Studies for the measurement of the mixing angle θ_13 with the Double Chooz experiment

Guillaume Pronost

Guillaume Pronost. Studies for the measurement of the mixing angle θ_13 with the Double Chooz experiment. High Energy Physics - Experiment [hep-ex]. Université Nantes Angers Le Mans, 2015. English. NNT : 2015NANT2039 . tel-01409566

HAL Id: tel-01409566
https://theses.hal.science/tel-01409566v1

Submitted on 6 Dec 2016

L'UNAM l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Guillaume PRONOST

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université de Nantes
sous le label de L'Université Nantes Angers Le Mans

École doctorale : 3MPL

Discipline : Constituants élémentaires et physique théorique
Spécialité : Physique des particules
Unité de recherche : SUBATECH - UMR 6457

Soutenue le 30 octobre 2015
Thèse N°:

Études pour la mesure de l'angle θ_{13} avec l'expérience Double-Chooz

Rapporteurs :	**Antoine KOUCHNER**, Maitre de Conférences, HDR, Université Paris Diderot Paris 7, APC **Alfons WEBER**, Professeur, University of Oxford
Examinateurs :	**Anatael CABRERA**, Chargé de Recherche, CNRS/IN2P3, APC **Dominique DUCHESNEAU**, Directeur de Recherche, CNRS/IN2P3, LAPP **Ginés MARTINEZ**, Directeur de Recherche, CNRS/IN2P3, SUBATECH (Président du Jury)
Directeur de Thèse :	**Jacques MARTINO**, Directeur de Recherche, CNRS/IN2P3, SUBATECH
Encadrant scientifique :	**Frédéric YERMIA**, Maitre de Conférences, Université de Nantes

Acknowledgments

Je souhaite tout d'abord remercier l'ensemble des membres de mon jury : Alfons Weber et Antoine Kouchner pour leur lecture attentive de mon manuscrit ainsi que pour leurs commentaires constructifs sur les améliorations à y apporter, Gines Martinez pour avoir présidé le jury, ainsi qu'Anatael Cabrera pour avoir accepté d'en faire partie.

Merci beaucoup à Frédéric pour avoir encadré ma thèse. Plus de trois années passées à travailler avec toi (depuis le stage de master 2), durant lesquelles j'ai pu apprécier tes conseils, nos discussions et ton sens de l'humour. Merci d'avoir été là et de m'avoir aidé aussi bien dans mon travail de thèse que dans tes conseils pour améliorer la clarté de mes présentations et de mon manuscrit. Merci à Jacques Martino pour avoir dirigé ma thèse, même si nous n'avons jamais vraiment travaillé ensemble.

Je souhaiterais remercier encore une fois Anatael, non seulement pour avoir fait partie de mon jury, mais aussi pour m'avoir soutenu durant ces travaux de thèse. Tes conseils et tes suggestions de travaux ont permis de faire des contributions de cette thèse ce qu'elles sont désormais. Merci pour ta confiance.

Ces trois années passées dans le groupe ERDRE et plus généralement au sein du laboratoire SUBATECH, ont été un réel plaisir. Merci à Muriel, à Amanda et à Nicolas pour leur accueil au sein du groupe ERDRE. Merci à Andi pour ses conseils sur l'étalonnage du détecteur de Double Chooz. Merci à Baptiste L. et à Baptiste M. pour nos discussions. Merci Tomo pour m'avoir découvrir de bons restaux à Tokyo ! Enfin merci Arnaud, Richard, Benoit, Jean-Luc, Virginia, Julien, Luca, Philippe, Nol, Elisabeth, Julie... La liste est longue et j'en oublie surement. Merci aussi aux membres des services informatique et administratif pour leur aide et leur sympathie, merci Stéphanie pour toute ton aide !

Merci aux thésards ou post-docs de SUBATECH. D'abord, à mes co-bureaux passés et présents, merci de m'avoir supporté et d'avoir inlasablement rassuré Fred qui me cherchait lorsque j'étais partis en conférence, meeting et autre. À Vincent pour les "ta gueule !" amicaux lorsqu'il nous arrivait d'éternuer, à Jennifer pour avoir supporté mes blagues pas drôles, à Charlotte aussi, même si tu as quitté notre bureau bien trop tôt ! Merci Florian pour ton humour et bon courage, c'est toi le prochain (du bureau) à soutenir ! Merci Grégoire pour nos discussions diverses et variées (notamment sur Cthulhu), et à Fanny et Loïc, bon courage à vous aussi ! Merci aussi à François, même si tu sembles toujours persuadé de t'appeler Benjamin G., à Benjamin S. pour ton humour et nos discussions sur la soupe et la salade. Enfin merci Maxime pour nos discussions sur les jeux de stratégie. Merci aussi à Anthony, Javier, Lucia, Martin, Zak, Lucille, Roland, Loick, Aurélien, José, Alexis, Antoine, Kévin, Alexandre, Alexandre et Benjamin. Bon courage à ceux qui n'ont pas encore soutenu !

I would also like to thank the members of the Double Chooz collaboration. DC collaboration is a cheerful and friendly community. I would like to thank specially Roberto S., for the help with the artificial neural networks. Thanks to the CIEMAT group of Double Chooz. I also want to thank Masaki, Yosuke (Abe-dono!), Kuze-sensei and all the members of Kuze laboratory, for the invitation to spend one month in TokyoTech. It was a great experience for me. A special thanks to you Ralitsa, good luck with the energy callibration. Let's have a drink soon! Merci aux membres du groupe Double-Chooz à l'APC où j'ai pu passer de temps en temps, pour une réunion, une soutenance de thèse ou autre.

Romain, merci pour ton accueil dans la collaboration. Je n'oublierai pas ces heures passées à travailler sur DataMigro. Nous avons été appelés sur Skype, sur téléphone (et effectivement, même à Montréal !) pour relancer les transferts. Fais-moi penser, si jamais on retourne un jour à New York, de ne plus prendre les lits du bas dans les auberges de jeunesse ! Merci Antoine C., pour tes commentaires c'était toujours sympathique de faire un meeting ou un repas avec toi. Et merci Valérian pour m'avoir montré que finalement, Antoine C. ne râle pas tant que cela. Enfin merci Adrien pour tes conseils avisés sur mon anglais. Je te souhaite beaucoup de courage avec le développement de RecoZoR et la prise en main de DataMigro.

Finalement, merci à mes amis et à ma famille. Merci à Romain "Carter", Seb et Tiphaine pour leur aide dans la relecture de ma thèse. Merci aux rôlistes de Pontivy et de Lorient (la liste est trop longue), je continuerai de venir de temps à autre pour une partie ou un verre ! Enfin, merci à mes parents, à mon frère et à mes soeurs pour votre soutien durant toutes ces années.

Contents

Introduction

In 1956, F. Reines and C.L. Cowan discovered an "elusive, a poltergeist" particle: the neutrino. Since their prediction by W. Pauli in 1930, the neutrinos, despite to be one the most abundant particle in the Universe, are difficult to observe. In the 30s, many physicists were even not convinced by the possibility to detect them. Thereafter, the discovery of the neutrino led to new developments of the particle physics.

Since the Super-Kamiokande experiment in 1998, there are several demonstrations of the existence of an oscillation phenomenon. This phenomenon is described by the theories of B. Pontecorvo and of Z. Maki, M. Nakagawa, and S. Sakata. It allows neutrinos to oscillate from one flavor to another. Therefore, an electron antineutrino, produced by the β decays within a nuclear reactor core could be detected as an antineutrino of a different flavor after a medium travel distance (several hundred meters). The Double Chooz experiment measures such oscillations in order to measure one of the parameters of the oscillation theory: the θ_{13} mixing angle. Until November 2011, this mixing angle was the last not-known mixing angle of the neutrino oscillation.

In this thesis, we present studies for the measurement of the θ_{13} mixing angle with the Double Chooz experiment.

The first chapter exposes a brief history of the neutrino research and details the theoretical bases of the neutrino physics. The Standard Model of particles physics is presented in order to introduce the neutrino particle. The Pontecorvo-Maki-Nakagawa-Sakata theory of neutrino oscillations is also reviewed. In this chapter, the lastest experimental measurements of neutrino oscillation parameters and of the neutrino masses are detailed.

The second chapter is the presentation of the Double Chooz experiment. The detection principle and the detector design are described. The online and offline reconstruction of the experiment is also presented. A particular focus has been made on the Inner Veto vertex reconstruction algorithm, which is one of the works developed during this thesis. This Inner Veto vertex reconstruction algorithm has been used in the Double Chooz analyses in order to developed a background rejection method, called Inner Veto veto, and to improve the method to measure the correlated background shape and rate.

Since the energy reconstruction is one of the critical point of the Double Chooz analyses in order to measure θ_{13}, the third chapter is the description of the energy reconstruction process in the Double Chooz analysis. This is one of the main topics of this thesis. The different parts of the energy reconstruction are described. The energy reconstruction of Double Chooz consists mainly in the application of several corrections to the detector response. The measurement of the so-called Uniformity correction and of the charge non-linearity correction were ones of the works made during this thesis. This chapter also presents the validation of the energy reconstruction, and a focus is made on the adaptation of the energy reconstruction for the Hydrogen analysis of Double Chooz. These two last analyses were performed during this thesis.

In order to allow the Double Chooz analysis to access an improve the resolution on the reconstructed number of photoelectrons (PE) produced in the photomultiplier and on their arrival time, we developed a new calorimetry method, detailed in the fourth chapter. This method uses the shape information from the photomultipliers waveforms, whereas the standard calorimetry method of the experiment is shape independent. The algorithm developed, called “RecoZoR”, is presented, as well as its results.

The fifth chapter is the description of the two types of Double Chooz analyses: the Gadolinium analysis and the Hydrogen analysis. The specific concerns and backgrounds of both analyses are presented, as well as the different applied selections. The development of a method to reject both accidental and correlated background is particularly developed. This method has been one of the developments of this thesis for the Double Chooz analyses. The computation of the analyses systematics and the extraction of θ_{13}, with a rate+shape and a Reactor Rate Modulation fits, are also detailed.

The measurement of the remaining correlated background has been one of the tasks performed during this thesis. It is presented in the sixth and last chapter for both Gadolinium and Hydrogen analyses. We performed a cross-check of the official measurement for the Gadolinium analysis, and the main measurement of the correlated background rate and shape for the Hydrogen analysis, which are both presented here. Finally, the energy spectrum shape of this background is discussed.

Chapter 1

Neutrino physics, state of the art

1.1 A brief history of the Neutrino research

Neutrino history is strongly linked to the radioactivity studies. In 1896, the α radioactivity was discovered by H. Becquerel, and in 1900, P. Villard discovered the γ radioactivity. Both were identified as two-body reactions, leading to a discrete energy spectrum. In 1914, J. Chadwick discovered the β^- radioactivity. This reaction was supposed to be a two-body reaction:

$$^A_ZX \rightarrow ^A_{Z+1}Y + e^- \tag{1.1}$$

However, the fact that it produces a continuous electron spectrum was puzzled. Several theories and experiments were developed and performed to understand it, without any success. This led, in 1924, to N. Bohr's suggestion that the energy was not conserved.

In 1930, W. Pauli proposed the existence of a new particle called "neutron". He postulated that this particle needed to be be very light ($< 0.01 \times m_{proton}$), to be neutral and to have a $\frac{1}{2}$ spin [1]. Pauli's neutron is, in fact, our nowdays neutrino. In 1932, J. Chadwick demonstrated the existence of another neutral particle, with a mass close to the proton one, following the results of I. and F. Joliot-Curie experiment with α particles bombardement on $^9_4$Be. This new particle was also called "neutron". Chadwick's neutron is the one we know today.

In order to fix the naming issue between both particles, in 1933, E. Fermi proposed to rename Pauli's neutron "neutrino" ("little neutron" in Italian). In the end of 1933, I. and F. Joliot-Curie discovered the β^+ radioactivity, with the emission of a e^+ instead of a e^-. In 1934, E. Fermi proposed a theory of Beta Decay [2] including a neutrino emission, allowing to write:

$$^A_ZX \rightarrow ^A_{Z+1}Y + e^- + \overline{\nu} \quad \text{for } \beta^- \text{ decays} \tag{1.2}$$

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z-1} Y + e^{+} + \nu \quad \text{for } \beta^{+} \text{decays} \tag{1.3}$$

1.1.1 Neutrino detection

In order to detect ν, the Inverse Beta Decay (IBD) reaction was proposed in 1934:

$$\overline{\nu} + p \rightarrow p + e^{+} \tag{1.4}$$

However, its cross-section $\sim 10^{-44}$ cm^2 was so low that many physicists were not convinced by the possibility to use it.

The atomic era, starting at the end the 2nd World War with the tragic events of Nagasaki and Hiroshima, led to the use of nuclear reactor cores in order to produce electricity. Many nuclei produced by fission are β^{-} emitters, and therefore emit $\overline{\nu}$. This makes nuclear reactor cores one of the most abundant sources of neutrinos on Earth (more details about it in section 2.5).

Taking advantage of this technologic development, in 1953, F. Reines and C.L. Cowan built a first neutrino detection experiment close to a Hanford nuclear reactor. The detector was a liquid scintillator doped with Cadmium nuclei. This detector was able to detect the scintillation lights from the positron annihilation and from the neutron capture on Cadmium nuclei. They observed an excess of events when the reactor was running, compared to the period when the reactor was off. However, the statistic was not sufficient to claim the detection of neutrinos [3]. A second experiment, in 1956, near the Savannah River Plant was built and was able to confirm the detection of neutrinos [4].

In the early 1960s, an experiment which can be consider as the first accelerator neutrino experiment was performed in Brookhaven National Laboratory, suggested independently by B. Pontecorvo [5] and M. Schwartz [6]. This experiment used neutrino produced together with muons (μ) from pion (π) decays. B. Pontecorvo suggested the possibility that these "muon" neutrinos were different from the neutrino from β decays [5]. A tentative to detect these "muon" neutrinos with the following reactions was performed:

$$\nu + n \rightarrow p + e^{-} \text{ or } \mu^{-} \tag{1.5}$$

$$\overline{\nu} + p \rightarrow p + e^{+} \text{ or } \mu^{+} \tag{1.6}$$

The experiment resulted in the detection of an excess of muon emissions above the background, and without excess of electron emissions [7]. This experiment was the demonstration of the lepton number conservation proposed in 1953 [8].

Finally, the τ neutrino and the τ charged lepton were discovered in 2000 by the DONUT collaboration [9]. Hints on their existence were previously brought in 1975 by the $e^{+}e^{-}$ collider at SLAC [10].

In the 1990s, the LEP experiments ALEPH, DELPHI, L3 and OPAL [11] determined the number of light neutrinos by looking at the Z^0 decays. Z^0 decays into quarks or charged leptons or neutrino-antineutrino pair. Whereas the quarks or the charged leptons can be detected, the neutrino-antineutrino pair is invisible. However, the decay can be "observed" through the observation of a missing energy. Therefore, the ratio between the visible and the invisible decays can be measured experimentaly. The comparison between the experimental and the theoritical ratios allowed to measure the number of light neutrinos:

$$N_\nu = 2.9840 \pm 0.0082 \tag{1.7}$$

Therefore, the final picture of the active neutrino flavors should be: ν_e, ν_μ and ν_τ. If other flavors of neutrino exist, they should not be able to interact under the weak interaction.

1.1.2 Anomalies

Neutrino research history is a history of anomalies. After the continuous β spectrum anomaly, which led to the discovery of the neutrino, the next discovery occured thanks to a new anomaly.

The Homestake experiment was an experiment aiming to measure the neutrino flux from the sun. Fusion reactions inside our star lead to the generation of several ν_e in both fusion cycles (*pp*-cycle and CNO-cycle). The Standard Solar Model (SSM) allowed (and allows) to provide an estimation of the different ν_e flux. In order to detect ν_e, the following reaction was used:

$$\nu_e +^{37} Cl \rightarrow^{37} Ar + e^- \tag{1.8}$$

In 1968, the Homestake experiment discovered a deficit of ν_e compared to SSM prediction [12]. This deficit was confirmed by two other radiochemical experiments: GALLEX [13] and SAGE [14] in the 1990s. Different experiments confirmed these results in the following decades. This was called the solar anomaly.

In 1988, another anomaly was discovered. Kamiokande experiment, looking at neutrino produced by cosmic rays, observed a deficit of ν_μ when compared with the prediction [15]. In 1998, Super-Kamiokande confirmed this deficit [16, 17]. However, as a water Cherenkov detector, Super-Kamiokande is able to determine the direction of the ν_μ. Since neutrinos interact with a very low cross-section, ν_μ can travel across the Earth and interact in a detector on the opposite side of the planet. Super-Kamiokande demonstrates that there was a discrepancy between the number of ν_μ events coming from upside and coming from downside. FIG. 1.1 shows the number of ν_μ events (μ-like) and of ν_e (e-like) as a function of the zenithal angle.

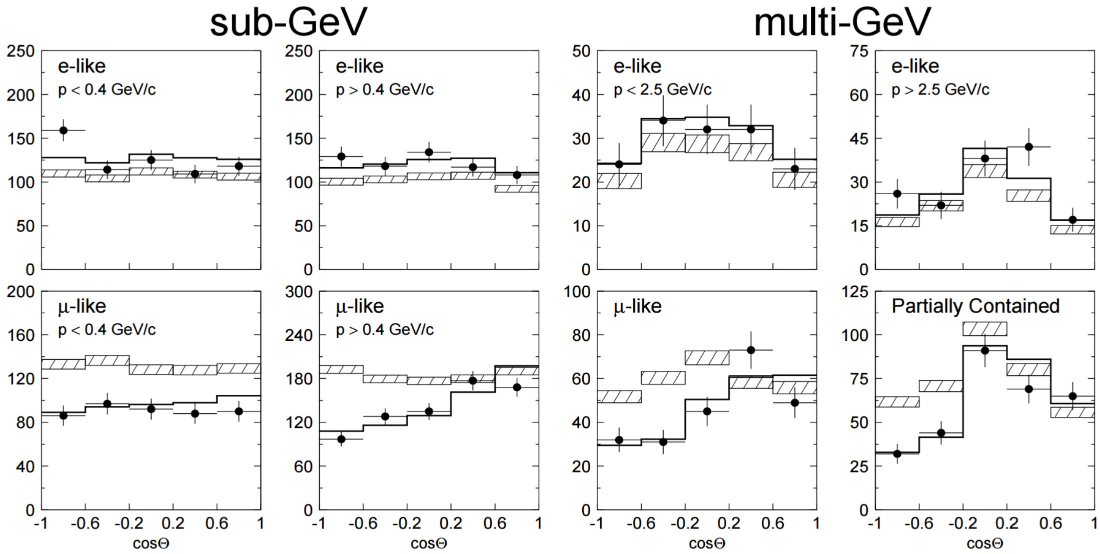


FIGURE 1.1: Super-Kamiokande zenithal angle distributions of μ-like and e-like events for sub-GeV and multi-GeV data sets from [16]. Upward-going particles have $\cos\theta < 0$ and downward-going particles have $\cos\theta > 0$. The hatched region shows the Monte Carlo expectation for no oscillation normalized to the data live-time. The bold line shows the best fit assuming $\nu_\mu \leftrightarrow \nu_\tau$ oscillations.

One would expect a ratio close to unity. This result was the first demonstration of the theory of neutrino oscillations, proposed by B. Pontecorvo, Z. Maki, M. Nakagawa and S. Sakata in 1957 and 1962, respectively [18, 19].

The solar anomaly was finally solved in 2002, by the SNO experiment [20]. SNO used both Neutral Current (NC) and Charged Current (CC) to detect neutrinos (see section 1.2.1). Whereas NC is sensitive to all neutrino flavors, CC is here only sensitive to ν_e. Other flavor neutrinos need to carry an energy higher than the one carried by solar neutrinos to be able to produce their associated charged lepton and interact through CC. Using CC, SNO confirmed the deficit observed by the previous experiments. However, using NC, SNO was able to detect the three flavors of neutrinos and demonstrated that the neutrino flux was consistent with the SSM expectation. This result was another demonstration of the oscillation theory.

1.2 Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a theory describing strong, weak and electromagnetic interactions. It describes elementary particles and their interactions in the context of the quantum field theory. The SM is an unification between quantum physics and special relativity. The SM does not describe the gravitationnal interaction.

The SM is a gauge theory under the local symetries group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where C, L and Y refer to the color charge, the left-handed chirality and the weak hypercharge, respectively. Each group owns a certain number of independent generators, called vector gauge bosons. Lie algebra allows to determine that each $SU(n)$ and $U(n)$ group owns $n^2 - 1$ and n^2 independent generators, respectively, where n is the group dimension. The 8 generators of $SU(3)_C$ group are the gluons, which are the force carriers of the strong interaction. The 3 generators of $SU(2)_L$ group are the bosons $W^\pm$ and Z^0 bosons, which are the force carriers of the weak interaction. The generator of $U(1)_Y$ group is the photon (γ), which is the force carrier of the electromagnetic interaction.

For mathematical reasons, boson in a gauge theory are described as massless particles. Since the range of an interaction depends on the mass of its vector, this implies that the interactions carried by the bosons are long-ranged. However, due to the fact the weak interaction has a very short range, the electroweak theory states that weak interaction bosons masses have to be non-zero. This requires a spontaneous breaking of a symmetry. The spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry through the Brout-Englert-Higgs mechanism allows to explain the masses of the $W^\pm$ and Z^0 bosons. The combination of the weak interaction gauge group, the electromagnetic interaction gauge group and the Brout-Englert-Higgs mechanism, is known as the Glashow-Weinberg-Salam model, or electroweak theory [21, 22, 23]. This model successfully predicted the $W^\pm$ and Z^0 masses.

The SM describes two kinds of elementary particles, fermions, which are the matter components, and bosons, which are the force carriers. Fermions are half-integer spin particles which are characterized by Fermi-Dirac statistics. They obey the Pauli exclusion principle. Bosons are integer spin particles characterized by Bose-Einstein statistics.

Fermions are divided into two sub-categories, the quarks and the leptons. Each category has three families[1] which are described in TAB. 1.1. Quarks are subject to all interactions, whereas leptons are not subjet to strong interaction. Quarks cannot be observed directly since they are always confined into hadrons, except in theoritical states of matter, like Quark-Gluon Plasma [24]. Hadrons are arranged with three quarks (baryons) or a quark/antiquark pair (mesons). Left-handed fermions form doublet of $SU(2)_L$, whereas right-handed fermions form singlet. This implies that right-handed fermions do not interact with weak interaction bosons. As a neutral particle does not interact through the electromagnetic interaction, the model does not consider right-handed neutral leptons (right-handed neutrino). Therefore, due to the Brout-Englert-Higgs mechanism neutrinos are

---

[1] The three families are sometimes called *generations*, the word *flavors* is also used.

Fermions	1^{st} family	2^{nd} family	3^{rd} family
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$
	u_R, d_R	c_R, s_R	t_R, b_R
Leptons	$\begin{pmatrix} \mathrm{e}^- \\ \nu_e \end{pmatrix}_L$	$\begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}_L$	$\begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}_L$
	e^-_R	μ^-_R	τ^-_R

Table 1.1: Fermions described by Standard Model. Particles in the three families have same properties, except mass.

not considered as massive particles in the Standard Model. Up to now, there is no experimental evidence of the existence of a right-handed neutrino. However, according to [25] the right-handed neutrinos could provide explanation to several beyond Standard Model experimental evidences.

1.2.1 Neutrino interaction with matter

As neutrinos are singlet over all standard interactions exept weak interaction. Neutrinos interact with matter only through weak interactions. Two kinds of interaction exist:

- Neutral current
- Charged current

Neutral current is the exchange of a Z^0 boson between the neutrino and another particle. It transfers some of the neutrino energy and momentum to the target particle. In case of a light and charged target particle, such as an electron, it may be accelerated to a relativistic speed and then emit a Cerenkov radiation, which can be detected. The three neutrino flavors can interact through neutral current. However, this detection does not give information on the neutrino flavors.

Charged current is the exchange of a $W^{\pm}$ boson between the neutrino and another particle. It transforms the neutrino into its charged lepton partner. However, the interaction energy needs to be enough to create the neutrino heavier partner mass. Due to the exchange of a charged boson, the target particle also changes character to respect reaction symmetries.

Both currents are illustrated by the Fig. 1.2:

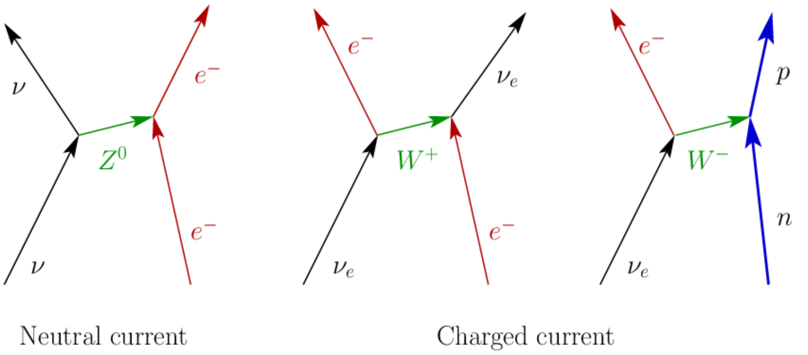


FIGURE 1.2: Neutrino scattering diagrams. In the first case, the neutrino transfered only energy and momentum to the target particle, through the exchange of a Z^0 boson. In the second case, the electron neutrino interacts through the exchange of a W boson, and converts into the equivalent charged lepton (an electron here), either by transforming an equivalent charged lepton into a neutrino, or by transforming a neutron into a proton.

1.3 Neutrino property: Oscillation

Neutrino oscillations were firstly proposed by B. Pontecorvo in 1957 [18], who introduced the possibility of $\nu \leftrightarrow \overline{\nu}$ oscillations by analogy to $K^0 \leftrightarrow \overline{K}^0$ oscillations. Neutrino-Antineutrino oscillations are still not observed but this idea was the foundation of a theory of neutrino flavor oscillations, which was developed in 1962 by Maki, Nakagawa, and Sakata [19] and completed by Pontecorvo in 1967 [26]. According to the theory, the oscillation phenomenom implies that at least two neutrinos are massive.

Since the Super-Kamiokande results in 1998 [16], the existence of neutrino flavor oscillations has been demonstrated.

1.3.1 Theory

The theory developed by Maki, Nakagawa, Sakata, and Pontecorvo states that the neutrino flavor eigenstates do not correspond to the neutrino mass eigenstates. Therefore, flavor eigenstates ν_α are a mixing of mass eigenstates ν_i. ν_α can be then represented as a linear combination of ν_i. The PMNS matrix allows to link the neutrino flavor eigenstates and the neutrino mass eigenstates:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.9}$$

where U_{PMNS} stands for the PMNS matrix. It is analogue to the CKM matrix, which acts on the quark flavor oscillations due to the weak interaction. The U_{PMNS} can be written as follows:

$$U_{PMNS} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \tag{1.10}$$

U_{PMNS} can be parametrized and developed as three matrix, representing the three oscillation sectors that will be developed in the following. The parametrization is done with the use of three angles, called *neutrino mixing* angles, and one phase, called CP-violating phase, δ. The parametrized matrix can be written as follows:

$$U_{PMNS} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}\mathrm{e}^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}\mathrm{e}^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.11}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$, with θ_{ij} the neutrino mixing angles. In the case of Majorana particles (cf. section 1.4.1), a fourth matrix shoud be added:

$$D_{Maj} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{i\alpha_1} & 0 \\ 0 & 0 & \mathrm{e}^{i\alpha_2} \end{pmatrix} \tag{1.12}$$

where α_1 and α_2 are two additionnal phases.

The first matrix parametrized the so-called atmospheric-accelerator sector, the second one the reactor-accelerator sector and the third one the solar-reactor sector. The mixing angles, θ_{ij}, characterizes the oscillation amplitudes. The CP-violating phase, δ, characterizes the matter-antimatter asymmetry in the leptonic sector, which could explain the observed matter-antimatter in the universe.

1.3.1.1 Oscillations in the vacuum

As neutrino flavor eigenstates are a linear combination of the mass eigenstates, it is possible to rewrite Eq. 1.9 as:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^*_{\alpha i} |\nu_i\rangle \tag{1.13}$$

where α stands for e, μ or τ, and $U^*_{\alpha i}$ is an element of the U_{PMNS} matrix. As U_{PMNS} should be unitary, it is possible to write:

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

where i stands for 1, 2 or 3, and α for e, μ or τ.

