Evidence of electron neutrino appearance in a muon neutrino beam. *Physical Review D*, 88(3):032002, 2013.

[37] J. H. Choi et al. Observation of Energy and Baseline Dependent Reactor Antineutrino Disappearance in the RENO Experiment. 2015.

[38] F. P. An et al. New measurement of θ_{13} via neutron capture on hydrogen at Daya Bay. 2016.

[39] Z.Y. Yu, on behalf of the Daya Bay Collaboration. Recent Results from the Daya Bay Experiment. *Neutrino 2016 conference*, 2016.

[40] Y. Declais et al. Study of reactor antineutrino interaction with proton at Bugey nuclear power plant. *Physics Letters B*, 338(2):383–389, 1994.

[41] Ettore Majorana. Teoria simmetrica dell'elettrone e del positrone. *Il Nuovo Cimento (1924-1942)*, 14(4):171–184, 1937.

[42] C Weinheimer. Direct determination of Neutrino Mass from Tritium Beta Spectrum. *arXiv preprint arXiv:0912.1619*, 2009.

[43] A.I. Belesev et al. Upper limit on additional neutrino mass eigenstate in 2 to 100 eV region from "Troitsk nu-Mass" data. *JETP letters*, 97(2):67–69, 2013.

[44] F. Vissani. Non-oscillation searches of neutrino mass in the age of oscillations. *arXiv preprint hep-ph/0012018*, 2000.

[45] H.V. Klapdor-Kleingrothaus et al. Evidence for neutrinoless double beta decay. *Modern Physics Letters A*, 16(37):2409–2420, 2001.

[46] M. Nakahata. SN1987A heralds the start of neutrino astronomy; http://cerncourier.com/cws/article/cern/29812, January 2007.

[47] M. Agostini et al. Spectroscopy of geoneutrinos from 2056 days of Borexino data. *Physical Review D*, 92(3):031101, 2015.

[48] S. Galata. KM3NeT/ORCA: Oscillation Research with Cosmics in the Abyss. *1st International Workshop for Neutrino Oscillation Tomography*, January 2016.

[49] Y Abe et al. First measurement of $\theta 13$ from delayed neutron capture on hydrogen in the Double Chooz experiment. *Physics Letters B*, 723(1):66–70, 2013.

[50] Y. Abe et al. Measurement of θ 13 in Double Chooz using neutron captures on hydrogen with novel background rejection techniques. *Journal of High Energy Physics*, 2016(1):1–29, 2016.

[51] P.Vogel and J.F. Beacom. The angular distribution of the neutron inverse beta decay, $\bar{\nu}_e + p \rightarrow e^+ + n$. *Phys. Rev., D*, 60, 1999.

[52] A Pichlmaier et al. Neutron lifetime measurement with the UCN trap-in-trap MAMBO II. *Physics Letters B*, 693(3):221–226, 2010.

[53] Th. A. Mueller et al. Improved Predictions of Reactor Antineutrino Spectra. *Phys. Rev.*, C83:054615, 2011.

[54] Y. Abe et al. Reactor $\bar{\nu}_e$ disappearance in the Double Chooz Experiment. *Phys. Rev., D*, 86(052008), 2012.

[55] Y. Abe et al. Background-independent measurement of in Double Chooz. *Physics Letters B*, 735:51 – 56, 2014.

[56] Y. Abe *et al.* Improved measurements of the neutrino mixing angle θ 13 with the Double Chooz detector. *Journal of High Energy Physics*, 2014(10):1–44, 2014.

[57] Y. Abe et al. Measurement of θ_{13} in Double Chooz using neutron captures on hydrogen with novel background rejection techniques. 2015.

[58] M. Ishitsuka, on behalf of the Double Chooz Collaboration. New Results of Double Chooz. 51^{st} *Rencontres de Moriond EW 2016*, 2016.

[59] Santamarina et al. The JEFF-3.1.1 Nuclear Data Library : JEFF Report 21. *OCDE - NEA*, 2006.

[60] Brookhaven National Laboratory. `www.nndc.bnl.gov`.

[61] Y. Abe et al. Improved measurements of the neutrino mixing angle θ 13 with the Double Chooz detector. *Journal of High Energy Physics*, 2014(10):1–44, 2014.

[62] O. Méplan et al. MURE: MCNP Utility for Reactor Evolution. User Guide - Version 1.9.2012.

[63] O. Méplan et al. MURE: MCNP Utility for Reactor Evolution, Description of the Methods, First Applications and Results. In *ENC 2005 Proceedings, Versailles, France*, 2005.

[64] G. Marleau et al. A user's guide for DRAGON Version 4. Technical report, Ecole Polytechnique de Montréal, Institut de Génie Nucléaire, Technical Report IGE-294, 2013.

[65] F. Von Feilitzsch et al. Experimental beta spectra from Pu-239 and U-235 thermal neutron fission products and their correlated anti-neutrinos spectra. *Phys. Lett.*, B118:162–166, 1982.