The mass eigenstates $|\nu_i\rangle$ are eigenstates of the Hamiltonian $\widehat{H}_0|\nu_i\rangle = E_i|\nu_i\rangle$, where $E_i = \sqrt{p_i^2 + m_i^2}$, with p_i and m_i respectively being the neutrino mass eigenstate i momentum and rest mass. Therefore, we have the solution of $|\nu_i(x,t)\rangle$:

$$|\nu_\alpha(x,t)\rangle = \mathrm{e}^{-i(E_i t - p_i x)}|\nu_i(0,0)\rangle \tag{1.15}$$

with x the position vector of the eigenstate. Then, it is possible to write the evolution of $|\nu_\alpha(x,t)\rangle$ as:

$$|\nu_\alpha(x,t)\rangle = \sum_{i=1}^{3} U^*_{\alpha i}\mathrm{e}^{-i(E_i t - p_i x)}|\nu_i\rangle \tag{1.16}$$

Using Eq. 1.14, we can write:

$$|\nu_\alpha(x,t)\rangle = \sum_{\beta=e,\mu,\tau}\left(\sum_{i=1}^{3} U^*_{\alpha i}\mathrm{e}^{-i(E_i t - p_i x)}U_{\beta i}\right)|\nu_i\rangle \tag{1.17}$$

This means that after its generation ($t > 0$), the neutrino in $|\nu_\alpha\rangle$ flavor state becomes a mixing of different flavor states. Therefore, it is possible to compute the probability to detect a $|\nu_\beta\rangle$ flavor state from a $|\nu_\alpha\rangle$ initial state.

$$P_{\nu_\alpha\to\nu_\beta} = |\langle\nu_\beta|\nu_\alpha(x,t)\rangle|^2 = \sum_{i,j=1}^{3} U^*_{\alpha i}U_{\beta i}U^*_{\alpha j}U_{\beta j}\mathrm{e}^{-i([E_i - p_i x]-[E_j - p_j x])(t)} \tag{1.18}$$

If we assume an ultrarelativist case, $E \simeq p_i >> m_i$, and consider $t = x = L$, where L the distance between the source and the detector, we can write $E_i \backsimeq E + \frac{m_i^2}{2E}$. The probability $P_{\nu_\alpha\to\nu_\beta}$ can be then written as:

$$P_{\nu_\alpha\to\nu_\beta} = \sum_{i,j=1}^{3} U^*_{\alpha i}U_{\beta i}U^*_{\alpha j}U_{\beta j}\exp\left(-i\frac{\Delta m^2_{ij}L}{2E}\right) \tag{1.19}$$

where $\Delta m^2_{ij} = m_i^2 - m_j^2$. Three mass-squared differences can be considered:

$$\Delta m^2_{21} = m_2^2 - m_1^2 \;\;;\;\; \Delta m^2_{32} = m_3^2 - m_2^2 \;\;;\;\; \Delta m^2_{31} = m_3^2 - m_1^2 \tag{1.20}$$

Therefore, we have the following relation between them:

$$\Delta m^2_{21} + \Delta m^2_{32} - \Delta m^2_{31} = 0 \tag{1.21}$$

Therefore, if oscillations happen, at least two neutrino masses have to be > 0. Currently, the ordonnance of the neutrino masses, i.e. the mass hierarchy, is not known. Thanks to current experimental results, we can reduce this mass hierarchy issue to two scenarii: the Normal Hierarchy (NH) and the Inverted Hierarchy (IH). The NH scenario is when the masses are ordered in a conventionnal way, with m_1 the lightest one: $m_1 < m_2 < m_3$. The IH scenario is when the lightest mass is m_3: $m_3 < m_1 < m_2$.

1.3.1.2 Oscillations in the matter

In the matter neutrinos can only interact via weak interactions (see section 1.2.1). Electron neutrinos can interact through Charged and Neutral Currents with electrons. Whereas muon and tau neutrinos interact by NC and need to carry enough energy to create a μ or a τ in order to interact through CC. Therefore the refraction is different for ν_e from the other neutrino flavors. This is described by the following potential [27]:

$$V = \sqrt{2} G_F N_e \tag{1.22}$$

where G_F is the Fermi coupling constant and N_e the electron density in matter. When considering antineutrino, this potential V will have an opposite sign.

1.3.1.3 Oscillations: Two-flavor case

Vacuum oscillations

It is possible to simplify the oscillation formalism in a two-flavor case. This approximation is valid for several experiments. Assuming two flavor eigenstates $|\nu_\alpha\rangle$ and $|\nu_\beta\rangle$, and two mass eigenstates $|\nu_1\rangle$ and $|\nu_2\rangle$, we can express EQ. 1.9 for oscillations in vacuum as:

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{1.23}$$

with θ the mixing angle considered. We can then express the flavor eigenstates as:

$$\begin{aligned} |\nu_\alpha\rangle &= \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \\ |\nu_\beta\rangle &= -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle \end{aligned} \tag{1.24}$$

which leads to the time evolution of $|\nu_\alpha\rangle$:

$$|\nu_\alpha(t)\rangle = \cos\theta\, \mathrm{e}^{-iE_1 t}|\nu_1\rangle + \sin\theta\, \mathrm{e}^{-iE_2 t}|\nu_2\rangle \tag{1.25}$$

The disappearance probability is then, in the ultrarelativist assumption:

$$P_{\nu_\alpha \to \nu_\beta} = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \tag{1.26}$$

where Δm^2 is the mass-squared difference, L the distance between the source and the detector and E the neutrino energy. Since we are in a two flavors oscillation case, we have:

$$P_{\nu_\alpha \to \nu_\beta} + P_{\nu_\alpha \to \nu_\alpha} = 1 \tag{1.27}$$

Which allows to write the survival probability as:

$$P_{\nu_\alpha \to \nu_\alpha} = 1 - \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \tag{1.28}$$

Using physic units, we can write it as:

$$P_{\nu_\alpha \to \nu_\alpha} = 1 - \sin^2(2\theta)\sin^2\left(1.27\frac{\Delta m^2(eV^2)L(m)}{E(MeV)}\right) \tag{1.29}$$

Oscillations in matter

Differentiating EQ. 1.15 with respect to time, we can write the following time development Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \widehat{H}_0 \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \tag{1.30}$$

with the common phase factor e^{ipx} ignored. In vacuum, the Hamiltonian $\widehat{H}_0$ is:

$$\widehat{H}_0 = \frac{1}{2E}\begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix} \tag{1.31}$$

It is also possible to write EQ. 1.30 in a flavor basis:

$$i\frac{d}{dt}\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = U^\dagger\, i\frac{d}{dt}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \widehat{H}_0 \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \widehat{H}_0 U^\dagger \begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} \tag{1.32}$$

or, multiplying by U:

$$i\frac{d}{dt}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = U\widehat{H}_0 U^\dagger \begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} \tag{1.33}$$

If we note, $\widehat{H}_f$ the transformed Hamiltonian in the vacuum, after trivial trigonometric computations, we have:

$$\widehat{H}_f = U\widehat{H}_0U^\dagger = \frac{m_1^2+m_2^2}{4E}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\Delta m^2}{4E}\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \tag{1.34}$$

where $\Delta m^2 = m_1^2 - m_2^2$. In the following, we will neglect $\frac{m_1^2+m_2^2}{4E}\mathbb{1}_2$. For oscillations in matter, we need to add the potential V to the transformed Hamiltonian:

$$i\frac{d}{dt}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = (\widehat{H}_f + V)\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} \tag{1.35}$$

This leads, after some trivial developments, to:

$$i\frac{d}{dt}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \frac{\Delta m^2}{4E}\begin{pmatrix} -\cos 2\theta + V & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} \tag{1.36}$$

After diagolization of this matrix, we can find:

$$\sin^2(2\theta_m) = \frac{\sin^2(2\theta)}{\left(\cos(2\theta) - \frac{EV}{\Delta m^2}\right)^2 + \sin^2(2\theta)} \tag{1.37}$$

$$\Delta m_m^2 = \Delta m^2\sqrt{\left(\frac{V}{\Delta m^2} - \cos 2\theta\right) + \sin^2 2\theta} \tag{1.38}$$

This allows to write the oscillation probability in matter (in unit physics) as:

$$P_{\nu_\alpha\to\nu_\beta} = \sin^2(2\theta_m)\sin^2\left(1.27\frac{\Delta m_m^2(eV^2)L(m)}{E(MeV)}\right) \tag{1.39}$$

It is interesting to rewrite Eq. 1.37 as:

$$\tan(2\theta_m) = \frac{\frac{\Delta m^2}{2E}\sin(2\theta)}{\frac{\Delta m^2}{2E}\cos(2\theta) - V} = \frac{\tan(2\theta)}{1 - \left(\frac{L}{L_m\cos(2\theta)}\right)} \tag{1.40}$$

with $L_m = \frac{2\pi}{V}$ and $L = 2\pi\frac{2E}{\Delta m^2}$. Then, a resonance effect occurs when $\frac{L}{L_m} = \cos(2\theta)$ happens, where the transition between both flavors is maximal.

1.3.2 Oscillation parameter measurement

Several experiments were built to measure the different oscillation parameters. Two kind of experiments can be separated: experiments looking at a deficit of ν_α from a source of ν_α, which are called *disappearance* experiments, and experiments looking at an exces of ν_β from a source of ν_α, which are called *appearance* experiments.

1.3.2.1 Experimental measurement of θ_{12} and Δm^2_{12}

The experimental measurement of θ_{12} and Δm^2_{12}, the so-called "Solar sector", can be made with experiments looking at solar neutrinos or with very long baseline reactor neutrinos experiments.

Fusion reactions within the Sun can generate electron neutrinos. Two types of fusion cycles exist within the Sun: the *pp* cycle and the *CNO* cycles [28]. Both generate several electron neutrinos. Solar experiments measure the rate and the energy of electron neutrinos arriving on Earth. From these measurements, a deficit is observed with respect to the prediction. It is then possible to extract the oscillation parameters.

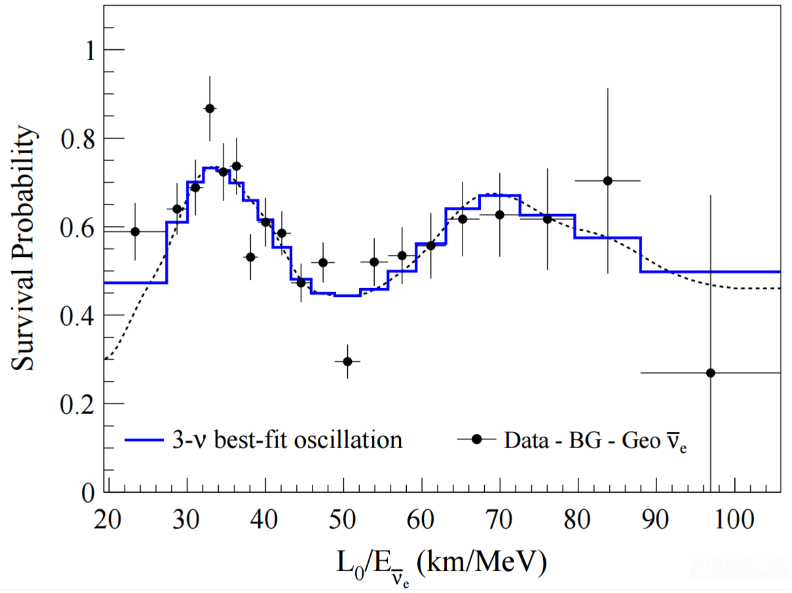


FIGURE 1.3: Ratio of the observed e spectrum to the expectation for no-oscillation as function of L_0/E for the KamLAND data from [29]. $L_0 = 180$ km is the flux-weighted average reactor baseline.

The current best precision on the "Solar sector" oscillation parameters is held by the KamLAND experiment since 2008. KamLAND is an experiment looking for electron antineutrino produced by nuclear reactor cores (cf. section 2.5). The neutrino sources are the 55 Japanese nuclear reactors, which are located in a flux-weighted average distance of 180 km. It is a 1 kton liquid scintillator detector located at Kamiokande experimental site. KamLAND detection principle relies on the Inverse Beta Decay (IBD) (cf. section 2.1). Due to its large E/L ratio, KamLAND is sensitive to the solar sector. In 2013, taking advantage of the long-term shutdown of commercial nuclear reactors in Japan, following the Fukushima disaster in 2011, KamLAND was able to improve its background constraints, leading to the following measurement [29]:

$$\tan^2\theta_{12} = 0.481^{+0.092}_{-0.080} \tag{1.41}$$

$$\Delta m^2_{12} = 7.54^{+0.19}_{-0.018}\,\mathrm{eV}^2 \tag{1.42}$$

Fig. 1.3 shows the ratio of observed electron antineutrinos over expected one for the KamLAND experiment.

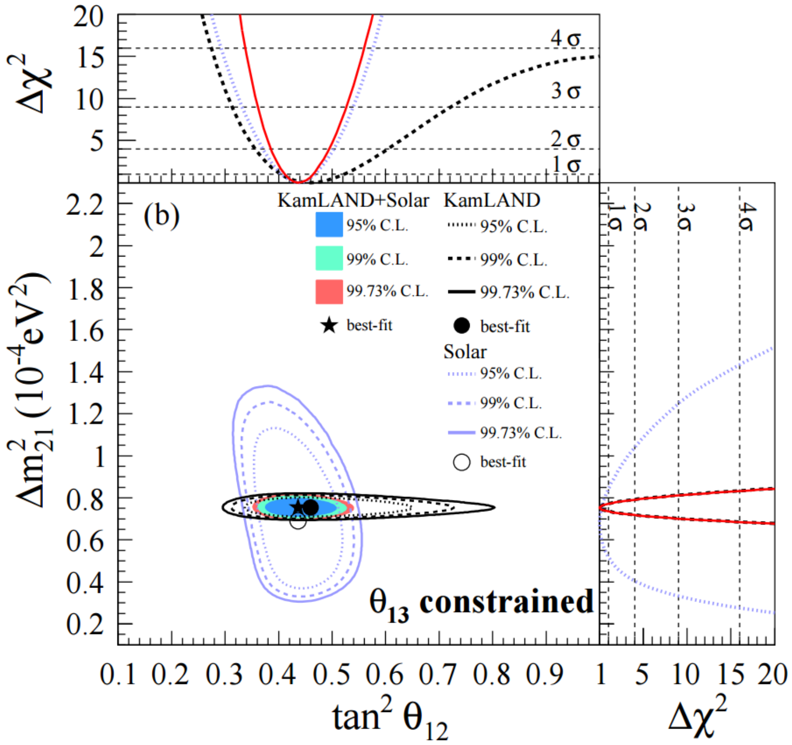


Figure 1.4: Allowed regions projected in the ($\tan^2\theta_{12}$, Δm^2_{21}) plane from [29], for solar and KamLAND data from the three-flavor oscillation analysis for θ_{13} constrained by accelerator and shortbaseline reactor neutrino experiments.

A combination of KamLAND results with solar experiments, short baseline reactor experiments and accelerator experiments, was done in 2011 [30] and in 2013. This combination assumed an invariance over the CPT symmetry, as solar experiments look at electron neutrino, whereas KamLAND looks at electron antineutrino. Fig. 1.4 shows the contours of the combination on both solar parameters. Whereas KamLAND has the best sensitivity on Δm^2_{12}, its sensitivity on $\tan^2\theta_{12}$ is lower than for solar experiments. It leads to an improvement of the precision [29]:

$$\tan^2\theta_{12} = 0.436^{+0.029}_{-0.025} \tag{1.43}$$

$$\Delta m^2_{12} = 7.53 \pm 0.18 \times 10^{-5}\,\mathrm{eV}^2 \tag{1.44}$$

1.3.2.2 Experimental measurement of θ_{23} and Δm^2_{23}

The experimental measurement of θ_{23} and Δm^2_{23}, the so-called "Atmospheric sector", can be made with experiments looking at atmospheric neutrinos or with long baseline accelerator neutrino experiments.

Cosmic ray interactions within the atmosphere generate high density particle showers. Pions can be created within such showers and can decay, emitting a muon and a muon (anti)neutrino. Experiments like Super-Kamiokande [31], IceCube-DeepCore [32] or MINOS [33], can detect such (anti)neutrinos and measure their flux in order to extract the oscillation parameters. Accelerators can also generate beams of ν_μ ($\overline{\nu}_\mu$), which are used by several experiments like NOνA [34], T2K [35] or MINOS [36]. Several atmospheric experiments can also measure these accelerator (anti)neutrinos, like Super-Kamiokande detector in the T2K, or MINOS. Presently, the best precision is held by MINOS and T2K experiments.

MINOS uses beams from NuMI facility in the Fermilab accelerator complex and detects them 735 km away with its far detector, MINOS+, located in the Soudan mine. It uses a near detector to normalize the neutrino flux coming from the neutrino beam. MINOS is able to detect ν_μ and $\overline{\nu}_\mu$ separatly. MINOS can also perform an appearance analysis, and is able to detect ν_e and $\overline{\nu}_e$ separatly. This experiment measures both accelerator and atmospheric (anti)neutrinos at the same time. MINOS last results for θ_{23} and Δm^2_{23} are [36]:

$$\text{N.H.:}\quad \sin^2\theta_{23} = 0.35 - 0.65\,(90\%\,\text{C.L.}) \tag{1.45}$$

$$\text{N.H.:}\quad |\Delta m^2_{23}| = [2.28 - 2.46] \times 10^{-3}\,\mathrm{eV}^2\,(68\%\,\text{C.L.}) \tag{1.46}$$

$$\text{I.H.:}\quad \sin^2\theta_{23} = 0.34 - 0.67\,(90\%\,\text{C.L.}) \tag{1.47}$$

$$\text{I.H.:}\quad |\Delta m^2_{23}| = [2.32 - 2.53] \times 10^{-3}\,\mathrm{eV}^2\,(68\%\,\text{C.L.}) \tag{1.48}$$

In the case of the T2K experiment, the $\overline{\nu}_\mu$ are produced at the J-PARC accelerator complex and detected in the 50 ktons Super-Kamiokande water Cherenkov detector at Kamiokande, 295 km away from the accelerator. It uses two near detector (INGRID and ND280) to normalize the neutrino flux coming from the neutrino beam. T2K experiment not yet performed a combination between atmospheric and accelerator data, neither between ν_μ and $\overline{\nu}_\mu$ disappearance analysis nor with $\overline{\nu}_e$ appearance analysis. Its most precise measurements are [37, 38]:

$$\text{N.H.:}\quad \sin^2\theta_{23} = 0.514^{+0.055}_{-0.056} \tag{1.49}$$

$$\text{N.H.:}\quad |\Delta m^2_{23}| = 2.51 \pm 0.10 \times 10^{-3}\,\text{eV}^2 \tag{1.50}$$

$$\text{I.H.:}\quad \sin^2\theta_{23} = 0.511 \pm 0.055 \tag{1.51}$$

$$\text{I.H.:}\quad |\Delta m^2_{13}| = 2.48 \pm 0.10 \times 10^{-3}\,\text{eV}^2 \tag{1.52}$$

Its last results on the $\overline{\nu}_\mu$ disappearance channel were presented recently [39]. A combination between ν_μ and $\overline{\nu}_\mu$ analysis is planned [40].

Fig. 1.5 shows the contours of MINOS and T2K experiments on the atmospheric sector. NOνA and IceCube-DeepCore experiments are expected to provide new results with increased precision in the upcoming years [41, 42].

1.3.2.3 Experimental measurement of θ_{13}

θ_{13} was the last not-known mixing angle until very recently (2011). θ_{13} is called "reactor sector" or "reactor-accelerator sector". It can be measured using short baseline reactor experiments ($\sim$ 1 km) or by accelerator experiments looking at ν_e appearance.

Since 1999, only an upper limit was measured by the CHOOZ experiment [44]. The discovery of a large θ_{13} value was an achievement reached by the three reactor neutrino experiments Double Chooz [47], Daya Bay [48] and RENO [49]. Its large value allows other experiments to access and measure the CP-violating phase, δ_{CP}, which depends on θ_{13}. Indeed, in the PMNS matrix, the δ_{CP} phase only appears as a factor of $\sin\theta_{13}$. Therefore, the δ_{CP} phase would not have been accessible if θ_{13} was 0.

Reactor neutrino experiments

Reactor neutrino experiments are disappearance experiments. They extract θ_{13} measurement from a comparison between an expected $\overline{\nu}_e$ flux in a non-oscillation case and an observed $\overline{\nu}_e$ flux from a distance of $\sim$ 1 km from the reactor core. All reactor neutrino experiments use the Inverse Beta Decay (IBD) to measure $\overline{\nu}_e$ (see section 2.1).

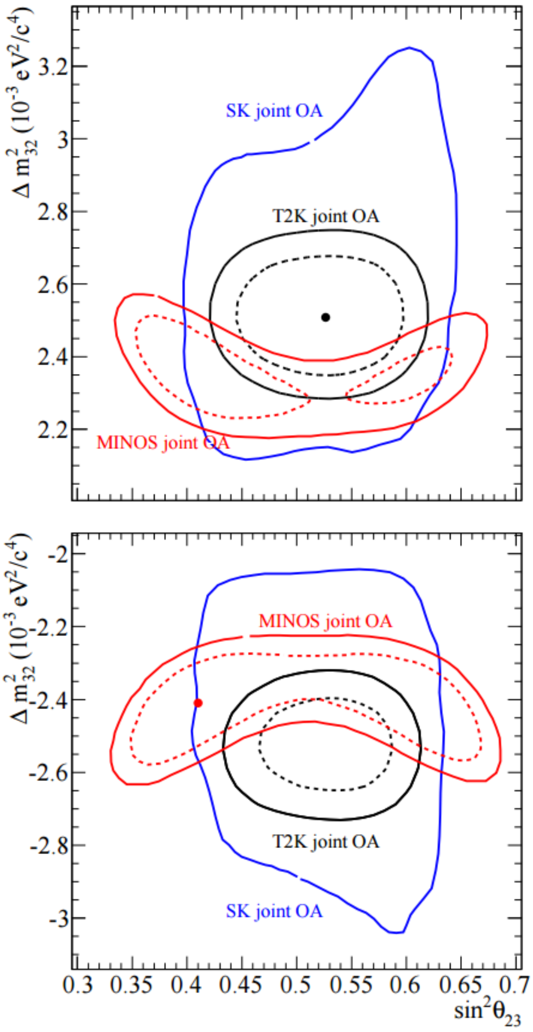


FIGURE 1.5: Contours from [38] showing T2K 68% (dashed) and 90% (solid) CL regions for normal (top) and inverted (bottom) mass hierarchy combined with the results from reactor experiments in the ($\sin^2\theta_{23}$, Δm^2_{32}) space compared to the results from the Super-Kamiokande [43] and MINOS [36] experiments.

The CHOOZ experiment and Double Chooz in its first phase used only one detector to measure the "oscillated" flux. In this case, the expected no-oscillated flux is predicted by nuclear core simulations (more details in section 2.5). New generation reactor experiments (Double Chooz 2nd phase, Daya Bay and RENO) use a minimum of two detectors: a *far* detector, $\sim$ 1km away, to measure the "oscillated" flux and a *near* detector, $\sim$ 400 m away, to measure the flux before oscillation. FIG. 1.6 shows the survival probability of a $\overline{\nu}_e$. The maximum survival probability is at low L, where near detectors are located, and the minimum one at $L \sim 1$ km.

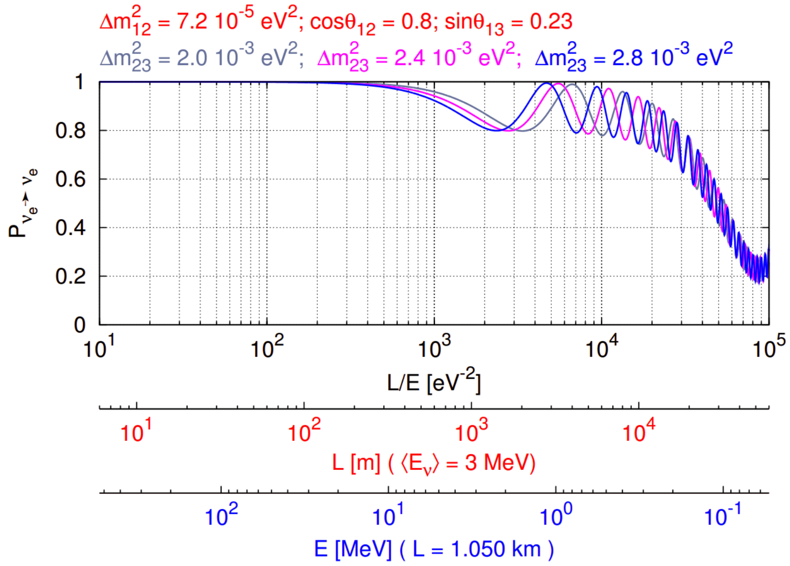


Figure 1.6: $\overline{\nu}_e$ survival probability as a function of L/E, of the mean distance from the reactor cores L (with $E = 3$ MeV) and of the neutrino energy E (with $L = 1050$ m), from [50]. Colors were modified to improve readability.

The design of the different reactor experiments is almost identical for each experiment[2]: a ν-target filled with a Gd-doped liquid scintillator, surrounded by a γ-catcher volume filled with a non-doped liquid scintillator to measure the electromagnetic energy and the γ from the IBDs occuring in the ν-target. A veto, equipped with several PMTs, and optically isolated from the two previous volumes, surround the whole volume to detect cosmic μ backgrounds. In some designs (CHOOZ and Double Chooz) this veto is filled with non-doped liquid scintillator, whereas in other designs (Daya Bay and RENO), it is filled with water.

The main improvement since the CHOOZ experiment is the addition of a non-scintillator buffer volume surrounding the γ-catcher volume, and where PMTs are located. This volume allows to protect the active volumes (ν-target and γ-catcher) from the PMTs radioactivity. It is filled with mineral oil.

More details on the detector design and on the Double Chooz experiment can be found in the next chapter (Chapter 2).

The **CHOOZ** experiment was the first reactor experiment which provides limits on θ_{13}. It was located close to the CHOOZ-B nuclear power plant, at the same

---

[2] The names of the different volumes can be different between each experiments. Here, the Double Chooz naming was used.

position as the present Double Chooz far detector. It ran from April 1997 to July 1998 and was able to perform a direct measurement of the background thanks to the fact that both reactors were not running at the beginning of the data taking. However, the liquid scintillator deteriorated after a few months, leading to the end of the experiment. The CHOOZ experiment was no able to find evidence for neutrino oscillation. It was able to set an upper limit on θ_{13} with one year of live time [44]:

$$\sin^2 2\theta_{13} < 0.17 \tag{1.53}$$

The **Daya Bay** experiment is located close to the Daya Bay, Ling Ao and Ling Ao-II nuclear power plants (China). It consists in one far experimental hall of 4 detectors and of two near experimental halls of 2 detectors each. Daya Bay was able to constrain θ_{13} over the last few years [48, 51, 52]. It provides the current most precise measurement of θ_{13} with its Gadolinium analysis [53, 54]:

$$\sin^2 2\theta_{13} = 0.084 \pm 0.005 \tag{1.54}$$

Using the simplified two-flavor case, Daya Bay has also been able to provide a direct measurement of the Δm^2 terms appearing in the $\overline{\nu}_e$ survival probability [53, 54]:

$$|\Delta m^2_{ee}| = (2.42 \pm 0.11) \times 10^{-3}\,\mathrm{eV}^2 \tag{1.55}$$

FIG. 1.7 shows the Daya Bay contours for $\sin^2 2\theta_{13}$ and $|\Delta m^2_{ee}|$.

Following Double Chooz experiment [55], Daya Bay performed a Hydrogen analysis. This analysis leads to the following results [56]:

$$\sin^2 2\theta_{13} = 0.083 \pm 0.018 \tag{1.56}$$

The **RENO** experiment is located close to the Yonggwang nuclear power plant (South Korea). It consists in one near detector and one far detector located 294 m and 1383 m away from the reactor array center, respectively. RENO last published results leads to [49]:

$$\sin^2 2\theta_{13} = 0.113 \pm 0.013(stat) \pm 0.019(syst) \tag{1.57}$$

Accelerator neutrino experiments

Accelerator neutrino experiments are sensitive to θ_{13} when looking for ν_e ($\overline{\nu}_e$) appearance. The probability of appearance is given by:

$$\begin{aligned} P(\nu_\mu \to \nu_e) &\simeq \sin^2(\theta_{23})\sin^2(2\theta_{13})\sin^2\left(\tfrac{\Delta m^2_{31}L}{4E}\right) \\ &+ \text{CP-violating phase, matter effect term} \end{aligned} \tag{1.58}$$

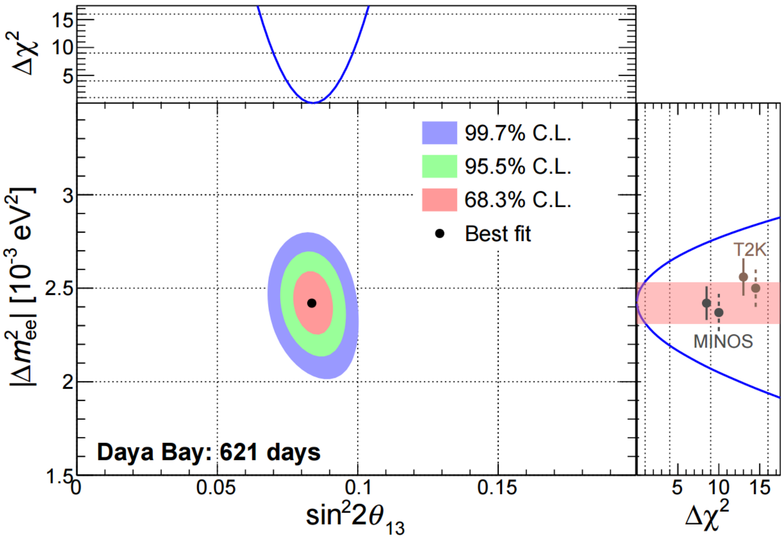


FIGURE 1.7: Regions in the ($|\Delta m^2_{ee}|$,$\sin^2 2\theta_{13}$) plane allowed at the 68.3%, 95.5% and 99.7% CL by the Daya Bay near-far comparison of $\overline{\nu}_e$ rate and energy spectra, from [54].

where L is the distance from the neutrino source and E the neutrino energy.

T2K is currently the only accelerator experiment to provide a θ_{13} measurement. Its last results were [57]:

$$\text{N.H.:} \quad \sin^2 2\theta_{13} = 0.140^{+0.038}_{-0.032} \tag{1.59}$$

$$\text{I.H.:} \quad \sin^2 2\theta_{13} = 0.170^{+0.045}_{-0.037} \tag{1.60}$$

1.3.2.4 Experimental measurement of δ_{CP}

Thanks to the relativly high value of θ_{13} the measurement of δ_{CP} is possible. Accelerator neutrino experiment looking at ν_e ($\overline{\nu}_e$) appearance are currently leading the way for this measurement. NOνA and T2K experiments released recently indications for a non-zero value of δ_{CP} [58, 38]. FIG. 1.8 and 1.9 show the contours of NOνA and T2K experiments, respectively. NOνA disfavors the IH by $\sim 2\sigma$ whereas T2K favors it. However, none of these experiments is currently able to provide a clear answer on either the value of δ_{CP} or the mass hierarchy.

DUNE experiment [61] is a new accelerator experiment planned for the upcoming years. FIG. 1.10 shows the expected sensitivity of this experiment after 10 years.

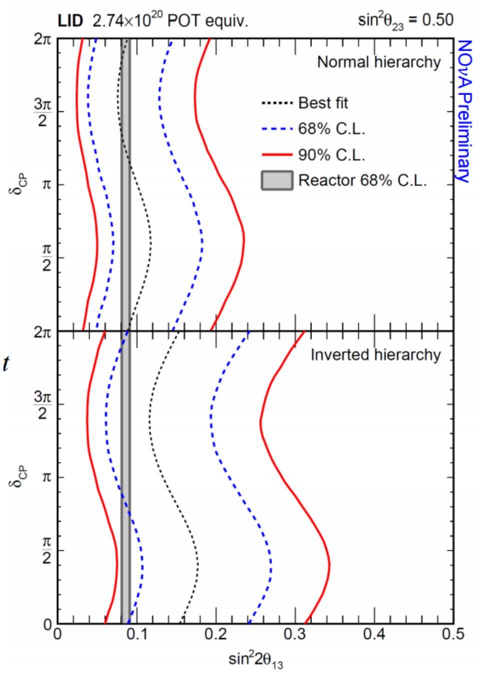


FIGURE 1.8: NOνA contours plot of 68% and 90% CL allowed regions for $\sin^2 2\theta_{13}$, as a function of δ_{CP} assuming normal hierarchy (top) and inverted hierarchy (bottom) from LID method in [58]. The dashed line represents the best fit. $\sin^2\theta_{23}$ was fixed to 0.50. The shaded region shows the average θ_{13} value from the reactor measurement.

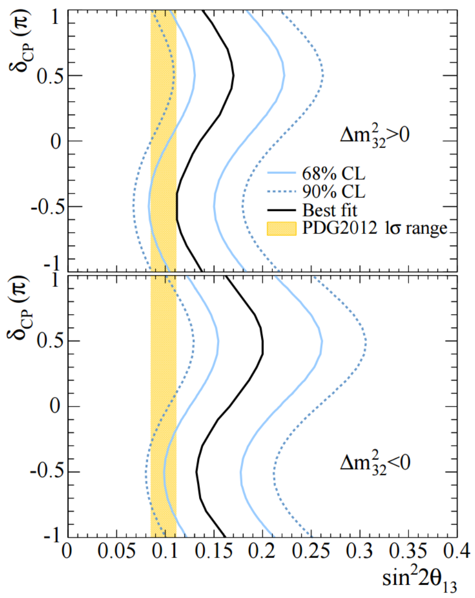


FIGURE 1.9: T2K contours plot of 68% and 90% CL allowed regions for $\sin^2 2\theta_{13}$, as a function of δ_{CP} assuming normal hierarchy (top) and inverted hierarchy (bottom) from [38]. The solid line represents the best fit. The values of $\sin^2\theta_{23}$ and Δm^2_{32} are varied in the fit with the constraint from [59]. The shaded region shows the average θ_{13} value from the PDG2012 [60].

δ_{CP} measurement is strongly correlated to the mass hierarchy measurement. The mass hierarchy can be determine independently from the δ_{CP} value (see section 1.4.2.3).

1.3.2.5 Anomalies: leading to sterile flavors?

Anomalies discovered in the previous century lead to the discovery of neutrino oscillations (see section 1.1). Since this demonstration, new anomalies, which do not fit in the current 3ν oscillation paradigme, were found by several experiments.

Several theorists suggest the introduction of additionnal neutrino flavors inside the model to explain these anomalies. As it was demonstrated that only three

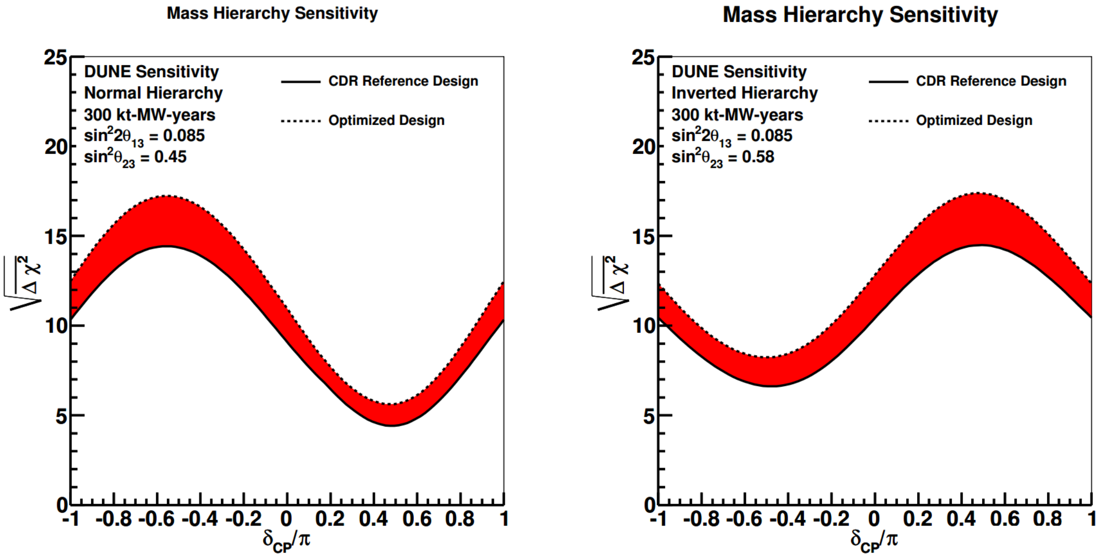


FIGURE 1.10: DUNE sensitivity to the mass hierarchy determination as a function of δ_{CP} for the N.H. (left) and I.H. (right), from [61]. The dashed and the solid lines show the different ν beam configuration.

active neutrinos exist by LEP experiments [11], these new flavors should be singlet under all the SM interactions and interact only through the neutrino oscillation and via the gravitationnal interaction due to their masses. Therefore, these new flavors are called *sterile* flavors.

Different models are proposed, some suggesting to add 1 or more new *sterile* flavors (called $3+1$ or $3+n_r$ *minimal* models) [62].

Accelerator anomalies

In 1996, the LSND experiment observed an excess of event in its $\overline{\nu}_\mu \to \overline{\nu}_e$ analysis. This excess can be interpreted as an oscillation. However, this oscillation would require a Δm^2 larger than the one observed for "standard" oscillations [63]:

$$\Delta m^2_{LSND} > 0.1\,\mathrm{eV}^2 \tag{1.61}$$

MiniBooNE experiment was designed to cross-check LSND results and analyze both $\overline{\nu}_\mu \to \overline{\nu}_e$ and $\nu_\mu \to \nu_e$ oscillations. MiniBooNE results lead to the observation of an excess with $\Delta m^2 \simeq 1$ eV^2 [64]. However, MiniBooNE observed the excess at a lower energy than LSND, and was then not able to confirm nor to refute the LSND results.

It is interesting to notice that the neutrino production methods used in LSND and in MiniBooNE are not the same. LSND neutrino production was done with μ

decays at rest method, whereas MiniBooNE used a proton on target method with the Booster Neutrino Beamline beam. The main advantage of MiniBooNE method is the ability to generate both $\overline{\nu}_\mu$ and ν_μ, whereas the μ decay at rest does not allow to generate ν_μ. JSNS^2 experiment [65], in Japan, and OscSNS experiment [66], in the USA, plan to check the LSND measurement using the same neutrino production method.

MicroBooNE [67] was designed to investigate MiniBooNE results at low neutrino energy (< 600 MeV). Its commissionning started mid-August 2015, neutrino beam at Fermilab is expected to be ready early November 2015 [68].

Source anomalies

GALLEX and SAGE experiments in 90s were designed to measure the solar neutrino flux with gallium detectors. Both experiments were calibrated using high activity ν_e sources.

However, in 2010, a re-analysis of these calibration data [69] showed that a deficit was observed between the number of detected events and the number of expected events [62]:

$$R = \frac{N_{obs}}{N_{exp}} = 0.86 \pm 0.05 \tag{1.62}$$

This deficit can also be explained by oscillations with a high Δm^2 sterile flavor. The source anomaly can be investigated using high activity source. The SOX experiment [70] plans to deploy a PBq $\overline{\nu}_e$ source close to the BOREXINO detector [71].

Reactor anomalies

A recent re-analysis of the $\overline{\nu}_e$ spectra emitted by nuclear reactor core has been performed by Mueller et al [72]. It re-analyses the reference electron spectra measured at the high flux ILL reactor in Grenoble (France) by Schreckenbach et al. [73, 74, 75, 76]. The conversion of these electra spectra in into $\overline{\nu}_e$ spectra was shown to be incomplete in this re-analysis. A new method has been proposed to perform this conversion and led to a shift of the spectra normalization by 3% on average [72]. This results was confirmed by a re-evaluation of the antineutrino spectra in [77].

This analysis was applied on older experiment data, and mostly short baseline reactor experiment (SBL) with a < 100 m distance from the reactor core. These experiments were performed in the 80s to probe the neutrino oscillations. However, as we now know, the current 3ν oscillations framework does not allow $\overline{\nu}_e$ oscillations at such low E/L. SBL experiments did not observe any oscillation in their time.

With the re-analysis of the $\overline{\nu}_e$ reactor spectra, Mention et al. [78] found that the ratio between observed and expected number of $\overline{\nu}_e$ falls to $R = 0.9430.023$. FIG. 1.11 show the new estimation of $R = \frac{N_{obs}}{N_{exp}}$ for different experiments.

This deficit can also be explained by the introduction of a new sterile flavor with $\Delta m^2 > 1.5$ eV^2. Authors find an agreement between their claim and MiniBooNE observations. However, in the highlight of the $[4, 6]$ MeV distortion observed by Double Chooz (see section 5.5.1 and [79]), RENO and Daya Bay, and its possible explanations [80, 81], one could consider that this deficit may come from a lack of knowledge on the reactor $\overline{\nu}_e$ spectra.

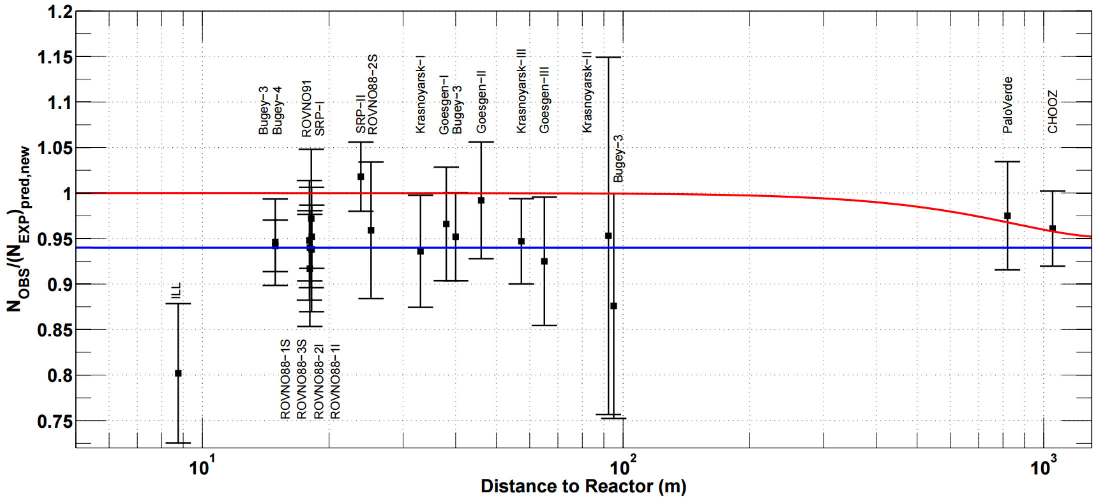


FIGURE 1.11: Illustration of the short baseline reactor antineutrino anomaly from [78]. The experimental results are compared to the prediction without oscillation, taking into account the new antineutrino spectra, the corrections of the neutron mean lifetime, and the off-equilibrium effects.

Indeed, the situation is still unclear and needs to be investigated. Many experiments plan to study sterile oscillations at very short distance from reactor cores. These experiments involve small detectors using well-known technologies like Gd-doped liquid scintillator in the case of experiments like STEREO [82], based near ILL research reactor, or innovative technologies like plastic scintillator with ^{6}Li nuclei in the case of experiments like SOLi∂ [83], based near BR2 research reactor at Mol[3].

---

[3] Many other experiments are developed, here only the experiments with EU laboratories were mentioned.

Cosmological constraints

Relic neutrinos from the earliest stage of the Universe can be indirectly used to provide information on the structure and on the composition of the Universe. Relic neutrinos affected the cosmic microwave background (CMB) radiation. Therefore, studies on the CMB anisotropies can allow to extract informations on neutrinos. Both effective number of neutrino species, N_{eff}, and the sum of the neutrino masses, $\sum N_\nu$, can be extracted from the studies on the CMB radition.

The Planck collaboration, using the Planck satellite to observe the CMB radiation anisotropies, provides recent measurements of both parameters, using also the baryon acoustic oscillation measurements [84]:

$$N_{eff} = 3.04 \pm 0.18 \tag{1.63}$$

$$\sum N_\nu < 0.17 \text{ eV} \tag{1.64}$$

Summary

Three different detection methods led to the measurement of anomalies with respect to the current 3ν oscillation model. However, the existence of a fourth flavor of neutrino is disfavored by the Planck results. In addition, the discovery of the $4-6$ MeV distortion by the reactor experiments [175, 176, 177] could indicate that a lack of knowledge on the reactor $\overline{\nu}_e$ spectra is an explanation of the so-called reactor anomaly.

Ongoing experiments should allow to clarify the situation on the existence of more than 3 neutrinos in the next decade.

1.4 Neutrino mass

Masses of the SM particles are generated by the Brout-Englert-Higgs mechanism. It consists in the introduction of a Higgs doublet $\Phi(x)$, which allows to generate masses through spontaneous symmetry breaking of $SU(2)_L \otimes U(1)_Y$ symmetry. $\Phi(x)$ is a complex doublet with a spin 0:

$$\Phi(x) = \begin{pmatrix} \Phi^+(x) \\ \Phi^0(x) \end{pmatrix} \tag{1.65}$$

where Φ^+ is the charged complex scalar field and Φ^0 the neutral one. With the addition of the Higgs field, the SM electroweak Lagrangian is written as:

$$\mathcal{L}_{EW} = \mathcal{L}_{gauge} + \mathcal{L}_{matter} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{1.66}$$

where $\mathcal{L}_{gauge}$ describes the interactions between the $SU(2)_L$ and $U(1)_Y$ gauge bosons. $\mathcal{L}_{matter}$ is the kinetic term for the Standard Model fermions. $\mathcal{L}_{Higgs}$ describes the Higgs field and $\mathcal{L}_{Yukawa}$ describes the Yukawa interaction generating the fermions masses.

The Higgs field Lagrangian is written as:

$$\mathcal{L}_{Higgs} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) \tag{1.67}$$

with $V(\Phi)$ the Higgs field potential:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 = \lambda \left(\frac{-\mu^2}{2\lambda} + \Phi^\dagger \Phi \right)^2 - \frac{\mu^4}{4\lambda} \tag{1.68}$$

where λ is the Higgs self-coupling strength, which has to be positive in order to have a potential bounded from below [85], and μ is the Higgs quadratic coupling parameter. μ is related to the Higgs boson mass $m_{Higgs} = \sqrt{-2\mu^2}$. This implies that $\mu^2 < 0$ in order to have spontaneous symmetry breaking. Therefore, if we neglect the $\frac{\mu^4}{4\lambda}$ term, there are two minima at $\Phi = v^\pm$, with $v^\pm$ defined in Eq. 1.69. These two minima correspond to two vacua, degenerated in energy. The Lagrangian is no longer invariant under $SU(2)_L \otimes U(1)_Y$ symmetries.

$$v^\pm = \pm \sqrt{\frac{-\mu^2}{\lambda}} \tag{1.69}$$

After this symmetry breaking we have:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.70}$$

where $H(x)$ is the physical Higgs boson. Then fermion masses come from a Yukawa coupling between the fermions field and the Higgs doublet:

$$\mathcal{L}_{Yukawa} = -\sum_{\alpha^-} \frac{G^\alpha_- v}{\sqrt{2}} (\overline{f}^\alpha_L \Phi f^\alpha_R + \overline{f}^\alpha_R \overline{\Phi} f^\alpha_L) - \sum_{\alpha^+} \frac{G^\alpha_+ H}{\sqrt{2}} (\overline{f}^\alpha_L \Phi^c f^\alpha_R + \overline{f}^\alpha_R \overline{\Phi}^c f^\alpha_L) \tag{1.71}$$

where f^α_L and f^α_R stand respectively for the left-handed doublet and the right-handed singlet of a fermion α, $G^\alpha_\pm$ is the α fermion Yukawa coupling for $T_3 = \pm\frac{1}{2}$, with T_3 the weak isospin. The first term is a mass term, with $m_\alpha = \frac{G^\alpha v}{\sqrt{2}}$. While the second term corresponds to the fermion coupling to the Higgs field $-i\frac{m_\alpha}{v}$.

This mechanism was postulated independently by P.W. Higgs, and F. Englert and R. Brout, in 1964 [86, 87]. ATLAS and CMS experiments announced, on July 4th, 2012, the discovery of the Higgs boson, confirming the Brout-Englert-Higgs mechanism [88, 89].

However, due to the current absence of right-handed neutrinos in the SM, this mechanism cannot explain the mass of the neutrino induced by neutrino oscillation (cf. section 1.3). Therefore, the current SM needs to be modified in order to find a way to generate the neutrino masses. One possibility is to add right-handed neutrino singlet in the theory, another is the use a Higgs triplet. In that sense, one can consider the evidence of neutrino flavor oscillations as a hint for physics beyond the SM.

1.4.1 Majorana or Dirac particles?

1.4.1.1 Dirac mass

By adding the right-handed neutrino singlets in the SM, we can generate the neutrino masses with the same Higgs mechanism as the one used for other fermions. The Lagrangian of mass is then extracted from Eq. 1.71, it can be written as:

$$\mathcal{L}^D_{mass} = -m^D_\nu(\overline{\nu}_R\nu_L + \overline{\nu}_L\nu_R) \tag{1.72}$$

In this case, neutrinos are considered as Dirac particles, as are the other elementary fermions. The right-handed neutrinos, as singlets under all SM interactions, are considered to be *sterile* particles.

However, if we consider right-handed neutrino singlets to explain the neutrino masses, the new SM free parameters, the neutrinos Yukama couplings, have to be in accordance with the observations: $G^\nu \simeq 10^{-12}$. This value is several orders of magnitude smaller than the other fermion Yukawa couplings ($G^e \sim 0.3^{-5}$). Such value is commonly considered to be *unnatural.*

1.4.1.2 Majorana mass

If neutrinos are Dirac particles, they are like any other particles of the SM, with their left-handed and right-handed chiral projections considered to be independent. However, neutrinos are the only neutral elementary fermions. Therefore, they are quite different from the other SM particles. It is then theoriticaly possible to investigate the possibility that they are Majorana particles.

E. Majorana proposed in 1937 the existence of neutral particles which are their own antiparticles under the CPT transformation [90]. In other terms, Majorana particles and antiparticles are the opposite chiral states of the same particle.

The Majorana mass cannot be explained via the Higgs coupling, as it mixes particles and antiparticles. In the case of neutrinos as Majorana particles, the origin of their masses is different to the one of quarks and charged leptons masses. The Majorana neutrino Lagrangian of masses can be expressed as following:

$$\mathcal{L}^M_{mass} = -\frac{1}{2} m^M_\nu (\overline{\nu}^c_L \nu_L + \overline{\nu}_L \nu^c_L) \tag{1.73}$$

where $\overline{\nu}^c_L$ and ν^c_L stand for the charge-conjugate of $\overline{\nu}_L$ and ν_L, respectively.

1.4.1.3 See-saw mechanism

Dirac and Majorana masses are two ways to explain neutrinos masses. However, they do not explain the smallness of these ones. The *see-saw* mechanism allows to explain it by the presence of high masses sterile neutrinos. Therefore, it is possible to write a general mass Lagrangian, using the two descriptions:

$$\mathcal{L}^{D+M}_{mass} = -\frac{1}{2} (\overline{\nu}^c_L, \overline{\nu}_R) \begin{pmatrix} m^M_L & m^D \\ m^D & m^M_R \end{pmatrix} \begin{pmatrix} \nu_L \\ \nu^c_R \end{pmatrix} \tag{1.74}$$

where m^M_L and m^M_R are the left and right-handed Majorana neutrino masses, respectively. This matrix can be diagonalized using an orthogonal matrix, which allows to have two mass eigenvalues, m_1 and m_2:

$$m_{2,1} = \frac{m^M_L + m^M_R}{2} \pm \sqrt{\left(\frac{m^M_L - m^M_R}{2}\right)^2 + (m^D)^2} \tag{1.75}$$

Due to the current absence of mass hierarchy or mass absolute values observations, different cases can be considered. In the *see-saw* mechanism, $m^L_L = 0$ and $m^M_R \gg m^D$, leading to:

$$m_1 \simeq m^M_R \tag{1.76}$$

$$m_2 \simeq \frac{(m^D)^2}{m^M_R} \tag{1.77}$$

The right-handed sterile neutrino ν_R of mass m_1 is therefore very heavy compared to the SM left-handed active neutrino of mass m_2. This allows to explain the smallness of the neutrino masses, while generating the Dirac mass m^D as for all the other SM fermions.

1.4.2 Neutrino mass experimental measurements

1.4.2.1 Absolute neutrino mass measurements

The absolute mass of the neutrino α can be express as:

$$m^2(\nu_\alpha) = \sum_j |U_{\alpha j}|^2 m^2(\nu_j) \text{ where } \alpha = e,\ \mu,\ \tau \tag{1.78}$$

Using the β decay, it is possible to measure the absolute mass of the neutrinos. In case of a β^- decay, we have:

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z+1}X + e^- + \overline{\nu}_e \tag{1.79}$$

If we neglect the ${}^{A}_{Z+1}X$ nucleus recoil, the energy emitted by the reaction is shared between e^- and $\overline{\nu}_e$. When measuring the e^- spectrum, which is simpler to measure than the $\overline{\nu}_e$ spectrum, it would reach a value $E_e^{max} = E_0 - m_{\nu_e}$ where E_0 is the reaction available energy and m_{ν_e} the absolute electron neutrino mass at rest. We can compute the E_0 value with simple computations. It should be then possible to deduce the $\overline{\nu}_e$ mass. FIG. 1.12 shows an example of an e^- spectrum after a β decay.

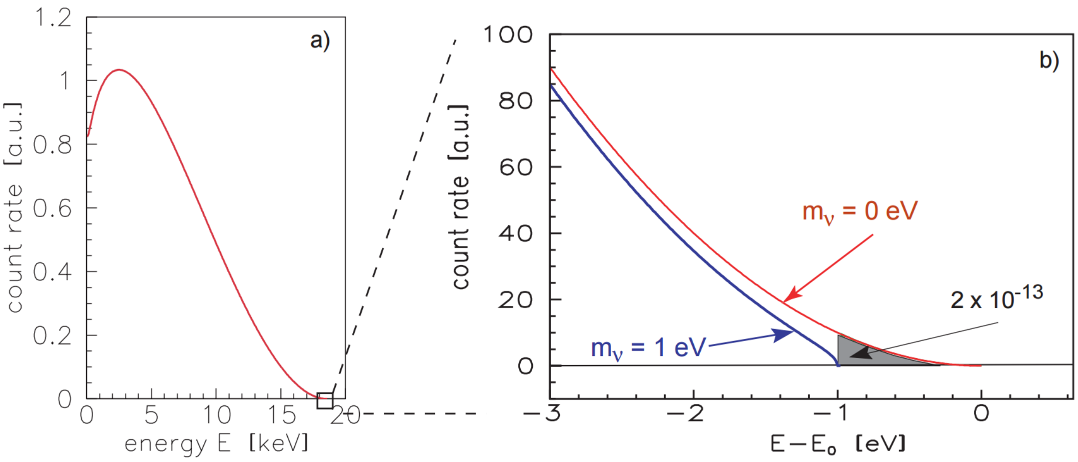


FIGURE 1.12: The electron energy spectrum of tritium β decay (a), with the narrow region around endpoint E_0 (b), from [93].

Several experiments use the tritium β decay to measure the absolute mass. Tritium, with a $E_0 = 18.6$ keV, a simple electronic shell and a half life of 12.3 years, is a good candidate for such experiments, as a low E_0 is mandatary to have the needed energy resolution.

Mainz [91], Troitsk [92] and the upcoming KATRIN [93] experiments are based on its spectroscopy. The current best precision measurement is the upper limit fixed by Troitsk experiment at $m_{\nu_e} \leq 2.12$ eV [92]. The KATRIN experiment is expected to have a sensitivity of $m_{\nu_e} \leq 0.35$ eV [93].

Less precise limits were also measured for the absolute mass of ν_μ and ν_τ. Using $\pi^+ \rightarrow \mu^+ + \nu_\mu$ decay, accelerator experiments found $m_{\nu_\mu} \leq 170$ keV [94]. And, using $\tau^- \rightarrow 2\pi^- + \pi^+ + \nu_\tau$ and $\tau^- \rightarrow 3\pi^- + 2\pi^+ + \nu_\tau$ decays, accelerator experiments found $m_{\nu_\tau} \leq 18.2$ MeV [95].

1.4.2.2 $\beta\beta \to 0\nu$ experiments

In order to probe the Dirac or the Majorana nature of the neutrino, experiments use the double β decays reaction. It consists in the simultaneous transformation of two neutrons into two protons in the same even-even nucleus. The simple β decay is not allowed for these nuclei.

On neutrino point of view, the common result of double β decay is the emission of two neutrinos ($\beta\beta \to 2\nu$), in addition to the two electrons:

$$^{A}_{Z}X \to ^{A}_{Z+2} X + 2e^{-} + 2\overline{\nu}_{e} \tag{1.80}$$

This reaction does not violate the lepton number conservation and can be observed with typical half-life of $10^{18}/10^{20}$ years.

If neutrinos are Majorana particles, double β decays could result in an emission of only two electrons and no neutrino, with both neutrinos being annihilated by each other. This is only possible if neutrinos are their own antiparticle.

$$^{A}_{Z}X \to ^{A}_{Z+2} X + 2e^{-} \tag{1.81}$$

This reaction violates the lepton number by two units, and is therefore not allowed in the SM. If this reaction is observed it would demonstrate the Majorana nature of the neutrino.

Several experiments look at $\beta\beta \to 0\nu$ signal with different $\beta\beta$ isotopes. Presently, no evidence of $\beta\beta \to 0\nu$ signal was found, experiments only put limits on the $\beta\beta \to 0\nu$ half-life for different isotopes.

FIG. 1.13 shows the possible correlation between the $m_{\beta\beta} = |\sum_i m_i \cdot U_{ie}^2|$ mass and the neutrino lightest mass. Present and near-future experiments with a $\sim$ 100 kg detector, like SuperNEMO [98], CUORE [99], KamLAND-Zen [100], etc., have a sensitivity until $m_{\beta\beta} \gtrsim 0.35$ eV. The next generation of detectors ($\sim$ 1 ton), like nEXO [101], THEIA [103] or SNO+ [102], will have a sensitivity until $m_{\beta\beta} \gtrsim 0.01$ eV. In case of I.H. hierarchy, since $m_{\beta\beta}$ should be > 0.01 eV, according to FIG. 1.13, the nature of the neutrino should be determined by the next generation of detectors.

1.4.2.3 Mass hierarchy measurements

Future long baseline experiments looking for the δ_{CP} measurement can be able to determine the mass hierarchy at the same time, as δ_{CP} is strongly correlated to the mass hierarchy [61]. However, it is also possible to take advantage of the matter effect on neutrino oscillations to determine the mass hierarchy. Matter effect alters the oscillation of ν and of $\overline{\nu}$ differently depending of the mass-hierarchy. From EQ. 1.40, it is possible to demonstrate that the matter effect depends of the mass hierarchy. Taking $V = \sqrt{2}G_F N_e$, EQ. 1.40 can be rewritten as:

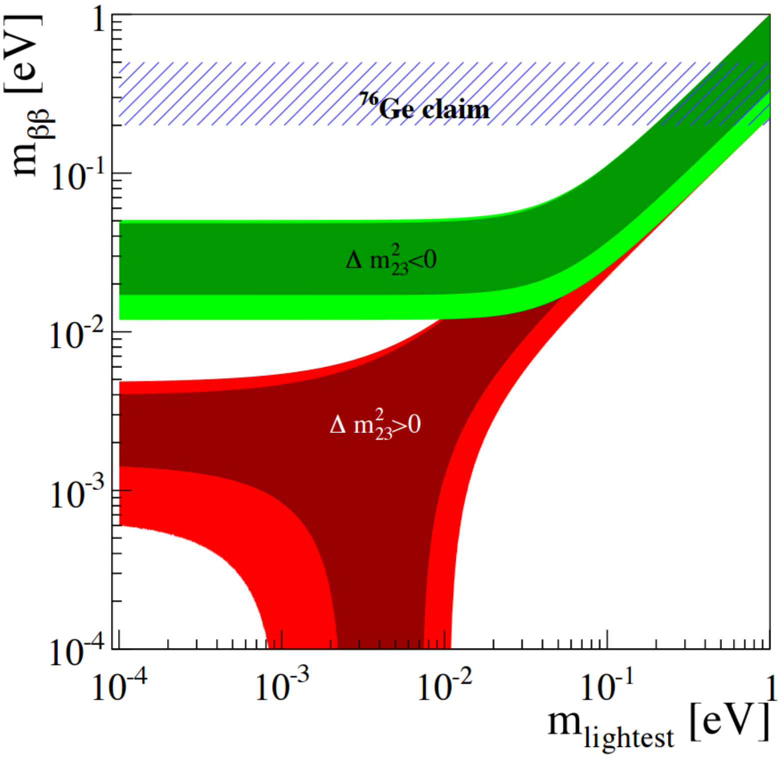


FIGURE 1.13: ($m_{\beta\beta}$, $m_{lightest}$) plane from [96], showing the possible contours of the N.H. (red) and of the I.H. (green), and the, now-excluded, ^{76}Ge claim [97].

$$\tan(2\theta_m) = \frac{\frac{\Delta m^2}{2E}\sin(2\theta)}{\frac{\Delta m^2}{2E}\cos(2\theta) - V_e} \tag{1.82}$$

in case of ν oscillations, and as:

$$\tan(2\theta_m) = \frac{\frac{\Delta m^2}{2E}\sin(2\theta)}{\frac{\Delta m^2}{2E}\cos(2\theta) + V_e} \tag{1.83}$$

in case of $\overline{\nu}$ oscillations. Therefore, the resonance effect, i.e. $\frac{\Delta m^2}{2E}\cos(2\theta) \mp V_e = 0$, occurs for ν in case of N.H., and for $\overline{\nu}$ in case of inverted hierarchy.