[66] V.I. Kopeikin et al. Reactor as a Source of Antineutrinos: Thermal Fission Energy. *Phys. Atom. Nucl.*, 67:1892–1899, 2004.

[67] X.B. Ma et al. Improved calculation of the energy release in neutron-induced fission. *Phys. Rev., C*, 99(014605), 2013.

[68] P. Huber and T. Schwetz. Precision spectroscopy with reactor antineutrinos. *Phys. Rev., D*, 70(053011), 2004.

[69] A. Onillon. *Prédiction des taux de fission des cœurs de Chooz et estimation des incertitudes associées dans le cadre de l'expérience Double Chooz.* PhD thesis, Université de Nantes Angers Le Mans, 2014.

[70] C Aberle et al. Large scale Gd-beta-diketonate based organic liquid scintillator production for antineutrino detection. *Journal of Instrumentation*, 7(06):P06008, 2012.

[71] F Beissel et al. The trigger and timing system of the Double Chooz experiment. *Journal of Instrumentation*, 8(01):T01003, 2013.

[72] A. Remoto. *Correlated background and impact on the measurement of* θ_{13} *with the Double Chooz detector.* PhD thesis, Université de Nantes Angers Le Mans, 2012.

[73] G. Pronost. *Études pour la mesure de l'angle* θ_{13} *avec l'expérience Double-Chooz.* PhD thesis, Université de Nantes Angers Le Mans, 2015.

[74] R. Roncin. *From the measurement of the* θ_{13} *mixing angle to the search for geo-neutrinos: studying* $\bar{\nu}_e$ *with Double Chooz and Borexino.* PhD thesis, Paris Diderot and Universita degli Studi dell'Aquila, 2014.

[75] C. Bauer et al. Qualification tests of 474 photomultiplier tubes for the inner detector of the Double Chooz experiment. *Journal of Instrumentation*, 6(06):P06008, 2011.

[76] T. Matsubara et al. Evaluation of 400 low background 10-in. photo-multiplier tubes for the Double Chooz experiment. *Nucl. Instrum. Meth.*, A661:16–25, 2012.

[77] Y. Abe et al. The Waveform Digitiser of the Double Chooz Experiment: Performance and Quantisation Effects on PhotoMultiplier Tube Signals. *JINST*, 8:P08015, 2013.

[78] B. Widrow et al. Statistical theory of quantization. *IEEE Transactions on Instrumentation and Measurement*, 45(2):353–361, 1996.

[79] P. Carbone et al. Mean value and variance of noisy quantized data. *Measurement*, 23(3):131 – 144, 1998.

[80] National Semiconductor. *ADC08D500 Data Sheet.*

[81] Tarek Akiri. *Flash-ADC test, optimization of the detector design and development of a new concept of spatial reconstruction in the Double Chooz neutrino experiment for the measurement of the* θ_{13} *mixing angle.* Theses, Université Paris-Diderot - Paris VII, September 2010.

[82] Y. Abe et al. Characterization of the Spontaneous Light Emission of the PMTs used in the Double Chooz Experiment. 2016.

[83] Th. A. Mueller et al. Improved Predictions of Reactor Antineutrino Spectra. *Phys. Rev. C*, 83:054615, May 2011.

[84] R. Carr. *Measurements of Electron Antineutrino Disappearance in the Double Chooz Experiment.* PhD thesis, Columbia University, 2015.

[85] A. Cabrera for the Double Chooz Collaboration. Double Chooz III : First Results. *Seminar in LAL (France)*, May 22nd, 2014.

[86] S.-H. Seo for the RENO Collaboration. New Results from RENO. *Neutrino 2014 conference*, June 2014.

[87] W. Wang for the Daya Bay Collaboration. The latest oscillation results from the Daya Bay reactor neutrino experiment. *ICHEP 2014 conference*, July 2014.

[88] A. Cucoanes et al. Reactor Neutrino Flux Uncertainty Suppression on Multiple Detector Experiments. 2015.

[89] A. Cabrera and on behalf of the Double Chooz Collaboration. Double Chooz (multi-detector results). *Neutrino 2016 conference*, 2016.

[90] A. Hourlier et al. IBD Background Rejection and Tagging at the Double Chooz Experiment. *Neutrino 2016 conference*, 2016.

[91] J. Lopez et al. A Prototype Detector for Directional Measurement of the Cosmogenic Neutron Flux. *Nucl. Instrum. Meth.*, A673:22–31, 2012.

[92] J. A. Formaggio and CJ Martoff. Backgrounds to sensitive experiments underground. *Annu. Rev. Nucl. Part. Sci.*, 54:361–412, 2004.

[93] Ho-yin Ngai. *Measurement of cosmic-ray muon induced neutrons in the Aberdeen Tunelunderground laboratory in Hong Kong.* PhD thesis, University of Hong Kong, 2012.

[94] Y.-F. Wang et al. Predicting neutron production from cosmic-ray muons. *Phys. Rev. D*, 64:013012, Jun 2001.