Another possibility to determine the mass hierarchy is to use $\overline{\nu}_e$ and take advantage of the small phase shift in the oscillation terms depending on Δm^2_{31} and on Δm^2_{32}. This leads to an interference in the $\overline{\nu}_e$ survival probability. This interference become largest at a baseline of $\sim$ 55 km, where the solar oscillations become maximums. The position of the maxima ad minima of the subdominant oscillations is shifted by 180 depending on the hierarchy [108]. This is illustrated in FIG. 1.14.

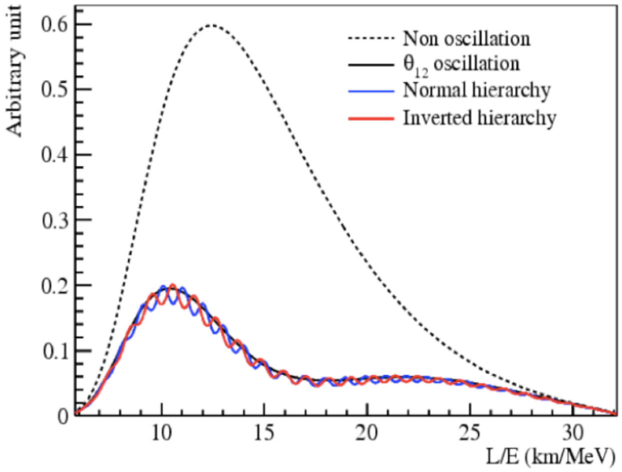


FIGURE 1.14: $\overline{\nu}_e$ survival probability as a function of L/E, from [108].

Several upcoming experiments plan to measure the mass-hierarchy in a few years: JUNO [108] in China, RENO-50 [109] in Korea, KM3Net-ORCA [104] in the Mediterranean Sea, PINGU [105] in Antartica and INO/ICAL [106] in India. These experiments are going to look at atmospheric neutrinos which travel across the Earth. Fig. 1.15 shows the sensitivity of PINGU and KM3Net-ORCA to mass hierarchy. JUNO and RENO-50 experiment are going to look at reactor electron antineutrino from a very long baseline (~ 55 km).

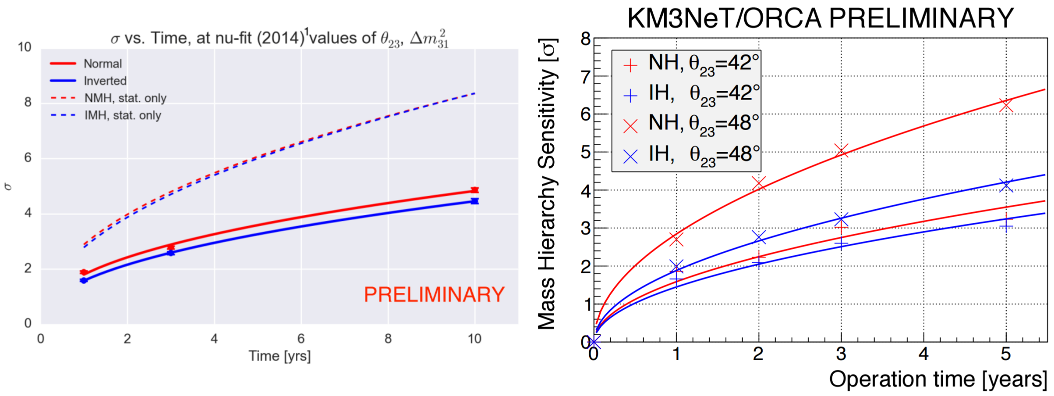


FIGURE 1.15: Sensitivity on the Mass Hierarchy determination of PINGU (*left*, from [42]) and of KM3Net-ORCA (*right*, from [107])

1.5 Summary

The status of the neutrino research has been presented in this chapter. Over the years the different experiments allowed to confirm the status of the neutrino as a massive particle with a mixing between its mass and flavor eigenstates. In the last years the value of the last not-known mixing angle, θ_{13}, has been measured by the reactor experiments (Daya Bay, RENO and Double Chooz) and the accelerator experiments (T2K). The analyses leading to the θ_{13} measurement of Double Chooz is presented in this thesis.

There are still some open questions on the value of δ_{CP} phase, on the mass hierarchy, on the absolute mass of the neutrino and on the existence of one or more sterile flavors. Concerning the δ_{CP}, the large value of θ_{13} will allow future experiment like DUNE to measure it. In the next decade, several experiments (ORCA, PINGU, JUNO, RENO-50...) plan to give a measurement of the mass hierarchy. Many experiments are also running, or are close to, in order to determine the existence of a sterile flavor with reactor short baseline, accelerator short baseline and source experiments.

Chapter 2

The Double Chooz experiment

Double Chooz is a reactor experiment, located in Chooz, in the French Ardennes. Its detectors are positioned near Chooz-B nuclear power plant. Double Chooz aims at the θ_{13} angle measurement by the detection of an electron anti-neutrino disappearance.

Nuclear reactor cores are the most abundant source of $\overline{\nu}_e$ on Earth, with a production of about 10^{21} $\overline{\nu}_e$/s. These $\overline{\nu}_e$ are emitted by the $\beta - decay$ of radioactive nuclei produced by the nuclear fissions. The anti-neutrino flux emitted is a pure $\overline{\nu}_e$ flux, as the maximal energy released by $\beta - decays$ is far smaller than the rest mass of μ or τ, which allows to study the $\overline{\nu}_e$ disappearance in order to measure θ_{13}.

Since the E/L ratio of the reactor neutrino in the configuration of Double Chooz, it is possible to neglect the oscillations driven by the smaller Δm^2. Therefore, the $\overline{\nu}_e$ survival probability, i.e. the probability for one $\overline{\nu}_e$ to remain in $\overline{\nu}_e$ state, with an energy E and at a distance L from its source, can be written from Eq. 1.29 as:

$$P\left(\overline{\nu}_e \longrightarrow \overline{\nu}_e\right) = 1 - \sin^2(2\theta_{13}) \sin^2\left[1.27\left(\frac{\Delta m^2_{13}\,(eV^2)\cdot L\,(km)}{E\,(GeV)}\right)\right] \tag{2.1}$$

Since Double Chooz is a disappearance experiment, its results do not depend on CP violation phase, δ_{CP} parameter vanished with the PMNS development to obtain the $\overline{\nu}_e$ survival probability.

In this chapter, the design of the Double Chooz experiment will be reviewed. The online system of the experiment will be presented, with a particular focus on the data handle system, whose development and maintenance were a part of my work during this thesis. The production of $\overline{\nu}_e$ by the nuclear cores, as well as the simulation of this production and of the $\overline{\nu}_e$ interactions in the detectors, are presented. Finally, we detail the DC offline reconstruction, particulary the Inner

Veto vertex reconstruction algorithm, which was one of the contribution of my thesis to the DC experiment. Despite the complex geometry of the Inner Veto, which was not designed to perform a vertex reconstruction, a resolution of 36.31 cm was reached with this algorithm.

2.1 $\overline{\nu}_e$ detection principle at Double Chooz

In the Double Chooz experimental setup, $\overline{\nu}_e$ are detected through **Charged current** (cf. section 1.2.1) in liquid scintillator:

$$\overline{\nu}_e + p \rightarrow e^+ + n \tag{2.2}$$

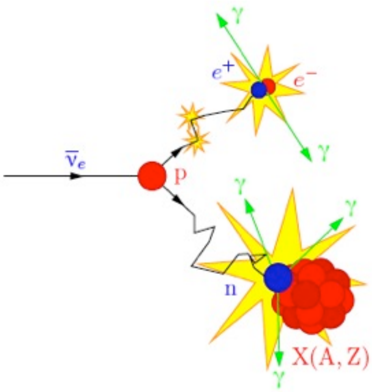


Figure 2.1: Inverse β-decay reaction with neutron capture on Gadolinium.

This charged current is called inverse β-decay reaction (IBD). The IBD reaction has a threshold of 1.8 MeV and provides a clear signature, shown in Fig. 2.1. The e^+ quickly loses its energy and annihilates itself with an electron from the liquid scintillator, producing a total energy ranging from 1 to 12 MeV (two γ of 511 keV each from positron annihilation and the positron energy deposited in the scintillator). The neutron can be captured on nuclei like Hydrogen from liquid scintillator, producing a 2.223 MeV γ, or the neutron capture signal can be improved by doping the scintillator with Gadolinium nuclei, since its thermal neutron cross-section is the highest among any stable nuclei (259 000 barns) and its capture signature is easier to discriminate from natural radioactivity (with the emission of several γ rays for a total energy $\sim$ 8 MeV). The neutrino signal consists in a coincidence between a prompt event, the positron annihilation and its loss of energy, and a delayed event, the neutron capture. The $\overline{\nu}_e$ energy can be computed easily using the kinematic relations. Eq. 2.3 shows the $\overline{\nu}_e$ energy, if the neutron recoil is neglected, as its mass is far greater than the positron mass:

$$E_{\overline{\nu}_e} = (m_n - m_p) + E_{e^+} \tag{2.3}$$

$$E_{\overline{\nu}_e} = \Delta + E_{e^+} \tag{2.4}$$

Where m_n is the neutron rest mass, m_p the proton rest mass and E_{e^+} the positron energy. Δ is the mass difference between the neutron and the proton, its value is $\Delta \simeq 1.293$ MeV. In a liquid scintillator, as in Double Chooz detectors, the visible positron energy is the kinetic energy deposited by the positron, T_{e^+}, and the energy released by its annihilation with an electron:

$$E_{vis} = T_{e^+} + 2 \times m_e = (T_{e^+} + m_e) + m_e \simeq E_{e^+} + m_e \simeq E_{\overline{\nu}_e} - \Delta + E_{0_e} \tag{2.5}$$

Where m_e is the electron rest mass. Given this, the relation between the $\overline{\nu}_e$ and the visible energy is described by the Eq. 2.6:

$$E_{\overline{\nu}_e} \simeq E_{vis} + (\Delta - m_e) \simeq E_{vis} + 0.782 MeV \tag{2.6}$$

Eq. 2.1 shows that the oscillation probability, and so the value of θ_{13}, depends on the ratio E/L. As shown in Eq. 2.6, $E_{\overline{\nu}_e}$ can be measured with the positron E_{vis}. Then, it is possible to extract θ_{13} from an energy spectrum analysis, using a comparison between the E_{vis} spectrum after oscillation, and the one from prediction. Such analysis is the main θ_{13} measurement done by Double Chooz. In order to provide it, a precise energy reconstruction and calibration is mandatory, this is detailed in chapter 3. In addition, in the first phase of Double Chooz, the prediction is given by a Monte-Carlo simulation, then a precise Monte-Carlo (MC) simulation is essential, the Double Chooz MC simulation is described in section 2.6. The θ_{13} energy spectrum analysis is detailed in chapter 5.

2.2 The Double Chooz detectors

Double Chooz uses two identical detectors located at different baselines, for $\overline{\nu}_e$ detection.

The Far Detector is settled in the lab hall used by the previous CHOOZ experiment, located 1050 m away from the reactor cores mean position. Far Detector is located near the maximum amplitude of $\overline{\nu}_e$ oscillation due to θ_{13}. The far lab hall is located 150 m under a hill, which leads to an overburden of 300 meter water equivalent (m.w.e.) for the detector. This overburden reduces the cosmic muon rate to ~ 30 Hz in the inner detector volume.

The Near Detector is settled in a new lab hall, located 400 m away from the reactor cores mean position. $\overline{\nu}_e$ oscillations due to θ_{13} are expected to be

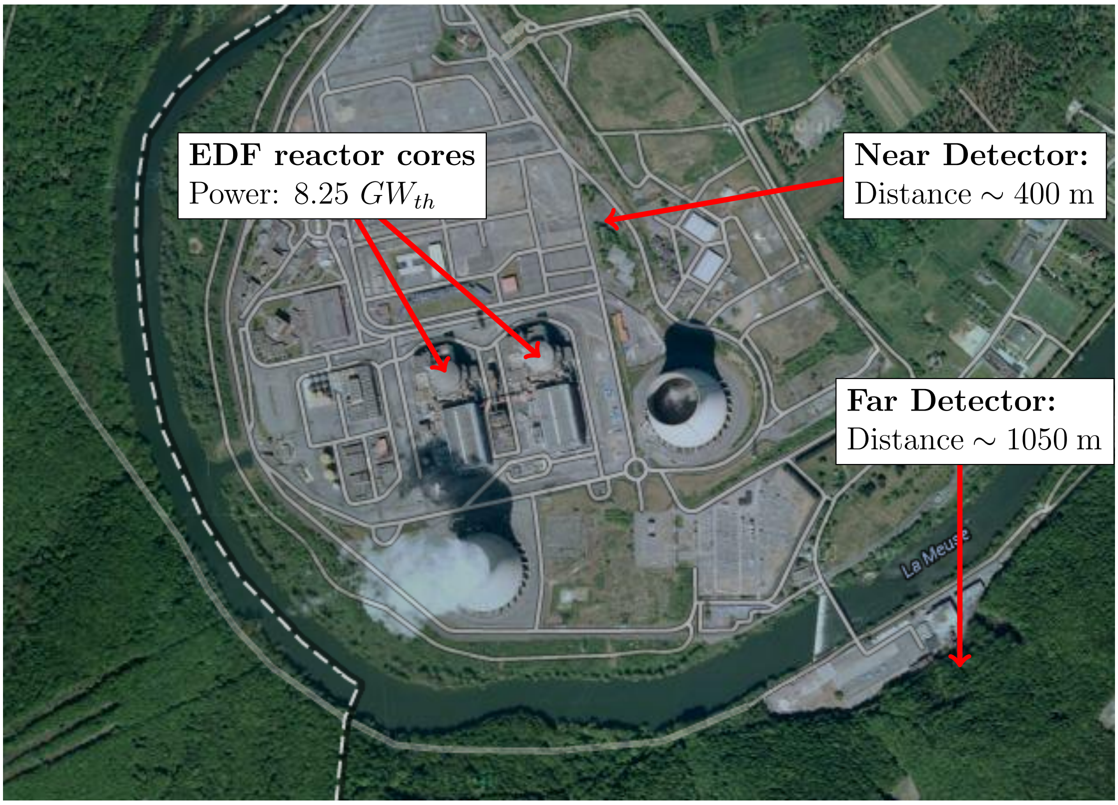


FIGURE 2.2: Aerial view of Double Chooz site (picture from Google maps)

negligible at this distance. The near lab hall is not built under a hill and has only an overburden of 120 m.w.e., which leads to a cosmic muon rate lowering to ~ 230 Hz.

The Far Detector was built between 2009 and 2010 and is taking data since April 2011. The Near Detector was build between 2011 and 2014 and is taking data since January 2015. An aerial view of the Double Chooz site is shown in FIG. 2.2.

2.2.1 Detector design

The Double Chooz detector design was inspired by the CHOOZ detector design. This design is shown in FIG. 2.3. It consists in three sub-detectors, each one being optically isolated from the others: the inner detector (ID), the inner veto (IV) and the outer veto (OV). Each one of these detectors is detailed in the following subsections. Despite their names, events with signal in IV or in OV are not vetoed

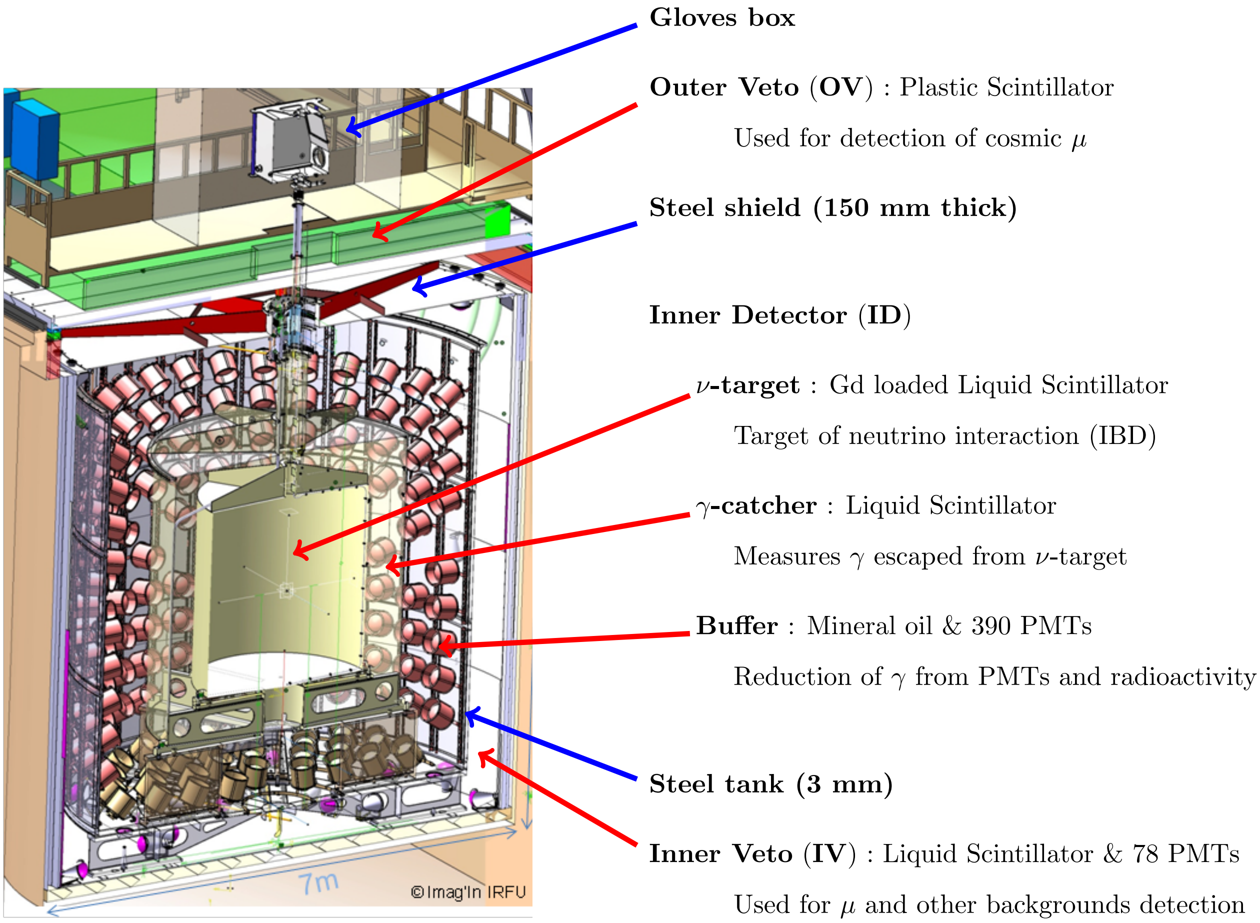


FIGURE 2.3: Double Chooz detector design

online, but the information is used in the off-line analysis.

The inner detector and inner veto volume is surrounded by a 150 mm thick low activity and demagnetized stainless steel shield. Its aim is to protect the detector from low energy gammas coming from the surrounding rock. This is one of the improvements with respect to the CHOOZ experiment, which used low activity sand to shield its detector.

This design has been optimized for a high $\overline{\nu}_e$ detection and a high signal-over-background ratio. The signal-over-background is > 20 in DC third publication for the Gadolinium analysis [79] and about ~ 11 for the Hydrogen analysis (more details are given in the chapter 5 about Double Chooz offline analysis). The detector materials have been selected in order to reduce the radioactivity level below 10 Bq in the Inner Detector.

2.2.1.1 The Inner Detector (ID)

The Inner Detector consists in three nested cylindrical volumes. From the inside to the outside: the ν-target, the γ-catcher and the buffer.

The **ν-target** volume is the primary fiducial volume for $\overline{\nu}_e$ detection. Its vessel is an 8 mm thick acrylic vessel, filled with 10.3 m^3 Gadolinium-loaded liquid scintillator. This liquid scintillator is composed of 80% dodecane ($C_{12}H_{26}$), 20% PXE (C_6H_{18}), $7g/L$ of PPO ($C_{15}H_{11}NO$), $20g/L$ of bis-MSB ($C_{24}H_{22}$), and of $1g/L$ of Gadolinium. The Gadolinium concentration in the Double Chooz experiment allows $\sim$ 85% neutron capture probability on Gadolinium. Time stability for the Gd-doped liquid is ensured by Gadolinium atoms encapsulation in a $Gd(dpm)_3$ molecule[1]. It ensures an efficient dissolving in the scintillator and a durable transparency.

Dodecane and PXE are aromatic components. Aromatic molecules are excited by charged particles. Following this excitation, its possible vibrational and rotational parts are transformed into heat by collision with other molecules. Their electrons deexcite by an internal radiationless conversion process into their first excited state. This first excited state then deexcites under photon emission either by fluorescence (fast photon emission) or less often by phosphorescence (slower photon emission).

As the absorption spectra of aromatic molecules overlap their own emission spectra, aromatic solutes are added in the liquid scintillator in order to shift the wavelength of the photon to longer values, for which the mixture is more transparent. Double Chooz wavelength-shifting aromatic solutes, also called “fluors”, are PPO and bis-MSB. These fluors are chosen to ensure a high overlap between the emission spectrum of PXE and the absorption spectrum of PPO, as well as between the emission spectrum of PPO and the absorption spectrum of bis-MSB. Given the scintillator high transparency at bis-MSB emission spectrum, most of the scintillation light is emitted by bis-MSB. This is illustrated in FIG. 2.4.

PPO and bis-MSB were also chosen because their emission spectra match the region of the highest quantum efficiency in the photomultipliers. Double Chooz target liquid scintillator was produced for Near and Far Detector in 2010 at MPIK. Both liquids are then the same, and with the same aging. This would leads to a complete cancelation of the uncertainties on the proton number when comparing Near and Far Detector. Since protons are the Inverse Beta Decay targets and lead to the thermalization time of the neutron, a good knowledge of the proton

---

[1] $(dpm)_3$ stands for “tris(dipivaloylmethanato)” which reflects the structure of the molecule.

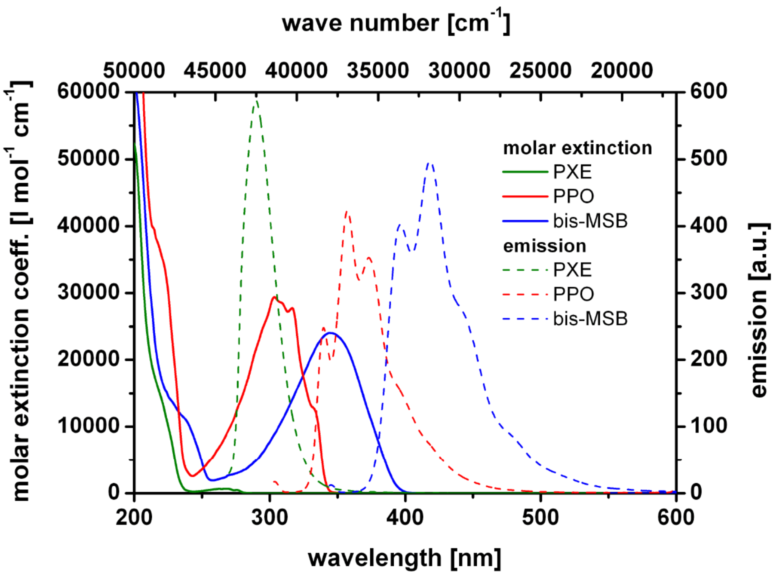


FIGURE 2.4: Molar extinction coefficients (i.e. absorption spectra) (solid) and emission (dashed) spectra of the aromatic components in the ν-target and in the γ-catcher scintillators, from [111].

number is mandatory to reduce the systematics on θ_{13} measurement. Then, a special care was taken to the knowledge of the proton number in the ν-target since the production of the liquid, leading to a good knowledge of this proton number. This proton number value is shown in TAB. 2.1.

Proton number in:	
Target	$13.60^{+0.02}_{-0.04}\%$ (computed)
γ-Catcher (Far Detector)	$14.6 \pm 0.2\%$ (measured)
γ-Catcher (Near Detector)	to be measured

TABLE 2.1: Proton number (or Hydrogen fraction) in the Double Chooz liquids [136]. Proton numbers were not computed for γ-catcher liquids since the mineral oil proton number was not known to a good enough precision.

The γ-**catcher** volume is the secondary fiducial volume for $\overline{\nu}_e$ detection. Its vessel is a 12 mm thick acrylic vessel, filled with 22.3 m^3 non-Gadolinium-loaded liquid scintillator. γ-catcher primary aim is to collect escaped gammas from the

IBD and neutron capture inside ν-target. This improves the energy containment efficiency of the detector. γ-catcher secondary aim is to be used in a second IBD selection, using neutron capture on Hydrogen, as a fiducial volume.

The γ-catcher liquid scintillator is composed of 30% dodecane, 66% mineral oil (Shell Ondina909), 4% PXE, $2g/L$ of PPO and 20 mg/L of bis-MSB. This liquid 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. Double Chooz γ-catcher liquid scintillator was produced at MPIK in 2010 for Far Detector and in 2013 for Near Detector. As Hydrogen IBD analysis was not initially planned for Double Chooz, γ-catcher liquid scintillators for Near and Far Detector were not produced at the same time. This should lead to a discrepancy between the γ-catcher proton number in Far Detector and in Near Detector. The proton number in the Near Detector was not yet measured.

The **buffer** volume consists in a 1.05 m thick volume filled with 114 m^3 non-scintillating oil. The buffer is one of the main improvements with respect to the original CHOOZ design. Its structure serves as support for the 390 Inner Detector photomultiplier tubes (PMT). These PMTs are based on the Hamamatsu R7081 model and have a 10 inches photocathode with an average Quantum Efficiency of 23%. The fraction of surface covered by these PMTs is about 13%, which leads in ~3% photon detection efficiency. The main purpose of the buffer volume is to shield the ν-target and the γ-catcher volumes from gammas radioactivity coming mainly from PMTs photocathode.

The buffer non-scintillating oil is composed of 47.2% n-alkanes and 52.8% Ondina917 oil. This composition has been chosen to ensure material compatibility with γ-catcher acrylic vessel and with PMTs, and an high transparency in the scintillator emission wavelengths.

2.2.1.2 The Inner Veto (IV)

The Inner Veto consists in a cylindrical volume surrounding the Inner Detector and optically isolated from it. It is filled with 90 m^3 liquid scintillator, composed of 50% n-alkanes and 50% linear alkyl benzene (LAB) with $2g/L$ of PPO and $20mg/L$ of bis-MSB. LAB is the aromatic compound of the liquid scintillator, with PPO and bis-MSB as fluors.

The Inner Veto main aim is to shield the ID from low energy radioactive background and to detect, with high efficiency, cosmic muons and spallation neutrons crossing its volume. The studies performed during my Ph.D. shown that it can be used also to detect radioactive gamma from Thallium and Potassium (see section 5.1.4).

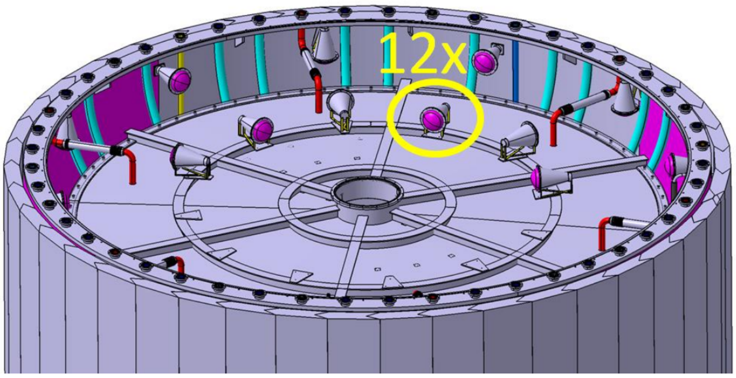


FIGURE 2.5: View of Inner Veto top part. 12 PMTs are positioned in an inner ring, 6 looking towards the center and the other half looking towards the outside. 12 other PMTs are positioned in an outer ring, one half looking inside the Inner Veto top part and the other half looking inside the Inner Veto lateral part. A similar configuration can be found in the bottom part, but with 24 PMTs in the external ring and 18 in the outer ring, in order to improve light collection from cosmic muons. In the lateral part there are 12 PMTs, half of these looking up and the other half looking down.

The Inner Veto vessel tank is equipped with 78 Hamamatsu R1408 PMTs fixed on its bottom, lateral surface and on the upper buffer lid. These PMTs have a 8-inch photocathode and are encapsulated in stainless steel capsules, filled with mineral oil, with a transparent PET window at their front end. This mineral oil was chosen to match the optical properties of the liquid scintillator. The PMTs distribution covered a fraction of the surface of 0.6%. Then, in order to maximize the light collection, reflective foils and white paint have been used to cover the outer buffer wall and the Inner Veto tank. This improves the reflectivity by more than a factor of 2. The PMT spatial configuration has been optimized through MC simulation in order to obtain an homogeneous detector response and maximize the light collection. FIG. 2.5 shows the PMTs configuration on the top part of the Inner Veto and FIG. 2.6 shows the PMTs configuration in the whole Inner Veto. However, this configuration was not optimized to allow vertex position reconstruction, more details on the vertex reconstruction with IV are given in sec. 2.7.2

2.2.1.3 The Outer Veto (OV)

The Outer Veto consists in 64 plastic scintillator strips of 50 × 10 × 3200 mm or 3600 mm with 1.5 mm diameter wavelength-shifting optic-fibers. The strips are

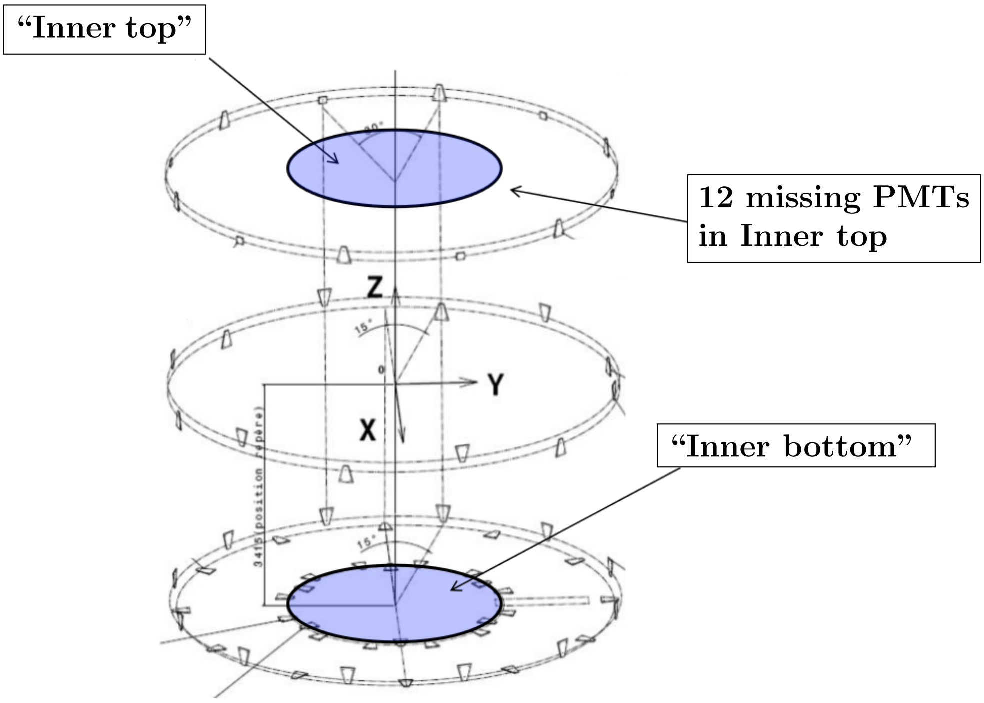


Figure 2.6: Inner Veto PMTs configuration. Squares in top part show PMTs looking inside the Inner Veto top part. There are 12 missing PMTs on this drawing located in the inner top ring: 6 looking towards the center and 6 looking outside, similarly positioned to those in the inner bottom ring as shown in Fig. 2.5.

arranged in order to form a 32 strips side plane. It is mounted above the stainless steel shield of the detector. The Outer Veto aims at detecting cosmic muons. Its dimensions exceed the IV diameter in order to detect the cosmic muons not entering the detector, as these muons could cause muon correlated backgrounds in the ID. In 2012, an upper Outer Veto was added above the glove box in order to allow the detection of cosmic muons passing by the chimneys.

2.2.1.4 The Chimneys

In order to enable the calibration tubes passage into the detector, the upper lid of each Inner Detector volumes is linked to the gloves box via a chimney. The three chimneys are nested into one another and filled with the same liquid than their linked ID volume. These chimneys can also be considered as a drawback of the Double Chooz detector as they create an hole in the Inner Veto and the Outer Veto. Low energy cosmic muons can sneak inside the Inner Detector and decay within without being detected by any of the two vetos. This issue was corrected in 2012 with the addition of the upper Outer Veto. RENO experiment [49] has similar issue.

The ZAxis tube is immersed in the ν-target volume and the Guide tube is immersed in the γ-catcher volume. More details on these calibration tubes are given in the Detector Calibration section.

2.2.2 The detector electronic systems

Far Detector electronic systems were upgraded during the first semester of 2014. The pre-upgrade Far Detector systems are called FD-I, and the post-upgrade ones are called FD-II

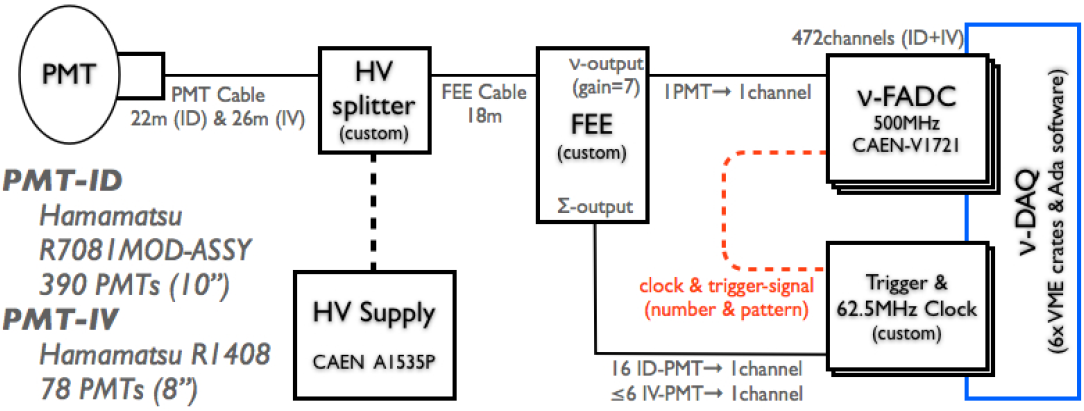


FIGURE 2.7: Schematic view of the ID and IV read-out chain.

2.2.2.1 Read-out system

The Inner Detector and Inner Veto read-out system is schematized on FIG. 2.7 from left to right. The Outer Veto read-out is an independent system.

The detector PMTs collect light emission from particle interactions in liquid scintillator. This light is converted into a millivolt charge signal by the PMTs and is carried outside the detector through a high voltage cable. Splitters then decouple the charge signal from high voltage (HV). The signal is then sent towards the front-end-electronics (FEE). The FEE amplifies the charge signal in order to send it towards the 8 bits flash-Analog to Digital Converter cards (FADC) which digitize the signal waveform, with a sampling at 500 MHz. Simultaneously FEE sums up the signal coming from groups of 16 ID-PMTs, or 5 IV-PMTs, and sends it towards the trigger system which, upon discrimination, releases the signal to the FADC. Triggered signal digitized waveforms are recorded by the data acquisition system (DAQ).

The FD-II upgrade aimed to increase the capability of the DAQ, allowing to handle higher event rates, up to ~ 600 Hz. It also allows basic event classification based on energy threshold. The FEE gains were also increased by a factor 2 with this upgrade, in order to give a more linear response for single photoelectron. More details on the importance of linearity in single photoelectron are given in chapter 4.

The FD-II and ND readout systems are identical.

2.2.2.2 The trigger system

The DC trigger system is a custom trigger system design which estimates the deposited energy by summing the signals provided by the FEE. The trigger system uses one trigger master board (TMB) which is connected to three sub-trigger boards. One trigger board is dedicated to the Inner Veto and the two others are dedicated to the Inner Detector (TB-A and TB-B).

The ID PMTs are either connected to TB-A or TB-B. The PMTs are divided in 12 geometric sectors of 32 PMTs each. For a given sector, half of the PMTs are connected to TB-A and the other half to the TB-B. Each PMT connected to the TB-A is surrounded by a PMT connected to TB-B and vice versa. Therefore, each TB look at the same volume, then both trigger boards are expected to give the same trigger decision [137].

The trigger decision is made when the signal satisfies one of the two thresholds on the deposited energy or on the number of PMTs groups activated.

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 IV. Thus the IV trigger condition is rather based on the hit pattern of the PMTs groups. As the IV was originally built to identify cosmic muons, the hit pattern in the bottom part of the IV is carefully studied.

All trigger boards send a 8 bits trigger word to the trigger master board, containing the trigger decision. The trigger master board reads four 8 bits trigger words, three from the trigger boards and one from the calibration system, the OV and the dead time monitor as an external trigger. Only one positive decision is necessary to register the signal in the flash-ADC.

2.2.2.3 The Outer Veto read-out system

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 optic 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. If the signals exceed a common threshold, the FEE send them to a 12-bit ADC. The ADC digitizes the signals which are read-out through an USB card.

The timing synchronization between the OV system and the main read-out system is performed by the TMB clock. This timing information is used in the off-line analysis in the search of a coincidence between the OV and ID / IV triggers.

2.3 The detector calibration

In order to reduce the systematics on the $\overline{\nu}_e$ detection, a very accurate knowledge of the detector is mandatory. The goal of Double Chooz is to achieve a 0.5% relative error on detection efficiency [45]. For such purpose, Double Chooz needs to determine accurately the liquid scintillator response, the detector optical model, the energy scale and the PMTs time offset and gain. These detector parameters are measured using well known particles.

The particles used in Double Chooz to calibrate the detector are light sources, radioactive sources and spallation neutrons. More details on the use of the calibration systems are given in the Energy reconstruction chapter (Chapter 3.

2.3.1 Inner Detector / Inner Veto Light Injection system

The Light Injection System consists of LED fibers running along the detector wall and positioned on the side of the PMTs photocathodes. These LED fibers are connected to a blue and a UV LED flasher. Light emission rate and intensity are controlled remotely.

There are two Light Injection Systems, one in the ID (IDLI) and one in the IV (IVLI). Both are run at least once a week to monitor the stability of the PMTs gain and timing response.

2.3.2 Radioactive Sources

Four radioactive sources are used in Double Chooz, three gamma sources, ^{137}Cs $\rightarrow$ 0.667 MeV, ^{68}Ge $\rightarrow e^+ \rightarrow 2 \times 0.511$ MeV and ^{60}Co $\rightarrow$ 1.173 and 1.333 MeV, and one neutron source ^{252}Cf. Sources have a 50 Hz emission frequency. These four sources have been deployed in the ν-Target and the γ-catcher volumes.

The deployment in the ν-Target is performed using motorized pulley-and-weight system, operated from the glove-box installed above the chimney. As this system allows to move the source only on the Z axis of the detector, it is called

ZAxis (ZA). ZAxis allows to move the source on a range span from 1 cm above the ν-Target bottom up to the ν-Target chimney, with a 1 mm precision.

The deployment in the γ-catcher is performed along a rigid loop tube using a motor driven wire. This system is called Guide Tube (GT). The GT traverses the γ-catcher by passing near ν-Target and Buffer boundaries. It allows a 1 cm precision along the loop on the source position and a 2 mm precision on the perpendicular distance between the target wall and the source.

The deployment of an articulated arm is investigated, in order to allow source deployment in the whole ν-Target volume.

2.4 The online system

The Double Chooz online system consists of the data acquisition and the processes handling the data files. The data acquisition is supervised by the shifter onsite or remotely, using an interface linked to the DAQ. Once the acquisition is done, the run files are processed and saved offsite.

2.4.1 Data taking

The Double Chooz data acquisition is performed by 5 read-out processors (ROP) connected to the FADC cards. When the trigger signal is received, the ROPs transfer the FADC 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 binary files. A binary file maximum size is about 20 GB. Since summer 2013, the trigger rate is about 250 Hz. Given an amount of data about 70 kB/trigger, for a one hour run, four binary files are written for a total size of about 65 GB.

The data acquisition is controlled by a Run Control and its Graphical user interface (GUI). This GUI is the interface between the shifter, both onsite or remote, and the data acquisition. Shifters can configure the run parameters and start and stop the data acquisition. For each run, the timing information such as start and stop time, the run length and the run-type are written in a database.

2.4.2 Data handling and processing

Onsite data handling is performed by a set of processes called “DataMigro”. The maintenance and the development of DataMigro was one of my task during this thesis.

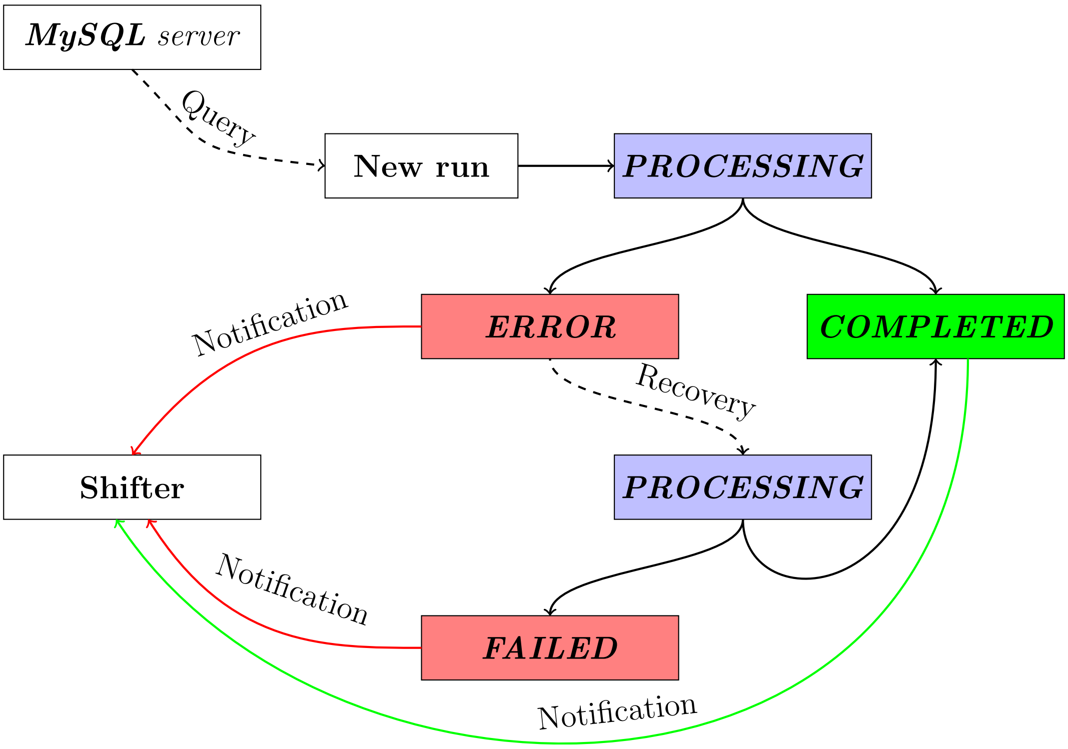


FIGURE 2.8: Schematic view of the DataMigro process structure

DataMigro is a daemon running on the background of Far Lab and Near Lab machines. It is able to perform several tasks on the files of a given run in order to process them and back them up.

All DataMigro processes have the same structure, described in FIG. 2.8. Each process runs on the background of the machines, using a cron system or a simpler bash scripts system. About every minute the process queries a local database to check if a new run can be processed. The status flag of the process 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.

There were two DataMigro configurations since the beginning of Double Chooz,

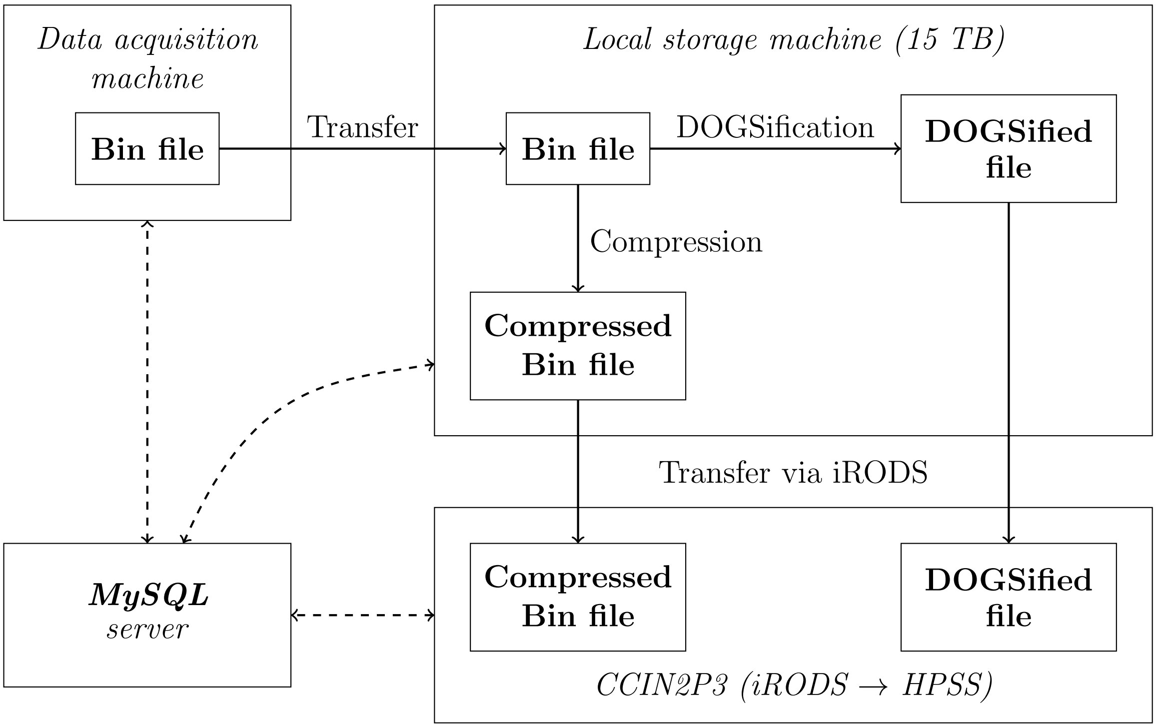


FIGURE 2.9: Schematic view of the DataMigro tasks

the first was running on until summer 2013 on. In this configuration, described in FIG. 2.9, the run binary files are transfered from the acquisition machine to a bigger local disk with 15 TB of disk space in RAID configuration. This disk acts like a buffer and is able to store about one week of raw data. Due to the low disk capacity of the acquisition machine, the binary files are removed from it once the transfer is successfully done. The network capacity allows a local transfer speed of about 100 MB/s.

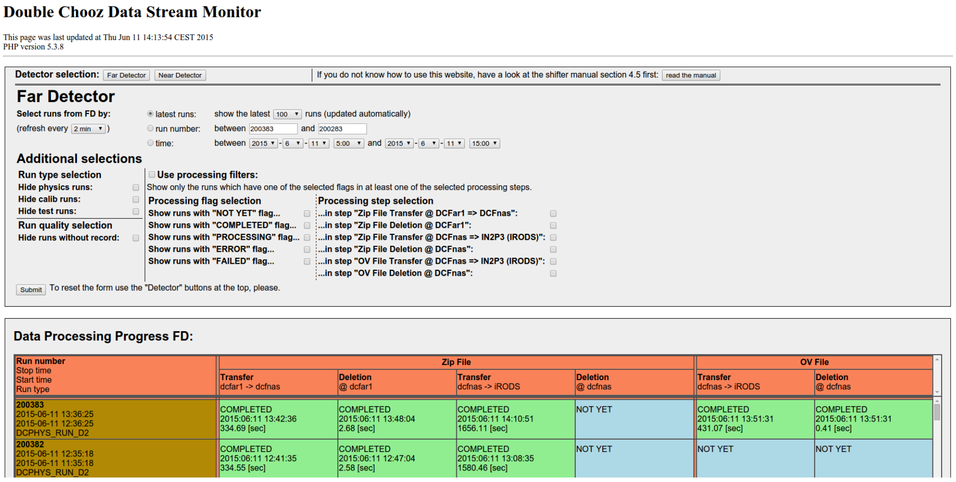
On the local disk, DataMigro applies two processes on the binary files. One of these processes is another transfer process, during which the binary files are compressed in bz2 format to reduce file size by a factor ∼5, and then transfered to CCIN2P3. The second process is called DOGSifier, it converts the binary files in a dedicated file format of the Double Chooz collaboration, based on ROOT. The DOGSified files are then transfered to the CCIN2P3 by DataMigro. Binary files and DOGSified files are removed from the local disk after their successful transfers.

The transfer to CCIN2P3 uses the Integrated Rule-Oriented Data System (iRODS), with a ∼10MB/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.

The second DataMigro configuration, running since summer 2013, is similar to the first one. The only differences are on the DOGSifier process which is now running on CCIN2P3. Then, DataMigro transfers only binary files to iRODS. This new configuration was mandatory in the FD+ND configuration, since the disk and the CPU capacity of DC machines could not handle the amount of data coming from both detectors running.

The DataMigro was installed on ND in summer 2014, and this installation leads to another upgrade of its configuration. The main upgrade is that the binary file compression is now handle directly during data taking. The machine configuration in Near Lab is also different than the one in Far Lab. In the Near Lab configuration the acquisition machine and the 15 TB local disk are the same. DataMigro was tuned to handle these different configurations.

In order to monitor the DataMigro processes a php-based framework has been developed to allow the monitoring of DataMigro status on a configurable and user friendly web page, as shown in Fig. 2.10.



Double Chooz Data Stream Monitor

This page was last updated at Thu Jun 11 14:13:54 CEST 2015
PHP version 5.3.8

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

Far Detector

Select runs from FD by:
(refresh every 2 min)
● latest runs: show the latest 100 runs (updated automatically)
○ run number: between 200383 and 200283
○ time: between 2015 - 6 - 11 5:00 and 2015 - 6 - 11 15:00

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": ☐

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**				**OV File**	
	Transfer dcfar1 -> dcfnas	**Deletion** @ dcfar1	**Transfer** dcfnas -> iRODS	**Deletion** @ dcfnas	**Transfer** dcfnas -> iRODS	**Deletion** @ dcfnas
200383 2015-06-11 13:36:25 2015-06-11 12:36:25 DCPHYS_RUN_D2	COMPLETED 2015:06:11 13:42:36 334.69 [sec]	COMPLETED 2015:06:11 13:48:04 2.68 [sec]	COMPLETED 2015:06:11 14:10:51 1656.11 [sec]	NOT YET	COMPLETED 2015:06:11 13:51:31 431.07 [sec]	COMPLETED 2015:06:11 13:51:31 0.41 [sec]
200382 2015-06-11 12:35:18 2015-06-11 11:35:18 DCPHYS_RUN_D2	COMPLETED 2015:06:11 12:41:35 334.55 [sec]	COMPLETED 2015:06:11 12:47:04 2.58 [sec]	COMPLETED 2015:06:11 13:08:35 1580.46 [sec]	NOT YET	NOT YET	NOT YET

Figure 2.10: View of the DataStream web-page, used to monitor the DataMigro status.

DataMigro development was initially done by Alberto Remoto. Romain Roncin and myself took over the development and ensure the stable operation mode in the beginning of 2013, after the end of Alberto Remoto thesis. Adrien Hourlier joined us in the beginning of 2015.

2.4.3 Monitoring system

In order to ensure the data taking quality, several monitoring tasks are performed. The first level of this monitoring is the online monitoring, which occurs during the data acquisition and is performed by a process called event builder (EBP). It ensures the faultlessness of the DAQ behavior and of the trigger system. The second level of the monitoring is a pseudo-online monitoring, which is done under the shifter supervision. It was previously performed by the DOGSifier which flagged the "bad" or "unexpected" behavior on an event and channel-wise basis. This pseudo-online monitoring was changed in summer 2013 due to its complexity and the amount of technical knowledge required to understand it. This monitoring is now performed at the end of the data acquisition. A run flagged as "bad" or "unexpected" will not be used in the offline analysis. The last level of the monitoring is an offline monitoring, which occurs during the offline data reconstruction.

Each level of the monitoring is done by many independent monitoring units, which perform plots of different observables. These plots are checked by the shifter crew to ensure the faultlessness of the data behavior. The output of these monitoring unit is available collaboration-wide to allow detector experts to diagnose and troubleshoot quickly and efficiently. This allows the experts to fix issues rapidly.

2.5 Reactor $\overline{\nu}_e$

2.5.1 Reactor operation

The $\overline{\nu}_e$ detected by DC are produced by the two Chooz-B reactor cores. Both reactors are N4 pressurized water reactor (PWR) with a thermal power about 4.27 GW_{th} each. FIG. 2.11 is a schematic view of a PWR. As suggested by their names, PWRs use light water as coolant. These reactor cores use 205 fuel assemblies each, the nuclear fuel used consists in uranium dioxide (UO_2) enriched with 3.45% of ^{235}U.

The main part of the reactor power comes from four main isotopes, ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu. Their fissions, initialized by a thermal neutron, release a mean energy of 200 MeV when the nuclei split into two or more fission-products (lighter nuclei). These fission-products are neutron rich β-emitters and initialize a radioactive decay chain of β^- emission. $\overline{\nu}_e$ are produced by the β-decays. FIG. 2.12 shows an example of a radioactive decay chain of β^- emission following a nuclear fission. Each of the four main isotopes generate a different set of fission products. Therefore, the emitted $\overline{\nu}_e$ spectra are different between each of these four main isotopes.

Fission reactions also produce neutrons, which feed a fission chain reaction. Nuclear power plant operators, like Electricité de France (EDF) for Chooz-B, aims

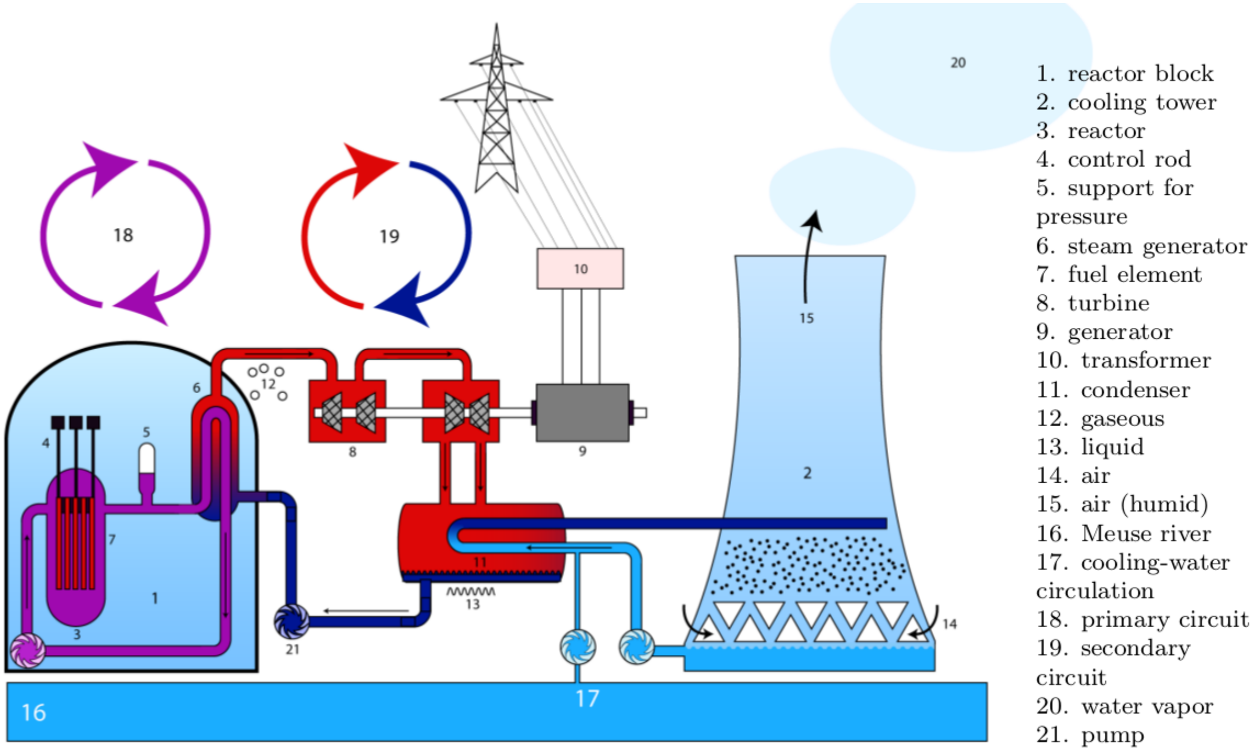


FIGURE 2.11: Schematic view of a PWR.

at maintaining a controlled fission chain reactor in order to produce energy in a safe way.

The electricity production in the nuclear power plant uses the kinetic energy deposition of reaction products (neutron and fission products). This deposition heats the water in the primary coolant circuit (label 18 on FIG. 2.11) to 320°C. To prevent water boiling, the primary coolant circuit is pressured about 150 - 160 bar. The heat is transferred through a steam generator (label 6) to the secondary coolant circuit (label 19). This secondary circuit is kept under a low pressure, allowing the water to boil and produce steam. The produced steam is used to activate the steam turbines (label 8) in order to produce electricity. Cooling of the secondary circuit is ensured by the cooling chimneys which pump water from the Meuse river to cool down and condense the water in the secondary circuit. Thermal power yield of a classic PWR is about 33%.

During the operation of a PWR there is no mixing between primary and secondary coolant circuits, nor than between coolant circuits and the external environment either. This feature is mandatory since the primary coolant is radioactive.