[95] R Hertenberger et al. Muon-induced neutron and pion production in an organic liquid scintillator at a shallow depth. *Physical Review C*, 52(6):3449, 1995.

[96] L.B. Bezrukov et al. Investigation of depth intensity curve of nuclear evens induced with cosmic ray muons. *Yadern. Fiz*, 17(1):98–103, 1973.

[97] F Boehm et al. Neutron production by cosmic-ray muons at shallow depth. *Physical Review D*, 62(9):092005, 2000.

[98] RI Enikeev et al. Hadrons generated by cosmic-ray muons underground. *Sov. J. Nucl. Phys.(Engl. Transl.);(United States)*, 46(5), 1987.

[99] M. Aglietta et al. Neutron flux generated by cosmic-ray mouns at 5200 hg/cm2 s.r. underground. Depth-neutron intensity curve. *Il Nuovo Cimento C*, 12(4):467–477, 1989.

[100] Lance M Garrison. *Measurement of Neutron and Muon Fluxes 100 m Underground with the SciBath Detector.* PhD thesis, faculty of the Graduate School in partial fulfillment of the requirements for the degree Doctor of Philosophy in the Department of Physics, Indiana University, 2014.

[101] D.-M. Mei and A. Hime. Muon-induced background study for underground laboratories. *Phys. Rev. D*, 73:053004, Mar 2006.

[102] J. F. Ziegler, M. D. Ziegler, and J. P. Biersack. SRIM - The stopping and range of ions in matter (2010). *Nuclear Instruments and Methods in Physics Research B*, 268:1818–1823, June 2010.

[103] Y. Abe et al. Muon capture on light isotopes in Double Chooz. *Phys. Rev. C*, 93:054608, May 2015.

[104] P.J. Chang et al. TechNote: Muon measurements in Double Chooz. Double Chooz internal note, 2016 July.

Abstract

Double Chooz is a reactor anti-neutrino experiment which measures the θ_{13} mixing parameter in the neutrino oscillation phenomenon thanks to two identical underground detectors located at different distances from the two reactors of the Chooz nuclear power plant in the French Ardennes. The Far Detector started data taking in april 2011 and the Near Detector began operations in january 2015. This thesis presents an analysis of both the single- and two-detector phases. Neutrinos interact in Double Chooz through inverse β decay on protons, where positron and a neutron are emitted, creating two coincident signals. Analyses were performed for each detector on independent samples using the neutron capture on Gadolinium or on Hydrogen. A combined analysis, using both neutron captures is also presented. Backgrounds to the neutrino sample originate from the ambient radioactivity and the abundance of atmospheric muons (cosmogenic (β-n) emitters, fast neutron showers, and muons stopping and decaying inside the detector). This thesis focuses on developing techniques for measuring and reducing the Stopping Muon background. With colleagues from APC and MIT, a 60 litre neutron TPC was installed in both underground laboratories at Chooz. This low pressure He/CF_4-based time projection chamber, called DCTPC, measures the flux, energy spectrum and direction of fast neutrons in both locations, with different overburdens (150 and 300 meters water equivalent for the Near and Far Laboratories respectively).

KEYWORDS: Double Chooz, neutrino oscillation, mixing angle, muons, DCTPC, neutrons

Résumé

L'expérience Double Chooz étudie les anti-neutrinos électroniques provenant de réacteurs nucléaires pour mesurer le paramètre de mélange θ_{13} dans le phénomène d'oscillation des neutrinos. Deux détecteurs souterrains identiques sont placés à différentes distances des réacteurs de la centrale de Chooz dans les Ardennes. Le Détecteur Lointain est entré en service en avril 2011, et le Détecteur Proche en janvier 2015. Cette thèse présente une analyse des phase simple et double détecteurs de l'expérience.

Les neutrinos sont détectés par désintégration β inverse dans Double Chooz, où un positron et un neutron sont émis, créant deux signaux coïncidents. Différentes analyses ont été menées sur des échantillons dans lesquels le neutron capture sur un atome de gadolinium ou d'hydrogène. Une analyse jointe des deux captures est également présentée.

Les bruits de fond proviennent de la radioactivité ambiante et des muons atmosphériques (émetteurs (β-n), neutrons rapides et muons s'arrêtant et se désintégrant dans le détecteur). Cette thèse se focalise sur le développement de techniques pour réduire le bruit de fond lié aux muons se désintégrant dans le détecteur.

Avec l'aide de collègues de l'APC et du MIT, un TPC à neutrons a été installé dans les deux laboratoires souterrains à Chooz. Cette chambre à projection temporelle, appelée DCTPC, remplie d'un mélange helium/CF_4 mesure le flux, le spectre en énergie, et la direction des neutrons rapides dans les deux laboratoires avec des blindages différents (respectivement 150 et 300 mètres équivalents d'eau pour le Laboratoire Proche et Lointain).

MOTS-CLÉS: Double Chooz, oscillation de neutrino, angle de mélange, muons, DCTPC, neutrons