One third of the nuclear fuel assemblies are changed every year. For this reason,

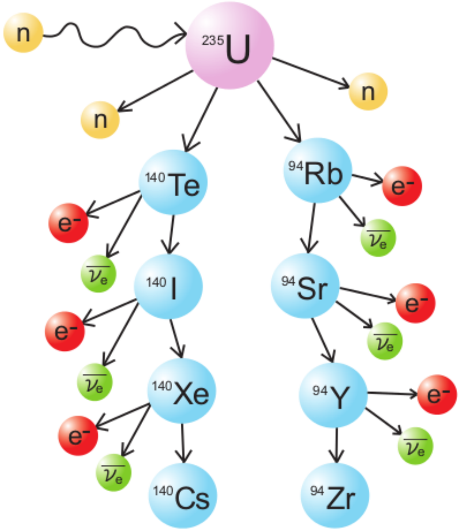


FIGURE 2.12: Example of a radioactive decay chain of β emission following a nuclear fission

each Chooz-B reactors are expected to be off one month per year. As Chooz-B has only two reactors, DC can benefit of a few periods with both reactors off. This allows the collaboration to check its analysis on signal-free data. The Double Chooz experiment is the only reactor experiment which can enjoy this feature. Daya Bay and RENO are near larger nuclear power plants [46, 49], where the probability to have all reactors off is very low.

2.5.2 $\overline{\nu}_e$ energy spectrum

Reactor $\overline{\nu}_e$ are produced by the fission-products β^- decays:

$$^A_Z X \rightarrow ^A_{Z+1} Y + e^- + \overline{\nu}_e \tag{2.7}$$

As the e^- and the $\overline{\nu}_e$ share the total released energy $E_{\beta_{max}}$, the $\overline{\nu}_e$ energy can be express as:

$$E_{\overline{\nu}_e} = E_{\beta_{max}} - E_\beta \tag{2.8}$$

Where E_β is the energy taken by the e^- (β particle). E_β is the only value which can be measured easily, as $\overline{\nu}_e$ has a very low interaction rate.

Due to the fissions occurring in the reactor, the nuclear fuel composition varies during its operation. Therefore, the emitted $\overline{\nu}_e$ flux and spectrum varies with time. The total spectrum of emitted $\overline{\nu}_e$ at a given time, $S_{\overline{\nu}_e,tot}(E,t)$, can be write as:

$$S_{\overline{\nu}_e,tot}(E,t) = \sum_k FR_k(t) \times S_{\overline{\nu}_e,k}(E) \tag{2.9}$$

Where $FR_k(t)$ is the isotope k fission rate at a given time t, and $S_{\overline{\nu}_e,k}(E)$ is the total $\overline{\nu}_e$ spectrum emitted after an isotope k fission. In a PWR, k can be ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu.

There are two different methods which allow to compute the total spectra $S_{\overline{\nu}_e,k}(E)$: *summation method* (also named "ab-initio" method) and *conversion method*.

The **summation method** consists in the summation of the emitted spectrum from all fission products coming from an isotope k:

$$S_{\overline{\nu}_e,k}(E) = \sum_p A_p \times S_{\overline{\nu}_e,p}(E) \tag{2.10}$$

Where A_p is the decay rate of the fission product p coming from the isotope k, $S_{\overline{\nu}_e,p}(E)$ is the emitted $\overline{\nu}_e$ spectrum from the fission product decays. The $S_{\overline{\nu}_e,p}(E)$ spectrum can also be deconstructed as the summation of the N_b β branches of the fission product decay:

$$S_{\overline{\nu}_e,p}(E) = \sum_b^{N_b} Br_p^b \times S^b_{\overline{\nu}_e,p}(E) \tag{2.11}$$

Where Br_p^b is the branching ratio of the β branch b and $S^b_{\overline{\nu}_e,p}(E)$ is the emitted $\overline{\nu}_e$ spectrum from the β branch b. However, in experiments, only the e^- spectrum (also called β spectrum) can be measured. Then, given the Eq. 2.8, the Eq. 2.11 should be rewriten as:

$$S_{\overline{\nu}_e,p}(E) = \sum_b^{N_b} Br_p^b \times S^b_{\beta,p}(E_{\beta_{max}} - E_\beta) \tag{2.12}$$

Where $S^b_{\beta,p}$ is the emitted β spectrum from the β branch b. The β branch spectra and the branching ratio are provided by dedicated experiments and can be found in nuclear databases. More details on the β spectrum from the β branch b and the branching ratio can be found in [129].

This prediction method allows to predict the $\overline{\nu}_e$ spectrum from all fissionable isotopes. However, this prediction implies the knowledge of many parameters, as the production rate of the fission products, the branching ratio and the maximum energy released. This gives a 10 to 20% uncertainty on the computed spectra

[72, 130]. Moreover, some fission product β decay data can be biased by a so-called "pandemonium" effect [131]. This effect comes from Germanium detectors used to measure the β spectrum of fission products which have decay branches in high energy states in the daughter nucleus. These branches are mainly underestimate due to the germanium detector low efficiency at high energy and to the low geometric efficiency, which does not allow to detect the gamma-ray cascade from the excited nucleus. This can be fixed by the use of a calorimetry method (Total Absorption Spectrometry, TAS). An international collaboration is working on this method to study the main nuclei decay properties. In [130], Fallot et al. showed that the correction of only 5 nuclei spectrum modifies the plutonium spectra by 8% in the 3 to 4 MeV range.

The **conversion method** consists in the experimental measurement of the total β spectra from the fissionable isotopes k. These β spectra are then converted into the $\overline{\nu}_e$ spectra, $S_{\overline{\nu}_e,k}(E)$, by the use of relation shown in EQ. 2.8. These measurements were done for ^{235}U, ^{239}Pu and ^{241}Pu in Grenoble ILL laboratory in the 1980s by Schreckenbach et al. [73, 74, 75, 76]. This method consists in the use of thin targets ($\sim$ 1 mg) of the given isotopes which are irradiated by a thermal neutron flux over 12 to 48 hours. Then, the spectrometer BILL was used to measure the β spectrum from the isotope.

The ILL infrastructure did not allow to measure the ^{238}U spectrum, as this isotope fissions only in fast neutron irradiation. Recently the Schreckenbach et al. method was also used by Haags et al. [132] to measure the ^{238}U spectrum, using the Munich II research reactor and a natural uranium target (99.7% ^{238}U, 0.3% ^{235}U) in order to measure a ratio between the ^{238}U spectrum and the ^{235}U spectrum. They irradiated two targets, one with a thermal neutron flux and the other with a fast neutron fluxes. During the thermal neutrons irradiation the ^{235}U was the only isotope to fission, which gave the ^{235}U spectrum. Whereas, during the fast neutrons irradiation the fissions were coming mainly from ^{238}U, with a low contribution from ^{235}U fissions.

These measurements allowed to have a compatible measurement of the ^{238}U spectrum with the ILL spectra. This spectrum was computed as:

$$^{238}U_{final} = F_{gamma} \times ^{238} U_{Haags} \times \frac{BILL}{^{235}U_{Haags}} \tag{2.13}$$

where F_{gamma} is a normalization factor, allowing to take into account the difference on the number of fissions between both experiments, $^{238}U_{Haags}$ and $^{235}U_{Haags}$ are, respectively, the ^{238}U and ^{235}U spectra coming from Haags et al. measurements. This spectrum has an energy range from 2.875 MeV to 7.625 MeV.

This method allows an higher precision on the $\overline{\nu}_e$ prediction than with the summation method. The typical uncertainty is $\sim 3\%$ with the conversion method. However, the main drawback of this method is that it measures the spectrum after a short irradiation. In nuclear reactor, as Chooz ones, the irradiation time is far longer than 48h. This can lead to off equilibrium effects due to the accumulation of long-life fission products. Mueller et al. [133] showed that these off equilibrium effects induce a 3% shift in the ILL spectra normalization. Recent discussions on the uncertainties related to this method, linked to the observed distortions (see section 1.3.2.5), hints the possibility of an underestimation of the uncertainties associated to Schreckenbach measurements [148].

In Double Chooz, the ^{235}U, ^{239}Pu and ^{241}Pu $\overline{\nu}_e$ spectra used are from P. Huber [134] with the off equilibrium correction coming from Mueller et al. [133]. For the first and the second publication, the ^{238}U spectrum was coming from the summation method shown by Mueller et al. [133]. For the third publication, N. Haags measurement was used with an extrapolation below 3 MeV and above 7.5 MeV, using a combination of the summation method and an exponential-polynomial fit on the data.

2.5.3 Reactor $\overline{\nu}_e$ flux simulation

During the first phase of the Double Chooz experiment, only FD was running. In order to obtain the un-oscillated $\overline{\nu}_e$ flux, the simulation of the reactor $\overline{\nu}_e$ flux was mandatory. The DC collaboration has developed a precise modeling of the reactor cores using the MURE framework (MCNP Utility for Reactor Evolution). MURE is a C++ interface for MCNP (Monte-Carlo N-Particle), a simulation code which models the particle transport. MURE allows to define a complex and configurable geometry and material composition, which are not featured in MCNP though they are mandatory for a proper simulation of a reactor core.

MURE most important feature is the ability to calculate nuclear fuel depletion and therefore the nuclear fission rate. Neutron transport inside the reactor is performed by MCNP and the depletion is calculated using a numerical integration of the Bateman equations via a Runge-Kutta algorithm. The Bateman equation is a model which allows to compute the nuclei activities (λ) and abundances (N) in a decay chain as a function of time. Eq. 2.14 shows the computation of the $i-th$ nucleus depletion:

$$\frac{\delta N_i}{\delta t} = -\lambda_i N_i + \sum_j \lambda_j^{j\to i} N_j + \sum_{j'} N_{j'} \langle \sigma_{j'} \rangle^{j'\to i} \langle \phi \rangle - \sum_r N_i \langle \sigma_i \rangle^{(r)} \langle \phi \rangle \tag{2.14}$$

Where N_i is the initial abundance of the $i-th$ nucleus, and λ_i its initial decay rate. j is the $j-th$ nuclei in the material, with $\lambda_j^{j\to i}$ the rate of its decays on the $i-th$ nucleus and N_j its initial abundance.

This computation is done after each MCNP execution: Particle transport is simulated via MCNP and then the nuclear fission rate is computed by MURE. The result on the fission rates is normalized by the reactor thermal power and the boron concentration in the reactor [129]. The boron concentration has a great influence over the fission rate as the boron is used as a neutron absorber inside the reactor, due to its high thermal capture cross-section.

These data are provided by EDF via the EXALT database. Due to the confidentiality of these data, the database access is allowed only onsite. In order to use these data, in-charge experts have to download the data from the onsite access and to fix the potential issues in these data. Indeed, these data can be unavailable in some cases (detector failure, data corruption...) and have to be corrected before being used by the collaboration. The time periods without data are not used to measure θ_{13}, yet it is difficult to exclude time periods from the reactor simulation. Then these data correction are critical to not bias the reactor simulation and so the fission rate prediction.

The data corrections were done by Anthony Onillon until 2013, before I took over this task.

2.5.4 $\overline{\nu}_e$ mean cross section per fission

The $\overline{\nu}_e$ flux in the detector is obtained considering their interaction cross-section. IBD cross-section formula is taking from [140]:

$$\sigma_{IBD}(E_{\overline{\nu}_e}) = E_{e^+}K\sqrt{E_{e^+}^2 - m_e^2} \tag{2.15}$$

where E_{e^+} is the positron energy, m_e, the positron rest mass, and the constant $K = \frac{2\pi^2/m_e^5}{f^R\tau_n}$ where τ_n is the measured neutron lifetime and $f_R = 1.7152$ is the phase space factor, including the Coulomb, weak magnetism, recoil and outer radiative corrections, expect the inner radiative corrections [141, 142]. Using the last measurement of neutron lifetime from [143] , $K = 0.961 \times 10^{-43}$ cm^2 MeV^{-2}.

The mean cross-section per fission is averaged over the $\overline{\nu}_e$ spectra from the different fission isotopes:

$$< \sigma_f >= \sum_k \alpha_k < \sigma_f >_k = \sum_k \alpha_k \int_0^\infty S_{\overline{\nu}_e,k}(E)\sigma_{IBD}(E)dE \tag{2.16}$$

where α_k is the fractional fission rate of the k isotope and $S_{\overline{\nu}_e,k}(E)$ is the reference $\overline{\nu}_e$ spectrum.

2.5.5 Bugey4 normalization

In the Far Detector-only configuration, the large uncertainties in the reference spectra ($\sim 3\%$) limit the sensitivity to θ_{13}. These uncertainties can be reduced using the Bugey4 rate measurement at 15m from Bugey reactor core [144] as an *effective* Near Detector, anchoring the normalization of the mean cross section per fission to this rate measurement:

$$< \sigma_f^{DC} >=< \sigma_f^{Bugey4} > + \sum_k (\alpha_k^{DC} - \alpha_k^{Bugey4}) < \sigma_f >_k \tag{2.17}$$

where the second term corrects for the different core inventories of the Bugey reactors, with respect to the Chooz reactors. The obtained correction is small ($0.9 \pm 1.3\%$), allowing to suppress the uncertainties on the reference spectra and letting the main uncertainty to come from the Bugey4 measurement, of about 1.7%.

With the Bugey4 measurement used as effective Near Detector, the reactor related systematics decrease from 2.7% to 1.7%. The comparison of the reactor related systematics with and without Bugey4 is shown in Fig. 2.13.

2.6 Detector response simulation

In order to study the neutrino selection, the detector response is modeled with a GEANT4 simulation. The electronic response is then modeled by a read-out system simulation (ROSS). The detector response simulation is calibrated as carefully as the detector itself, in order to ensure the compatibility between the simulations and the physics. This was mandatory for the first phase of the experiment, where the un-oscillated $\overline{\nu}_e$ flux was given by simulation (reactor and detector response simulations).

2.6.1 $\overline{\nu}_e$ interaction simulation

For the $\overline{\nu}_e$ interaction simulation, each physics run of the experiment is reproduced in a MC run. The number of expected IBD is computed as a function of the run length and of the simulated $\overline{\nu}_e$ flux. Each $\overline{\nu}_e$ is generated with a random position inside the reactor cores, a random interaction position inside the detector, with an energy randomly computed according to the simulated flux. Its interaction inside the detectors depends on the material proton density.

We used the Vogel et al. Inverse Beta Decay kinematics study [140] as model for the IBD simulation. The IBD itself is simulated separately as a positron and then as a neutron. A random direction vector is generated for the positron, the

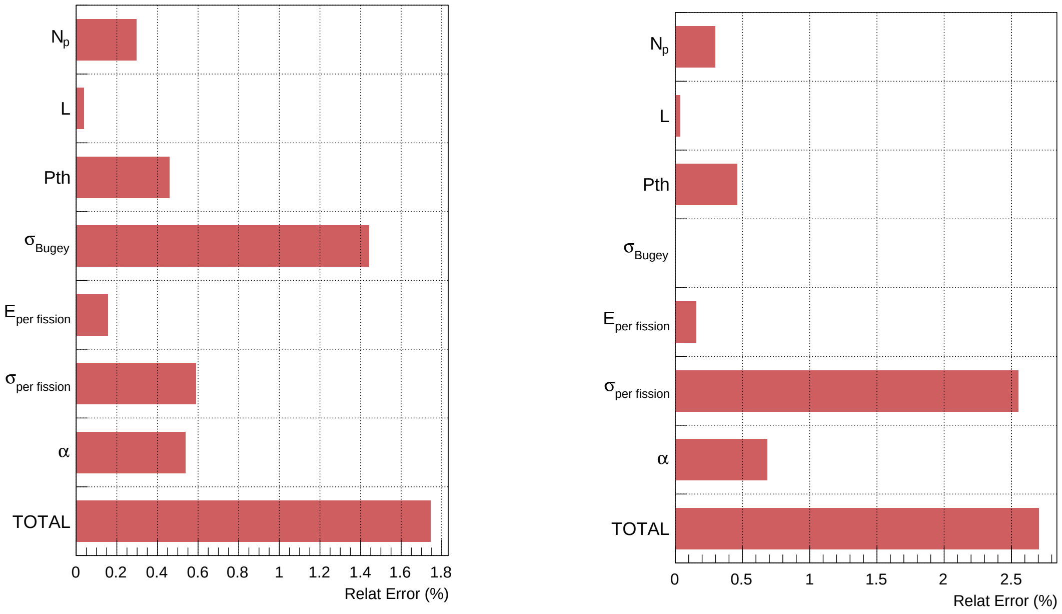


FIGURE 2.13: Reactor $\overline{\nu}_e$ rate prediction uncertainty budget. On the left with Bugey4 normalization, where the main contribution comes from this normalization, of about 1.4%. On the right without the Bugey4 normalization, where the main contribution comes from the mean cross section per fission, of about 2.5%, due to the uncertainties related to the reference spectra. The use of Bugey4 normalization allows to reduce the overall reactor related systematics from 2.7% to 1.7%.

positron and the neutron momentum vectors are defined using the interaction kinematics and the energy conservation. The GEANT4 simulation is fed with these momentum vectors to simulate the IBD.

2.6.2 GEANT4 simulation

The GEANT4 simulation itself is a DOGS module called DCGLG4sim. Its aim is to model the physics processes which occur in the DC liquids. The GEANT4 framework used in DC comes from the KamLAND experiment with improvements on the scintillation process, the thermal neutron interaction model and the photocathode optical surface model.

The liquid scintillators properties have been measured in laboratory [sources] and are periodically updated with the radioactive sources measurements to take into account the liquid aging. These properties are included in the simulation.

2.6.3 Read-out system simulation : RoSS

The GEANT4 simulation models the liquids response when energy deposition occurs. In order to model the PMTs and the electronics response when a photoelectron (PE) hits a PMT, Double Chooz developed a read-out system simulation (RoSS). RoSS takes as input some results from the GEANT4 simulation: the deposited charge and the time at which each PE hits PMT photocathode.

For each PMT, RoSS reconstructs the waveform as if it was digitized by the FADC. It takes into account the response of the different elements inside the read-out system as the PMTs, FEE, FADC, trigger system and DAQ (gains, baselines, etc.). RoSS simulates the response of the PMTs to each single PE as a probability distribution function (PDF), a normalized landau function. This PDF was measured via a dedicated setup of one read-out channel. This setup was also used to tune the design of the full read-out chain. After calibration data and MC energies agree within 1%.

More details on the technical working of RoSS can be found in section 4.1.

2.7 Double Chooz offline reconstruction

Double Chooz data are taken during the data taking shifts and then transfered to the CCIN2P3. At CCIN2P3, the offline reconstruction algorithms are applied. This offline reconstruction allows the Double Chooz collaborators to access to understandable variables and then, to perform the analysis ; the deposited energy, the vertex interaction and the muon track (in the case of a muon) for each event are reconstructed by this offline reconstruction. The final version of the data files is called *Common Trunk*.

In the previous layout, the data were formatted on site to be readable by our framework ("DOGSification") and then *CTed* (i.e. transformed in Common Trunk files) in Lyon. In the present layout, the files are *DOGSified* and *CTed* at the same time. The MC files are also *CTed* right after their production.

Spatial reconstruction

The event vertex position is a useful information to select the $\overline{\nu}_e$ during the data analysis. It allows to reject some background and then to improve the related systematics on the θ_{13} measurement.

In Double Chooz detectors, there are two separate detection volumes: the Inner Detector and the Inner Veto. Both have a different geometry and PMT distribution. Then, the Double Chooz collaboration needs two different algorithms to reconstruct the vertex position inside these volumes. During my Ph.D. work, we developed the Inner Veto reconstruction algorithm.

2.7.1 Inner Detector vertex reconstruction

Several algorithms were developed for the Inner Detector vertex reconstruction. RecoBAMA is the one used for the official analysis, thanks to its performances.

This vertex reconstruction uses a maximum likelihood algorithm on time and on charge. RecoBAMA assumes event as a point-like source of light described by $\overrightarrow{X} = (x_0, y_0, z_0, t_0, \Phi)$, where (x_0, y_0, z_0) represents the event position vector inside the detector, t_0 is the event time and Φ is the amount of emitted photon by steradian. RecoBAMA predicts the light arrival time to a PMT i as:

$$t_i = t_0 + \frac{r_i}{c_n} \tag{2.18}$$

Where r_i is the distance between the PMT i and the event vertex position and c_n is the effective light velocity in the current liquid. The amount of light received by the PMT i, i.e., the charge, is predicted as:

$$\mu_i = \Phi \times \epsilon_i \times \Omega_i \times e^{r_i/\lambda} \tag{2.19}$$

Where ϵ_i is the PMT i quantum efficiency, Ω_i is the solid angle subtended by the PMT i from the event vertex position and λ is the light attenuation length.

The likelihood used by RecoBAMA is expressed as:

$$L(\overrightarrow{X}) = \prod_{q_i=0} f_q(0; \mu_i) \prod_{q_i>0} f_q(q_i, \mu_i) \times f_t(t_i^{reco}; t_i; \mu_i) \tag{2.20}$$

where the first product is performed over all PMTs in the ID which have not been hit (i.e., with a charge q_i equal to zero), and the second product is performed over all PMTs which have been hit. The time t_i^{reco} is the reconstructed event time. f_q and f_t are the charge and time probability density functions (PDF). These PDFs were computed from MC simulation and validated with physics and calibration data. The reconstructed position is computed by finding the vector $\overrightarrow{X}_{min}$ which minimizes the negative log-likelihood function:

$$F(L) = -\ln L(\overrightarrow{X}) = -\sum_i \ln f_q(q_i; \overrightarrow{X}) - \sum_{q_i>0} \ln f_t(t_i^{reco}; \overrightarrow{X}) = F_q(\overrightarrow{X}) + F_t(\overrightarrow{X}) \tag{2.21}$$

RecoBAMA is able to compute a vertex reconstruction using only F_q or F_t, this gives a charge-only reconstruction or a time-only reconstruction. However, in order to improve the accuracy of RecoBAMA it is better to use the charge and time informations.

RecoBAMA precision was measured using calibration source and MC simulation. The resulted resolution is around 12 cm for the source positions, as shown in Fig. 2.14 for the ^{68}Ge source.

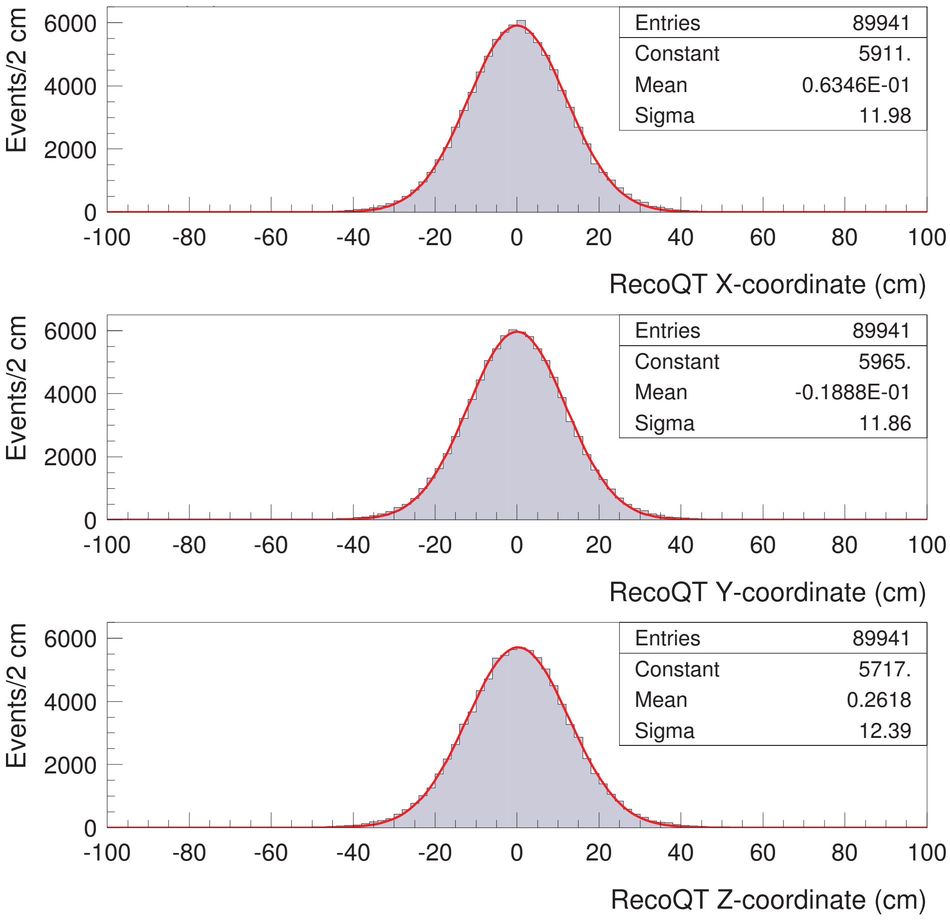


Figure 2.14: Reconstructed vertex as obtained from MC simulations of the Ge-68 source at the target center.

2.7.2 Inner Veto vertex reconstruction

At the beginning of my thesis, a reconstructed vertex position inside the Inner Veto was not available. No algorithm or method were developed to compute it. For the third analysis of DC, in order to improve the θ_{13} measurement, a new method to reduce some backgrounds, using the Inner Veto, was developed, called Inner Veto veto. To reduce the inefficiency of this method (i.e., the rejection of true $\overline{\nu}_e$

events) the use of the spatial correlation between Inner Veto and Inner Detector was mandatory (see section 5.1.4). We developed a method to reconstruct the vertex position inside the Inner Veto.

Inner Veto has a very complex geometry (as shown in section 2.2.1.2) which was not designed for a spatial reconstruction. Due to the PMT distribution, in a cylindrical distribution, the simplest reconstruction methods are not able to give a correct reconstruction. In addition, due to the high light reflections, the scintillation light from energy deposition can hit many PMTs and, like on the case of events occuring in the Inner Veto corner, the first PMT hit by the scintillation light can be the PMT located at the opposite position, as illustrated by FIG. 2.15.

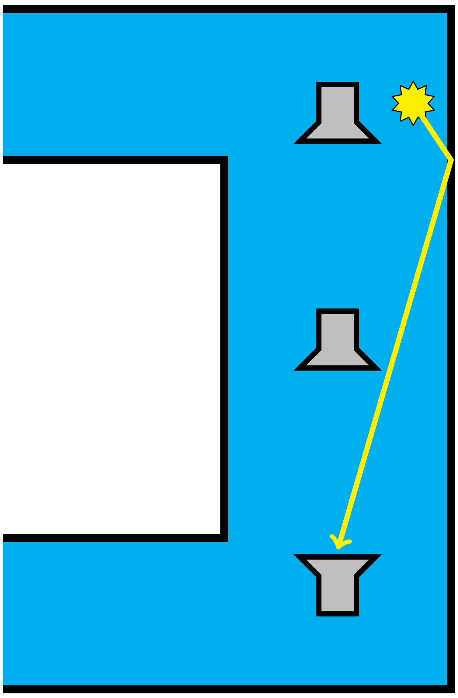


FIGURE 2.15: Schematic view of a possible scintillation light path for an energy deposition in a corner of the Inner Veto (The schematic diagram is not to scale).

Following several tries to develop an algorithm able to handle this complexity, we decided to develop an Artificial Neural Network algorithm. This was done in collaboration with Roberto Santorelli from CIEMAT (Madrid).

2.7.2.1 Artificial Neural Network

"Artificial" Neural Network are algorithms inspired by human and animal nervous system, like brain. This kind of algorithm is able to learn and to perform pattern recognition. They consist of a number of interconnected processors called neurons. The neurons process information separately and pass information to another neuron via connections. These connections have generally adapted parameters which weight the output passing along them. There are several types of artificial neural network. What they have in common is the idea of learning about a problem through relationships intrinsically present in data, rather that through a set of predetermined rules. [122, 123]

For this algorithm, we used a *feedforward multilayer perceptron*. The term *perceptron* is historical, and refers to the "activation" function performed by the neurons. *Feedforward* means that there is a definite input and a definite output, and that the data flows in one direction. This is in opposition to recurrent neural network in which data flows in a loop. This latter kind of neural network is useful when the problem is consistent with a time evolution.

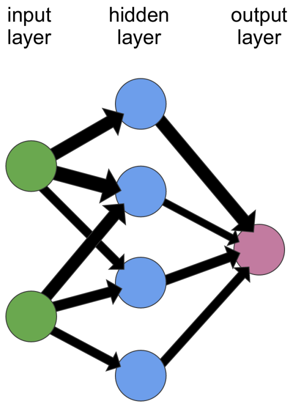


FIGURE 2.16: Simple *feedforward multilayer perceptron* with one hidden layer. The connection widths are related to the weight of each connection.

FIG. 2.16 shows a simple *feedforward multilayer perceptron* with two inputs, one output and one hidden layer. Each neuron in the input layer hold a value, x_i which can be modify by an "activation" function, such that the output of the i^{th} input neuron of F_i activation function is

$$p_i = F_i\left(x_i\right) \tag{2.22}$$

Each of the input neurons connects to every neuron in the next layer of neuron, here an "hidden" layer. And each of these connections has a weight $w_{i,j}$, associated with it. Each neuron of the hidden layer forms a weighted sum of its inputs, and passes this through an "activation" function, such that the output from the j^{th} hidden neuron of F_j activation function is:

$$q_j = F_j \left(\sum_i w_{i,j} \cdot x_i \right) \tag{2.23}$$

In the Fig. 2.16, these outputs q_j are passed directly to the output layer, but more complex neural network can have several hidden layer. In this case, the outputs are passed to the next hidden layer which performs a similar processing, until they reach the output layer.

The output layer then also performs a similar processing, so that the network output, y, is:

$$y_l = F_l \left(\sum_k w_{k,l} \cdot q_k \right) \tag{2.24}$$

Activation function of the neuron can be linear or non-linear function. In most cases "activation" function of input layer neuron are identity function. Neurons can also have an "output" function, which are simpler functions than "activation" functions. These functions are applied after the "activation" functions. In most cases and by default, the "output" function of neurons is the identity function, giving:

$$y_l = \sum_k w_{k,l} \cdot q_k \tag{2.25}$$

Artificial Neural Network training

The main purpose of Artificial Neural Networks is their learning ability. Artificial Neural Network are trained using training algorithm to extract correlations between their inputs. This training algorithm computes the appropriate value of the different ANN free parameters (weights and function parameters). This is done with a high statistics sample of events for which the outputs are already known. This is called *supervised learning.*

Learning proceeds by an error minimization taking into account the network free parameters. The typical error function used is the sum-of-squares error, described, for a single input vector, n, as [122]:

$$e^{\{n\}} = \frac{1}{2} \sum_l \beta_l (y_l^{\{n\}} - T_l^{\{n\}})^2 \tag{2.26}$$

where T_l is the target output value for the output neuron l. β_l allows to assign different weights to the different outputs, however in our case this feature was not used.

The most popular training method is the *backpropagation* method, developed by Rumelhart, Hinton and Williams in 1986 [124, 125]. The free parameters are first initialized to small random values. Each training event is then passed through the network and the outputs, y_l, are evaluated as well as the output values of all intermediate neurons in the ANN. In order to process the error minimization we determine the error gradient with respect to a weight, $w_{j,k}$, so from Eq. 2.23 and Eq. 2.26, we get [122]:

$$\frac{\partial e}{\partial w_{j,k}} = \sum_l \beta_l (y_l - T_l) \frac{\partial y_l}{\partial w_{j,k}} = \frac{\partial q_k}{\partial w_{j,k}} \beta_l (y_l - T_l) w_{k,l} \tag{2.27}$$

The remaining partial derivative can be evaluated from neurons activation function. Then, the error gradient can be evaluated with respect to any weight in the network by propagating it from the end to the beginning of the network. This is why this learning process is called *backpropagation*. The neurons parameters are then updated using this error gradient in a loop over all events in the training sample.

From the error gradient, there are several ways to train the network. In our case, we used the *Resilient backpropagation* (Rprop) [126]. This learning method is one of the fastest existing method [127]. In this method, we used a *batch* learning, meaning that we update the network parameters only after the error gradient computation for the whole training sample. In this case, our error is the average error, computed as:

$$E = \frac{1}{N} \sum_{n=1}^{N} e^{\{n\}} \tag{2.28}$$

where N is the number of event in the training sample. The average error gradient is computed from this equation. In the Rprop method, we only take into account the sign of the error gradient and update independently each neurons weights. A neuron weight $w_{j,k}$ is updated as follows:

$$\begin{array}{lll} w_{j,k}^{(t+1)} = w_{j,k}^{(t)} + \Delta w_{j,k}^{(t)} & if\ \frac{\partial e}{\partial w_{j,k}} < 1 \\ w_{j,k}^{(t+1)} = w_{j,k}^{(t)} - \Delta w_{j,k}^{(t)} & if\ \frac{\partial e}{\partial w_{j,k}} > 1 \\ w_{j,k}^{(t+1)} = w_{j,k}^{(t)} & else \end{array} \tag{2.29}$$

where t is the iteration number and $\Delta w_{j,k}^{(t)}$ is the weight-update value. For each weight, if the error gradient sign changed since the previous iteration, the weight-update is multiplied by a factor η^- with $\eta^- < 1$, and if the error gradient sign is the same, the weight-update is multiplied by a factor η^+ with $\eta^+ > 1$. If the error gradient sign changed since the previous iteration, it indicates that the previous step was too large and that the minimum was missed. Therefore, in addition to multiply $\Delta w_{j,k}^{(t)}$ by a factor η^-, the weight is reverted to its previous value: $w_{j,k}^{(t+1)} = w_{j,k}^{(t-1)}$. The weight-update values are commonly initialized to 0.1, and η^- and η^+ are commonly set to 0.5 and 1.2, respectively.

The main issue in this network training is the risk of *over-training*. A way to explain over-training, or over-fitting, is to say that it occurs when the network fit the training sample as well as possible. It means that the network will be able to guess the output for events from the training sample but not for other events.

In order to minimize this issue, we used two data samples for the training. One, the *training* sample, used as define previously, and the other, the *validation* sample, is used after. We applied the same process to the network with the validation sample but we do not update the network parameters with it. Validation sample is here to ensure that when the error with the training sample decreases, the error with the validation sample also decreases. If the error with the validation sample increases, the training has to be stopped.

One difficulty of the network training is that the training can also find no minimum in a reasonnable amount of time. It occurs when the parameters starting values are to far from their minima. To avoid it, we applied a brute force approach, which consists in retraining the network, with new parameter starting values, until getting a convergence.

2.7.2.2 Neural Network Inner Veto Reconstruction

For our algorithm, we optimized the ANN and defined a complex structure able to handle the difficulties of vertex reconstruction inside IV. We tried two types of reconstruction: one with one ANN computing the x, y and z coordinates, called *1NN*, and another one with three neural networks, one for each coordinate, called *3NN*. In our neural networks we used two different activation functions, a linear one, called Identity plus Bias and detailed in Eq. 2.30, and a non-linear one, called Logistic and detailed in Eq. 2.31.

$$F(x, bias) = x + bias \tag{2.30}$$

$$F(x, bias) = \frac{1}{1 + e^{-(x+bias)}} \tag{2.31}$$

where *bias* is a parameter which is tuned during the network training. As input for our neural networks, we chose to use the relative charge seen by PMTs as shown in Eq. 2.32. Using only the relative energy, we are loosing some informations. However, it is crucial to use it in order to allow our neural network to perform the reconstruction equally on low energy events and high energy events.

$$Q^{rel}_{PMT\,i} = \frac{Q_{PMT\,i}}{Q^{Total}_{IV}} \tag{2.32}$$

As we have 78 PMTs in IV, our neural networks have 78 "input" neurons, taking $Q^{rel}_{PMT\,i}$ as input value. The activation function of input neurons is the linear function Identity plus Bias. Input neurons are clustered in three groups, one for each PMT position: Top part of the IV, Bottom part and Lateral part. This was decided in order to use the correlation between PMTs in the same IV part to improve the reconstruction.

We have then two hidden layers. The first one consists in three groups of 10 neurons, each group linked to one input neurons group. The second hidden layer consists in 10 neurons, linked to all neurons of the first hidden layer. We have 5 neurons with a linear activation function (Identity plus Bias) and 5 neurons with non-linear activation function (Logistic) per group of 10 neurons. Due to the presence of reflections in the Inner Veto and due to the PMT distribution, we know that the problem will be both partially linear and non-linear. In the case of an energy deposition near a PMT, the problem is expected to be linear (direct path of light). However, in the case of an energy deposition in a corner, or close to a support structure, the problem is not expected to be linear (indirect path of light). Therefore, we decided to include both options and to let the algorithm learns it from correlations between the input values. In a previous configuration we used a Hyperbolic Tangent function instead of Identity plus Bias and Logistic functions in hidden layers. Hyperbolic Tangent can be linear or non-linear given its parameters. However, the convergence of our neural network was more difficult in that case and we then decided to use Identity plus Bias and Logistic.

Finally, we have one or three output neurons in the last layers, giving the x, y, or z coordinate as result. The activation function of these neurons is the Identity plus Bias function. The structure of the neural network with 1 output neuron is schematized in FIG. 2.17.

We initialized activation functions parameters and neurons weight with random numbers and trained our neural network with the Resilient Backpropagation training algorithm. We fed the trained algorithm with several independent samples of simulated 1 MeV e^-. We chose electron simulation in order to simplify the problem. As we used the relative charge, it was not necessary to use samples with different energies, otherwise we would had to use a sample with an uniform energy distribution. The training sample consisted in $100,000$ events uniformly distributed within IV and the validation sample in $19,000$ events. Such a high number of event was used in order to handle the complexity of the IV.

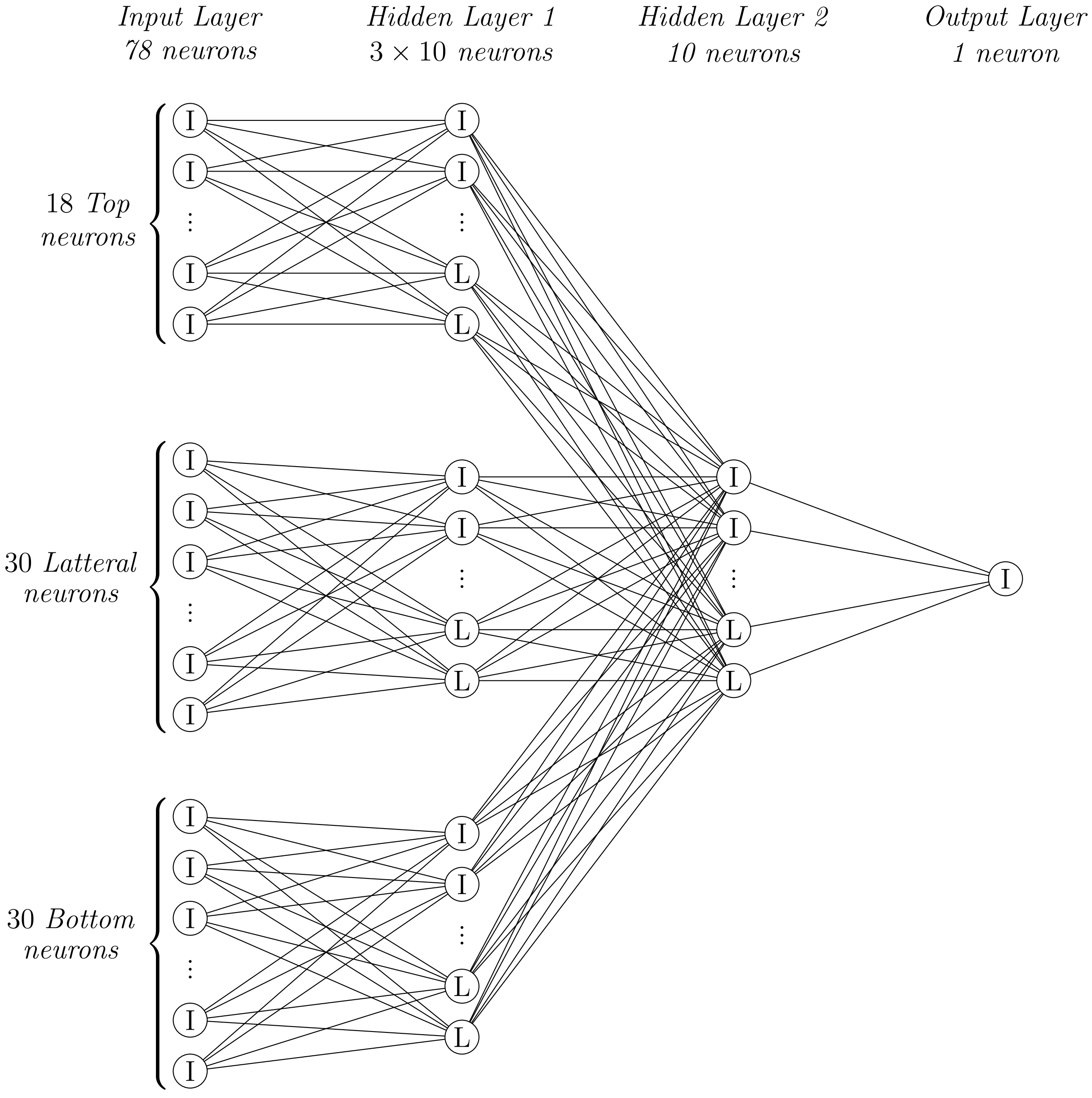


FIGURE 2.17: Schematic view of the neural network structure in the case of one output neuron. *I* stands for Identity plus Bias activation function and *L* for Logistic activation function.

2.7.2.3 Reconstruction resolution

From our neural network algorithm, we computed the reconstruction resolution with a third independent samples of $20,000$ events. Tab. 2.2 shows the resolution with the *1NN* configuration and Tab. 2.3 shows resolution with the *3NN* configuration. The performances of the reconstruction are better with the *3NN* configuration.

Coordinate	**Resolution**
X	289.9 ± 3.5 mm
Y	283.1 ± 2.9 mm
Z	353.0 ± 3.1 mm

Table 2.2: Summary of the reconstruction resolutions with the *1NN* configuration

Coordinate	**Resolution**
X	270.8 ± 3.1 mm
Y	274.4 ± 2.6 mm
Z	238.0 ± 2.9 mm

Table 2.3: Summary of the reconstruction resolutions with the *3NN* configuration

We also performed several tests to confirm the consistency of the results in different cases. One test was done to check the stability of the resolution at high energy. This was done on the *1NN* configuration and with a sample of $20,000$ simulated 10 MeV e^-, Tab. 2.4 shows the resulted resolutions. At high energy, more PMT are hit within IV, we expected a better reconstruction and this result demonstrate it. Another test was done to validate the consistency of the reconstructed position in case of an issue in the IV PMTs calibration. This shows a deviation < 1 cm.

Fig. 2.18 shows the true and the reconstructed vertex positions of simulated events with the *3NN* configuration. It is a demonstration that, even if the issue is complex, the artificial neural network is able to provide a relatively correct answer.

On the ZR^2 plot, NNIVReco tends to reconstruct events in the middle of the IV, i.e. on the planes where IV PMTs are located. This behavior can be understand as a bias due to the PMT distribution. It appears complex, even with an ANN, to discriminate events occuring on the PMT plane with events occuring above or under it (or at its left or right). One possibility to avoid this bias would be to use the light arrival time to each PMT as an input for the ANN. However, there are some doubts about the exactness of the PMT time calibration within the IV. It is also visible on FIG. 2.18 that $\sim 4\%$ of the events are misreconstructed: reconstructed outside the Inner Veto. This is a good result giving the complexity of the reconstruction within IV. A correction was proposed for such events, it is presented in the following section. FIG. 2.19 shows the reconstructed event positions for data events. It demonstrates the same behavior as with the MC.

Coordinate	**Resolution**
X	207.2 ± 2.3 mm
Y	202.9 ± 2.0 mm
Z	291.2 ± 4.6 mm

TABLE 2.4: Summary of the reconstruction resolutions with the *1NN* configuration with the 10 MeV sample

2.7.2.4 Events reconstructed outside IV

Our vertex reconstruction algorithm reconstructs $\sim 4\%$ events outside the Inner Veto. We fine tuned the algorithm with a had hoc correction for these events.

FIG. 2.20 shows the true and the reconstructed vertex positions of events reconstructed outside IV. FIG. 2.22 shows the mean distance between the true vertex position and the reconstructed one. From the first figure, it is visible that the majority of misreconstructed events are from the lateral part of the IV. From the second figure, we can see that the offset between truth and reconstructed vertex positions increases as R^2 decreases. Then, we cannot simply shift the reconstructed position to the closest IV border.

FIG. 2.21 shows distributions of the distance between true and reconstructed coordinates, X, Y and Z, for events reconstructed outside IV. We can see a non-Gaussian shape for X and Y and a Gaussian shape for Z. This highlight the difficulty of the reconstruction for X and Y. In addition, FIG. 2.23 shows the distribution of the differences between truth and reconstructed Φ, for events reconstructed outside IV, with Φ, the azimuthal angle, defined as $\Phi = \arctan(\frac{Y}{X})$. Whereas X

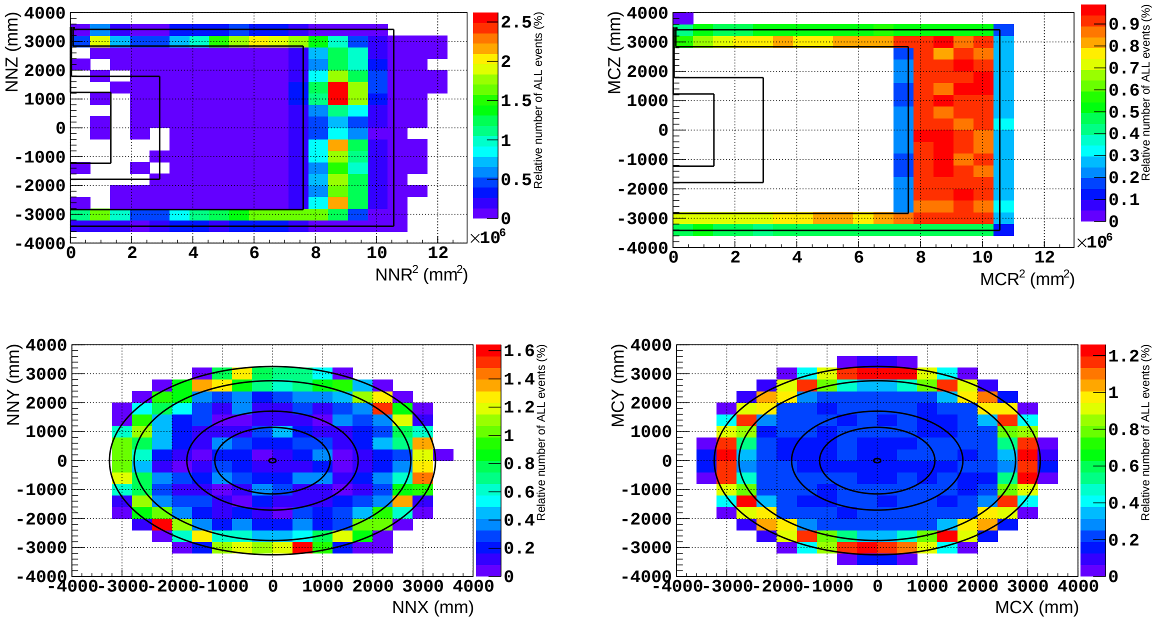


FIGURE 2.18: Reconstructed (*left*) and true (*right*) vertex position for a sample of 1 MeV electrons generated within the IV. Color scale shows the ratio of the number of events inside a given bin over the total number of events. The different lines show the detector volumes borders.

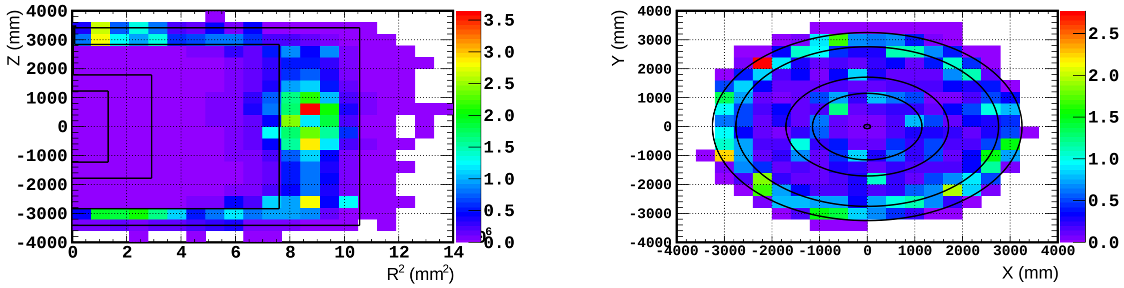


FIGURE 2.19: Reconstructed vertex position for data events. Color scale shows the ratio of the number of events inside a given bin over the total number of events.

and Y reconstruction seems to be biased, Φ reconstruction seems not to be biased, with most of the events reconstructed outside IV having $\Delta\Phi < 30\,\text{deg}$ (3σ).

From these results, we concluded that the reconstructed Φ and Z were valid. Therefore, in order to keep the Φ value we needed to apply the same correction to the reconstructed X and Y. We decided to arbitrary shift misreconstructed events

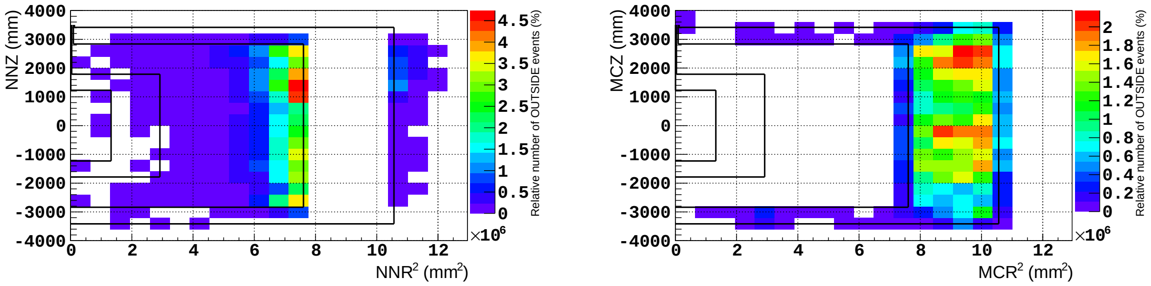


FIGURE 2.20: Reconstructed (*left*) and truth (*right*) vertex position for a sample of 1 MeV electron generated within the IV. Only events reconstructed outside the IV were selected. Color scale shows the ratio of the number of events inside a given bin over the total number of events misreconstructed. The different lines show the detector volumes borders.

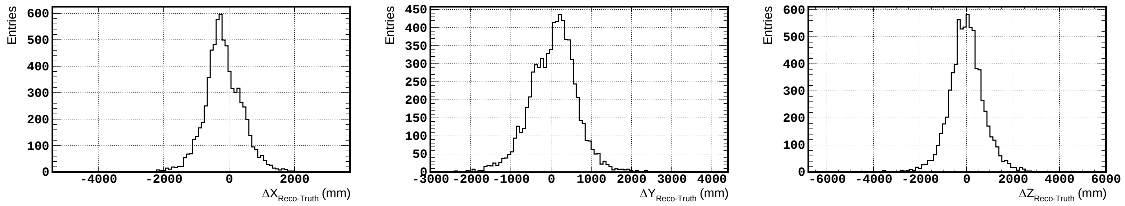


FIGURE 2.21: ΔX_{NN-MC} (*left*), ΔY_{NN-MC} (*middle*) and ΔZ_{NN-MC} (*right*) distributions for events reconstructed outside IV.

position to the middle of the IV lateral part with the same azimuthal angle, Φ. This gives the following new coordinates:

$$\vec{X}_{reco,corrected} = \begin{cases} R_{Latteral,middle} \times \cos(\Phi_{Reco}) \\ R_{Latteral,middle} \times \sin(\Phi_{Reco}) \\ Z_{Reco} \end{cases}$$

NNIVReco resolution[2] for events reconstructed outside IV was 75.96 cm. After application of this correction the NNIVReco resolution is 60.90 cm. As matter of comparison, for the full set of events (reconstructed inside or outside IV) the NNIVReco resolution is 36.31 cm before and after application of the correction. A resolution of 60.90 cm could appear high, however, we should keep in mind that this resolution affects only 4% of all events and represents an improvement of 20%

---

[2] This resolution was measured using the most probable value of a landau distribution fitted on the distance (Reco-True) distribution.

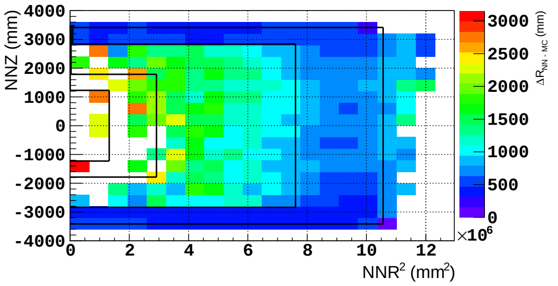


FIGURE 2.22: Mean distance between truth and reconstructed vertex positions as a function of the reconstructed vertex position.

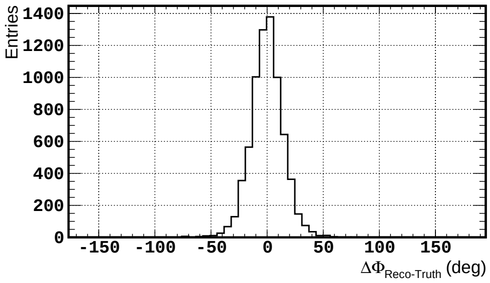


FIGURE 2.23: $\Delta\Phi_{NN-MC}$ distribution for events reconstructed outside IV. Events with a very high $\Delta\Phi_{NN-MC}$ are not shown on this figure. These later events were generated near the detector center axis where a bias on the Φ reconstruction can be expected.

compared to the previous resolution. Several other corrections were investigated, but none of them allowed a better correction than the present one [128]. FIG. 2.24 shows the reconstructed vertex position after application of the correction.

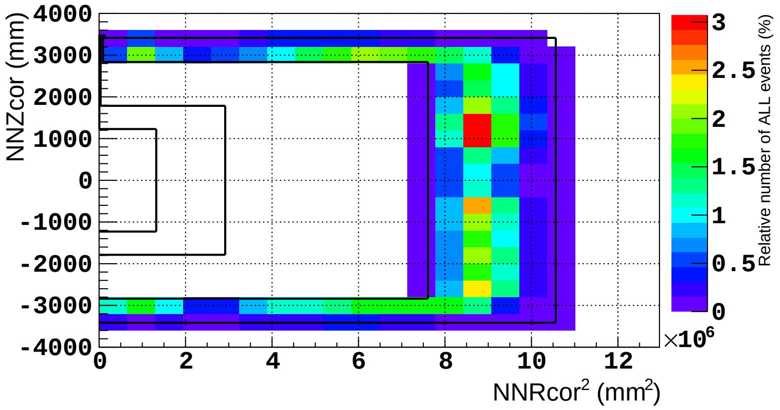


FIGURE 2.24: Reconstructed vertex position for a sample of 1 MeV electron generated within the IV. Color scale shows the ratio of the number of events inside a given bin over the total number of events. The different lines show the detector volumes borders.

2.7.2.5 Summary

This new reconstruction algorithm is able to reconstruct vertex position in the IV with a resolution of 36.31 cm. It is interesting to notice that the reconstruction is worse in the lateral part of the IV, as expected from effect similar to the one shown in FIG. 2.15.

However, we trained the network with 78 PMTs in the IV, and it appeared that 2 of these IV PMTs are dead in the FD. This fact is taken into account in the resolutions presented in section 2.7.2.3.

The reconstruction could be improved by doing a training with only active PMTs, which means to develop two set of neural network, one for FD and another one for ND.

In addition, it was shown that the NNIVReco tends to reconstruct events on the PMT plane. This bias could be corrected by using the light arrival time to each PMT as an additional input of the ANN. However, in order to use it, one should be careful and ensure the validity of the IV PMT time calibration stability over the time.

The deployment of a calibration source in the IV is not possible due to the design of the detector. Therefore, it is difficult to perform a correct evaluation of the NNIVReco algorithm with data. the validity of the NNIVReco has been evaluated on data, using IVLI calibration runs. The IVLI system does not allow to test the reflection of the IV walls, since the LED fibers are directed to a PMT, but it allows to test the consistency of the reconstruction. Consistent results were found with the different LEDs [149].

IV vertex reconstruction is used in a new method for the rejection of some backgrounds contaminating the IBD candidate sample. The development of this method is a part of my PhD work described in the Chapter 5. This method applies cuts on the distance between ID signal and IV signal of several meters (> 3 m), the bias due to a $\sim$ 35 cm resolution is then expected to be low. In addition, taking into account the fact that IV was not designed for such reconstruction, a $\sim$ 35 cm resolution can be consider as a good result.

2.7.3 Energy reconstruction

The energy reconstruction of Double Chooz is detailed in a specific chapter (Chapter 3), since a part of my PhD work was dedicated to this task.

2.8 Summary

The Double Chooz detector design, electronics, online system and offline reconstruction have been described in this chapter. The Far Detector was built between

2009 and 2010 and is taking data since April 2011. The Near Detector construction finished in 2014, and it is taking regular data since the beginning of 2015.

Before ND construction, the simulation of reactor $\overline{\nu}_e$ was required in order to perform the θ_{13} measurement. The analysis described in chapter 5 was done in this situation. This simulation introduces a systematic uncertainty of $\sim 2.7\%$, which is reduced to $\sim 1.7\%$ thanks to its normalization with Bugey4 measurement used as an effective near detector. The use of ND should decrease such uncertainty below 1% thanks to the relative comparison between ND and FD to measure θ_{13}.

During my thesis, I contributed to the data taking tasks, taking several data taking shifts. I also performed in the maintenance and development of the data transfer system for FD and ND. Finally, I developed a vertex position reconstruction algorithm for the IV, which allows to access to the vertex position in the IV with a resolution of $\sim$ 35 cm. For the last Double Chooz $\overline{\nu}_e$ analysis the NNIV reconstruction algorithm was used and mandatory to provide an additional handle on the background rejection. The $\overline{\nu}_e$ analysis is discussed in chapter 5.

Chapter 3

Energy Reconstruction

Double Chooz θ_{13} analysis is a rate+shape analysis. Therefore, detector calibration and especially energy reconstruction are critical for the success of the analysis. The IBD analysis relies mainly on neutron capture energy cuts to select IBD candidates. A good energy reconstruction, $< 10\%$, allows to increase the neutron detection efficiency (see section 5.3.2.2).

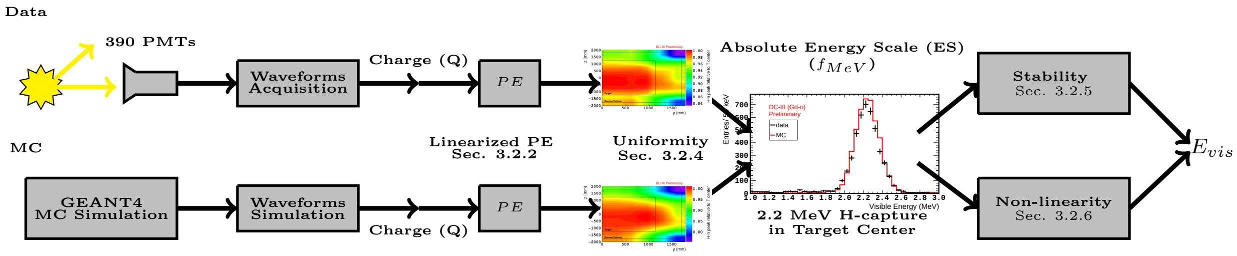


FIGURE 3.1: Scheme of the processes to compute the Visible Energy in Inner Detector. Charge extraction from waveforms is done by RecoPulse (cf. sec. 3.1.1)

FIG. 3.1 illustrates the structure of this chapter. The Double Chooz visible energy (E_{vis}) reconstruction processes are detailled. The visible energy is calculated and calibrated independently for data and Monte Carlo (MC), following the same sequence of steps and treating the MC like a second detector. During its first phase Double Chooz only relies on the Far Detector for its analysis. Therefore, the expectation of the $\overline{\nu}_e$ flux and shape without oscillation is given by Monte Carlo simulation. In order to not bias the analysis, Double Chooz has to calibrate its Monte Carlo as precisely as its detector. This is one aim of the energy calibration.

The Double Chooz energy reconstruction starts by the measurement of the number of photoelectrons produced by the PMTs. This is performed by the pulse reconstruction of the PMTs output signal and the "linearized PE" calibration. Several corrections are applied to correct for the fluctuations on the PMT output signal due to the electronics, and for non-stability, non-uniformity and non-linearity of the detector response. The extracted number of PE is converted in MeV in the process.

In this chapter, we present these different processes, and we detail how the different corrections are computed. We also describe the selections of the calibration samples used to compute the different calibrations. The validations of the energy reconstruction is described, presenting one of my studies about the Double Chooz energy reconstruction. During my thesis, I was in charge of the Uniformity correction, my work was used in the last Gadolinium and Hydrogen analysis of Double Chooz (see chapter 5). I also contributed to the determination of the non-linearity corrections, in particular the "Charge non-linearity" which was developed during my thesis.

Finally, the adaptations of the energy reconstruction, needed for the Hydrogen analysis, are presented. These adaptations led to the development of a new method to compute the uncertainty linked to the Uniformity correction, which can be used for both Hydrogen and Gadolinium analysis. During its development, we discovered a potential issue with the Uniformity correction, due to a bias in the electronics induced by the high energy depositions of the muons. A correction of this issue is presented, in addition to some hints for a future improvement. It was also one of my contributions to the energy reconstruction of Double Chooz. Some information is given about the energy model used by the final fit to extract θ_{13}.

3.1 Pulse Reconstruction

The pulse reconstruction is the computation of the charge and the timing information for each PMT in each event. We extract them from the digitized waveform. Waveform refers here to the output signal from a PMT as recorded by dedicated flash-ADC electronics (see section 2.2.2). Fig. 3.2 is an example of such waveform. When a photoelectron (PE) is produced in the PMT, a negative pulse is generated and recorded in the waveform. The amplitude of this pulse depends on the number of simultaneous PEs in the PMT, and can fluctuates due to the electronics noise.

Pulse reconstruction has to provide an estimation of the number of PE, as well as an estimation of the time when the PMT was hit. In order to do it, pulse reconstruction also has to determine, in the waveform, which part is the PE signal, so the peak, and which part is the baseline. Baseline refers to the mean intensity of the PMT output signal when no charge deposition occurs in, i.e. the flat line around 210 ADC counts (or DUI) in Fig. 3.2.

Double Chooz standard pulse reconstruction is shape independent. It relies on an integration of the waveform in order to compute the charge. Its shape independance is its main advantage, as this allows to avoid a bias due to a lack of knowledge

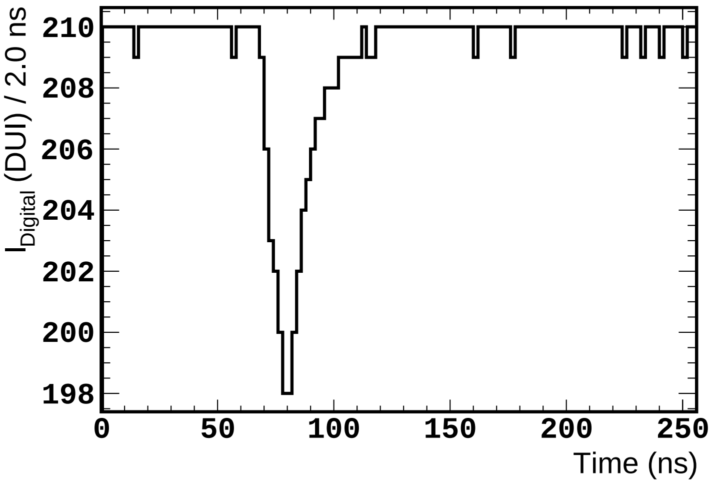


FIGURE 3.2: Double Chooz digitalized waveform from one PMT. Time binning is 2 ns. DUI stands for *digital units of intensity* and is equivalent to ADC counts.

on the electronics response. However, this is also its main limitation, as it is not able to deal with the electronics fluctuations and needs to rely on an averaged conversion factor (PMT gain) to determine the number of PE. Gain calibration (cf. section 3.2.2) is done for each PMT, and corrects for the channel to channel variations in the PE to ADC conversion. However, single-photoelectrons are in a non-linearity region where this calibration is harder. Therefore, the single photoelectron time and charge resolution are reduced. By being shape independent, this method induces a non-linearity issue in the charge reconstruction. More details on this algorithm are given in section 3.1.1. This algorithm is called *RecoPulse*.

Instead of this method, we could try to do a shape dependent method. This method, new for Double Chooz, could take advantage of our increased knowledge on the electronics response in order to perform the pulse reconstruction. I developed an algorithm based on such method, in order to reconstruct each photoelectron separatly. The goal of this algorithm, called *RecoZoR*, is to improve the resolution on the reconstructed number of PE and on the timing of these PEs. This algorithm is described in chapter 4.

3.1.1 Shape independent Pulse Reconstruction

Our shape independent algorithm performs baseline, pulse charge, and pulse time reconstruction. [120]

The baseline analysis is the first step of this pulse charge reconstruction. The baseline value is needed to compute the actual signal, using it to be subtracted from the charge. The baseline estimation is performed for each channels in two ways:

- The *floating baseline estimation*, which consists in an estimation from a time window of few nanoseconds ($\sim$ 16 ns) prior to each signal. This estimation takes advantage of the capabilites of our flash-ADC electronics. As no signal is expected in the first time sample (i.e. time bin), the baseline can be safely measured here. However, as this estimation depends on the pulse position, the time window size uses for this analysis could be limited, reducing the accuracy of the baseline estimation.

- The *external baseline estimation*, which uses special event taken with the two external triggers at a rate of 1 Hz. This estimation takes advantage of the full 256 ns time windows. It relies on the computation of the deviation between the mean intensity on each time sample intensities. The sample with largest deviation is rejected. This process is iterated until the largest and the lowest intesities of the remaining time samples are the same, within a tolerance of 1 ADC counts. However this method is known to be biased if the external trigger occurs after a large energy deposition, like muon. Nonetheless, the probility of such coincidence is low due to the muon rate (cf. section 2.2), expected to be about $\sim 2 \cdot 10^{-3}$ in the Far Detector[1].

In order to prevent these disavantages, the actual baseline is estimated using both methods, the *floating baseline* method is used by default and the *external baseline* method is used if the RMS of the baseline from the *floating baseline estimation* is larger than the one from the *external baseline estimation* by 0.5 a.u.

The pulse charge reconstruction is performed using a method called “sliding window”. This method consists in the sum of ADC counts after baseline substraction, during an integration time window. This integration time window length was choosen to be 112 ns [121] to optimize the charge resolution of single photoelectron signals, the energy resolution and the charge integration efficiency. The start time of the integration time windows is determined in order to maximize the integrated charge for each PMT for each event.

Most PMTs detect only one photoelectron for few MeV events within the Inner Detector. In order to discriminate actual photoelectron and noise fluctuations occuring in its absence, two conditions are required: $\geq$ 2 ADC counts in the maximum time sample and $q > B_{rms} \times \sqrt{N_s}$, where q is the integrated charge, B_{rms} is the baseline fluctuation as RMS, and N_s is the number of time samples in the integration window (56 for a 112 ns window).

---

[1] After muon effects biasing the baseline remain between [0, 150] μs after the muon. This estimation is done assuming the main part of such effects vanished after 50 μs. For near detector, the probility is expected to be about two times higher due to the muon rate.

3.2 Energy reconstruction

Double Chooz aims to measure θ_{13} using both rate and shape analysis. IBD analysis relies mainly on neutron capture energy cuts to select IBD candidates. In addition, the shape analysis is mostly useful to constraint the background systematics using their spectra. The combined rate and full background shape informations can be used to further constrain background systematics, leading to a significant improvement of the overall precision. Therefore, detector calibration and especially energy reconstruction are critical for the success of the analysis.

In addition, the Double Chooz analysis requires then an excellent control of the e^{+} energy scale systematics, otherwise deteriorating both rate and shape information on θ_{13}.

The Double Chooz visible energy (E_{vis}) is computed from photoelectrons (PE). It is calculated and calibrated independently for data and Monte Carlo (MC), following the same sequence of steps, as shown in the Fig. 3.1, and treating the MC like a second detector. The following equations show the relation between the E_{vis} and the total number of PE:

$$E_{vis}^{0,m} = PE^{m} \times f_{u}(\rho, z) \times f_{MeV} \tag{3.1}$$

$$E_{vis}^{data} = E_{vis}^{0,\,data} \times f_{s}^{data}(E_{vis}^{0,\,data}, t) \tag{3.2}$$

$$E_{vis}^{MC} = E_{vis}^{0,\,MC} \times f_{nl}^{MC}(E_{vis}^{0,\,MC}) \tag{3.3}$$

where $E_{vis}^{0,m}$ is a notation for the visible energy before the application of the stability correction or of the non-linearity corrections. m refers to either *data* or MC, $f_{u}(\rho, z)$ is the correction coming from the uniformity of the detector response, with (ρ, z) the reconstructed event position in the detector in cylindrical coordinates. f_{MeV} is the conversion factor from PE to MeV, extracted from the Hydrogen capture peak of neutron coming from a ^{252}Cf deployed at the center of the detector, during a long calibration run. $f_{s}^{data}(E_{vis}^{0,\,data}, t)$ is the correction coming from the stability of the detector response, with t the reconstructed event's time. Finally, $f_{nl}^{MC}(E_{vis}^{0,\,MC})$ is the correction coming from the non-linearity of the detector response. The different part of the energy reconstruction process are detailed in the following sections for the Gadolinium analysis. Adaptation of the energy reconstruction for the Hydrogen is described in section 3.2.8.

One part of my PhD works was about the energy reconstruction for the DC-III (Gd-n) analysis, it included the selection of the used event samples (spallation neutrons and radioactive sources analysis), as well as the Uniformity calibration and the computation of a part of the energy non-linearity. In addition, I also worked on the validation of the energy reconstruction.

3.2.1 Calibration event samples

The calibration of the detector consists in the comparison between the energy measured by the detector and the true energy deposited in it. In order to calibrate the detector, it is mandatory to have event samples with a well known energy position. As shown in section 2.3, Double Chooz deployed several radioactive sources in the detector during several calibration campaigns. These well-known radioactive sources are used to calibrate the energy scale and some detector parameters. The IDLI system (see section 2.3.1) is also used to calibrate each PMT parameters. The Double Chooz collaboration also uses the spallation neutron events i.e., neutron emitted by the muon spallations in the detector. These events occurred during the whole year, in the whole detector and have a well-known energy position: the Hydrogen and the Gadolinium neutron capture peaks.

All these samples are complementary. As it provides a mono-wavelength light at a given time, IDLI system is useful to calibrate the PMTs timing and the PMTs charge gain. However, as the IDLI light does not look like physic signal, it is not useful to calibrate the other parts of the energy reconstruction. Thanks to their high activity and the fact that both their energies and position are well-known, radioactive sources can provide different reference energies and position to compute the PE to MeV conversion factor and its stability over the energy scale, as well as to calibrate the position reconstruction. However, their deployment is short ($1-60$ minutes) and occurred only few days per year and as defined position. Therefore, they cannot be used to provide a calibration over the running time of the detector or over its full volume. Spallation neutrons beneficited from these advantages, since they occur in the full detector and over the whole detector's running time. However, the precision of their interaction positions depends of the reconstruction algorithm, and then does not to allow to measure with precision some detector caracteristics like the light yield.

As Double Chooz wants to calibrate its Monte-Carlo as precisely than its detector, two event samples are also used to calibrate it. The radioactive source deployment are simulated. However, due to the difficulty to simulate spallation neutrons, we used the IBD simulation to replace the spallation neutron sample.

The main informations from the calibration samples (i.e., peak position and standard deviation of the peak) are extracted from the selected sample via a fit. The fit function and the fitted energy range depend of the sample used.

We detail here the selection and the fit of the event samples from radioactive sources and spallation neutrons. The IDLI events selection is not detailled, as it was not used during my PhD works.

3.2.1.1 Radioactive sources

In order to get a sample of events from the decays of the radioactive sources deployed inside the detector, we need to separate these events from the background events. Two different strategies are used (detailled below): one for the 3 sources producing gamma rays (^{137}Cs, ^{68}Ge and ^{60}Co) and another for the ^{252}Cf source. As the studied physical events are not $\overline{\nu}_e$, we did not used all the selection cuts used for the antineutrino selection, described in the chapter 5. However, the cuts used are based on the $\overline{\nu}_e$ selection cuts.

A basic selection is applied in both strategies. This selection includes a tag of the muon events, the rejection of the events occurring in the millisecond after the muon event, the rejection of non-physical trigger (see section 5.1.2) and a basic rejection of light noise events (see section 5.1.1.3). As these selections were defined before the DC-III (Gd-n) analysis, the cuts used are from the DC-II publication [147]. More details on these cuts are given in the antineutrino selection discussion, chapter 5. The cuts used are described in the Tab. 3.1.

Some cuts applied need to use a conversion from PE to MeV. In the common case where the energy calibration is not performed yet before the selection of these samples, a conversion factor from PE to MeV is arbitrarily set at an expected value and then refined using the same method described in section 3.2.3.

Muon event	$Q_{IV} > 10^4$ a.u. **or** $PE_{ID} \times f_{MeV} > 30$ MeV
Light-noise event	$MQ/TQ > 0.09$ **or** $RMS(Tstart) > 40$ ns
Valid event	$PE_{ID} \times f_{MeV} > 0.7$ MeV **and** not a random trigger **and** not a light-noise event **and** $\Delta t_{\mu} > 1$ ms

Table 3.1: Basic selection cuts used for the radioactive source calibration runs. Q_{IV} is the charge deposited in the IV, in arbitrary units. PE_{ID} is the number of PE recorded by the ID's PMTs. MQ/TQ is the ratio between the maximum charge received by a PMT in the ID and the total charge received by the ID PMTs. $RMS(Tstart)$ is the root mean square of the distribution of the reconstructed ID PMTs starting times (i.e. the time at which they start to register a pulse). f_{MeV} is the conversion factor from PE to MeV (see section 3.2.3) or an estimation of this factor if the analysis is made before the energy calibration, Δt_{μ} is the time difference since the last tagged muon. Random triggers are used for pulse reconstruction calibration (see section 3.1) and occured at a fixed rate. The values of these cuts are based on the selection cuts applied in the DC-II analysis [147].

In addition to these basic cuts, we select only events which are reconstructed within a 1 m radius sphere around the source true position.

Gamma source selection strategy: In addition to the basic selection, we substract the remaining background using a statistic method. We apply the same basic selection on a control run took in the same conditions than the calibration source (i.e., few hours before or after and with the same detector configuration) but without radioactive source deployed. The energy spectrum obtained is scaled to the duration of the calibration run and then subtracted to the energy spectrum obtained from the calibration run.

^{252}Cf selection strategy: The ^{252}Cf spontaneous fissions emit gammas and neutrons. We are interested only in the neutrons coming from these decays, as the emitted gammas do not have a well-known energy. Due to this simultaneous emissions, the signal will be a coincidence between a prompt event (the emitted gammas) and a delayed event (the neutron capture). Then, the statistical background subtraction used in the previous strategy cannot be applied to the ^{252}Cf selection. Instead, the background can be reduce by the selection of pair of prompt-delayed events in correlations. The following time coincidence cut is applied: $\Delta t \,\epsilon\, [2\,\mu s, 1000\,\mu s]$, where Δt is the time difference between the delayed and the prompt event. ^{252}Cf emits an average of 3.73 neutron per fission [110]. It is possible, for a given fission, a first pair of events where the prompt is the gamma emission and the delayed a first neutron capture, followed by a second pair where the prompt is the first neutron capture and the delayed is a second neutron capture, etc. This feature allows to maximize the statistics and allows us to not try additionnal method to reject background.

The different fits used to extract information from the radioactive source runs are:

- For the 3 gamma sources (^{60}Co, ^{68}Ge and ^{137}Cs), a first fit with a Gaussian function is applied to measure the maximum value position. Then, the fit depends on the sources:
  - ^{68}Ge: the used fit function is a Gaussian function, with a fitted range of $[max_{fit} - 0.08\,MeV, max_{fit} + 0.08\,, MeV]$, the expected peak position is $\sim$ 1.022 MeV [146].
  - ^{137}Cs: the used fit function is a Gaussian function, with a fitted range of $[max_{fit} - 0.08\,MeV, max_{fit} + 0.08\,, MeV]$, the expected peak position is $\sim$ 0.661 MeV [146].
  - ^{60}Co: the used fit function is a Gaussian + tail function, with a fitted range of $[max_{fit} - 0.7\,MeV, max_{fit} + 0.25\,, MeV]$, the expected peak position is $\sim$ 2.505 MeV [146].

- For the ^{252}Cf, each sample is fitted two times, one for the Gd capture peak and the other for the H capture peak:
    - For the H capture peak: the used fit function is a Gaussian + tail function, with a fitted range of $[2.\,MeV, 2.44\,MeV]$, the expected peak position is $\sim$ 2.223 MeV.
    - For the Gd capture peak: the used fit function is a double Gaussian + tail function, with a fitted range of $[7.54\,MeV, 8.34\,MeV]$, the expected peak position is $\sim$ 7.937 MeV.

A Monte-Carlo simulation of the calibration runs is performed. We apply the same selection cuts than for data runs on the resulted Monte-Carlo runs. However, as no background is expected in Monte-Carlo runs, the statistical background subtraction is not applied on gamma source runs.

3.2.1.2 Spallation Neutron

The spallation neutrons are a useful source to characterize the detector response. As muon spallations occur every second in the detector and in the whole detector, the neutrons from these spallation can probe the detector response over time and over space (cf. section 3.2.4 and 3.2.5).

The spallation neutron capture rate is expected to be high: as illustration, according to this reference [112] at 500 m.w.e. the expected spallation neutron capture rate on Hydrogen was measured to be 8.0×10^4 $events/kton/day$ so $\sim$ 0.9 $Hz/kton$. In the case of Double Chooz, the Far Detector has a Target mass of 10.16 tons [113]. Given the DC-III (Gd-n) live time of 467.90 days, a simple computation gives a number of spallation neutron capture of $\sim 3.75 \times 10^6$. This number is a very rough estimation: as the Far Detector is located at 300 m.w.e. the rate should be far higher. In addition, due to the mixing between Gadolinium and Hydrogen in the Target volume, the neutron capture cross-section is higher than the one used in the reference case. This should lead to a higher expected spallation neutron capture rate. Thank to this high statistical power, a high efficiency selection, like the one need for $\overline{\nu}_e$ is not mandatory.

To select spallation neutron capture events, we need to tag the muon which produced them and then select the following events. The Inner Veto allows to detect the muons crossing the detector, as well as fast neutrons, i.e. spallation neutrons, produced by muon in the surrounding rock. Both should be tagged as "muon" for the purpose of spallation neutron selection: since the number of fast neutrons produced by muon spallations is high [118], we can expect fast neutron interactions in the IV simultaneously with fast neutron interactions in the ID. This is more detailed in section 5.1.4. Due to the hole in the Inner Veto, i.e. the chimney, this volume cannot detect all muons crossing inside the detector. In the Double-Chooz analysis a cut on the deposited energy inside the Inner Detector can also allow to tag muons. However, this last cut tags also Michel electron from stopping muon decays. In order to avoid the Michel electron selection, a tighter cut on this ID deposited energy should be applied in the spallation neutron capture selection. Then, the muon definition for the spallation neutron is:

- $Q_{IV} > 10^4$ a.u. (from DC-II analysis)
- $E_{ID} > 70$ MeV or $PE_{ID} \times f_{MeV} > 70$ MeV before the energy calibration

The 70 MeV cut was chosen as, according to various references [114, 115], the Michel electrons maximum energy is expected to be around 60 MeV.

Then, given this muon definition, we select all the events which validated the criteria from TAB. 3.2. Due to the amount of deposited energy by the muon, the detector electronics is overloaded right after the muon crossing. Then, the first $50\,\mu s$ following the tagged muon are not used. A selection with a $[30\,\mu s, 3\,ms]$ time-window was also done and used for a test of the detector response Uniformity calibration, see section 3.2.4. However, my studies for the last Double Chooz Hydrogen analysis (DC-III (H-n)) shown that the $50\,\mu s$ time cut was not enough to avoid the electronic overload effect, the analysis and its results are discussed in section 3.2.8.1.

Light-noise event	$MQ/TQ > 0.09$ **or** $RMS(Tstart) > 40$ ns
Valid event	$PE_{ID} \times f_{MeV} > 0.7$ MeV **and** not a random trigger **and** not a light-noise event **and** $\Delta t_\mu \in [50\,\mu s, 3\,ms]$

TABLE 3.2: Selection cuts used to select the valid events for the spallation neutron selection

The selected events are then divided in 9 time-window sub-samples of $300\,\mu s$. Only the first sub-sample, with $\Delta t_\mu \in [50\,\mu s, 450\,\mu s]$, will be used as signal for the different analysis. Six of the other sub-samples are used as background sample for a statistical background subtraction. The $[450\,\mu s, 750\,\mu s]$ and the $[750\,\mu s, 1050\,\mu s]$ time-window sub-samples are not used as background sample as some spallation neutrons are expected in these samples.

A similar selection is made with the Monte-Carlo, using the antineutrino simulation. Due to the difficulty to simulate spallations, we use the neutron capture events from the antineutrino MC. We select only neutron captures occurring in a time-windows of $\Delta t \in [50\,\mu s, 450\,\mu s]$ after the prompt event, i.e. the positron annihilation. No statistical background suppression is done with the Monte-Carlo, as no background is simulated in it.

Each sample can be then fit with the same functions and energy ranges than for ^{252}Cf neutron captures.

3.2.2 PE calibration

The total number of PE is given by the following relation:

$$PE^m = \sum_i pe_i = \sum_i Q_i / gain_i^m(Q_i, t) \tag{3.4}$$

Where pe_i is the number of PE for the readout channel i, Q_i is the reconstructed charge from RecoPulse for the channel i and $gain_i^m(Q_i, t)$ is the channel i gain, i.e., the charge-to-PE conversion factor extracted by calibration, taking into account both its time variation and charge dependence, i.e., its gain non-linearity. Some quality criteria are used to select only *good channels*. The gain non-linearity is due to the limited sampling of the waveform digitizer, which can bias the channel baseline estimation within 1 ADC count [116]. The gain for each channel, at given charge Q_i and a given time t, is given by the following function:

$$\begin{aligned} gain_i^m(Q_i, t) &= g_i^m(t) + l_i^m(t) \times (Q_i - c_i^m(t)) & (Q_i < c_i^m(t)) \\ &= g_i^m(t) & (Q_i \geq c_i^m(t)) \end{aligned} \tag{3.5}$$

Where (g_i^m, l_i^m, c_i^m) parameters are determined for each channel i from a fit of the measured gain for different charge at different time t.

The PE calibration, called *linearized-PE calibration*, aims to estimate, for each channel in an event, 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).

Double Chooz found that the gain and its non-linearity change after each power recycle of readout electronics [116]. In many experiments, the PMT gains are extracted by measuring single PEs, which in our case is in the nonlinear regime. Instead, in Double Chooz the PMT gains are measured in a wider PE range using the IDLI calibration system (IDLI) several times per week (cf section 2.3.1). Thereby the charge non-linearity, as well as its time dependency, is fully characterized on a per-channel basis. This was one of the major improvements in data-MC agreement over a previously adopted Single PE calibration. The relation between the calibration data and the gain is given by the following function:

$$gain = \alpha_0 \times \frac{\sigma_i^2}{\mu_i} \tag{3.6}$$

where μ_i and σ_i are the mean and standard deviation (RMS) of the observed charge distribution from the calibration data. α_0 is used to correct for the intrinsic spread in σ_i due to Single PE width and electronic noise. It is considered to be constant for all readout channels and is chosen by making the number of photoelectrons in the Hydrogen capture of spallation neutrons equal to the hit PMT multiplicity (n).

α_0 is the mean value of the $\alpha = n/PE$ distribution. Non-single PE contributions are taken into account using Poisson statistics with a *Poisson corrected number of hit PMTs*. It is computed from the number of *good PMTs* (N_{PMT}) (i.e. excluding dead or killed PMTs) and the number of hit PMTs (N_{hits}) as:

$$n = -N_{PMT} \times \ln(1 - \frac{N_{hits}}{N_{PMT}}) \tag{3.7}$$

The correlations between the total number of photoelectrons for $\alpha_0 = 1$ and the Poisson corrected number of hit PMTs is shown in Fig. 3.3.

For the DC-III (Gd-n) period, α_0 was measured to be 1.053 and to be constant for all readout channels and in time to equalize the total number of photoelectrons to n. This measurement was done using spallation neutrons captured on Hydrogen. Details on the spallation neutrons selection are given in the section 3.2.1.2. Fig. 3.4 shows, for a typical readout channel, the measured gain, overlaid with the gain correction function, from Eq. 3.5, as a function of the integrated charge. The gains for the MC simulation are extracted following the same procedure as the ones for data.

3.2.3 Absolute Energy Scale calibration

The absolute energy scale f^m_{MeV} was measured using the ^{252}Cf neutron capture peak on Hydrogen in photoelectron. This was done with a “candle” run where the ^{252}Cf source was deployed at the center of the detector at the 398th day since data taking started. A simple computation allows to obtain the PE to MeV factor:

$$(f^m_{MeV})^{-1}\ (\text{PE/MeV}) = \frac{PE_{ID}\ (\text{PE})}{2.223\ (\text{MeV})} \tag{3.8}$$

$(f^m_{MeV})^{-1}$ was found to be 186.2 PE/MeV for the data and 186.6 PE/MeV for the Monte-Carlo. As shown in Fig. 3.5 the Hydrogen capture peak is at 2.237 MeV in data and at 2.241 MeV in MC. This gives a data-MC agreement $< 0.25\%$.

The ^{252}Cf analysis was done as explained in section 3.2.1.1.

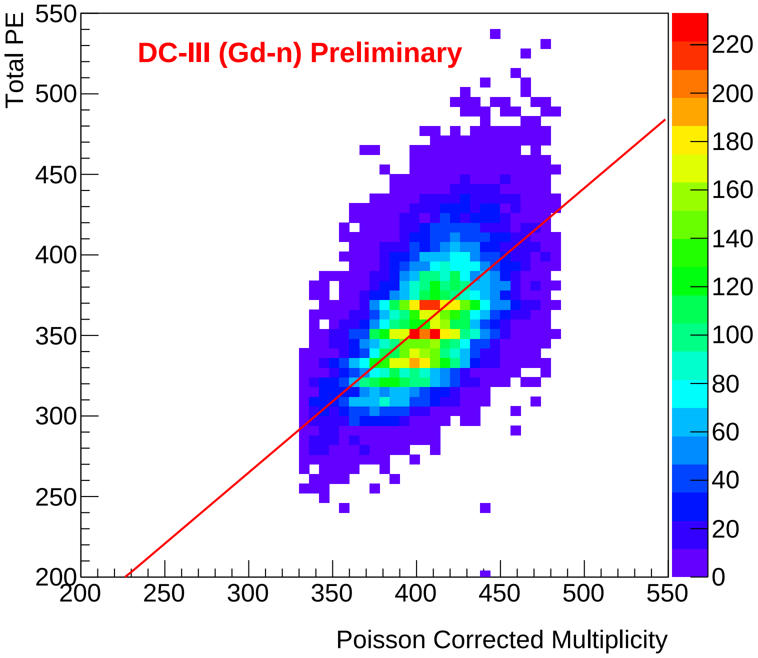


FIGURE 3.3: Correlations between the total number of photoelectrons for $\alpha_0 = 1$ and the Poisson corrected number of hit PMTs.

A second method can be used to measure the absolute energy scale. Using the spallation neutron sample it is possible to select a sample of spallation neutrons captured close to the center of the detector. The fit, performed as explained in section 3.2.1.2, allows to measure an estimation of f^m_{MeV}.

3.2.4 Uniformity calibration

The Uniformity calibration, f^m_u was introduced to correct for the position dependence of the number of PE, PE^m. It relies on the vertex position reconstruction algorithm RecoBAMA. The correction is applied as a function of ρ and z to convert PE^m into that at the center of the detector. It was one of my tasks during this Ph.D.

This correction was computed as a map, which gives the correction value for each (ρ, z) position in the detector. It was done by using the 2.223 MeV spallation neutron captures on Hydrogen, selected as described in section 3.2.1.2. In order to use the γ-catcher to measure the escaping γ from IBD reactions inside the ν-target, we need to calibrate the full detector (i.e. including the ν-target and the γ-catcher).

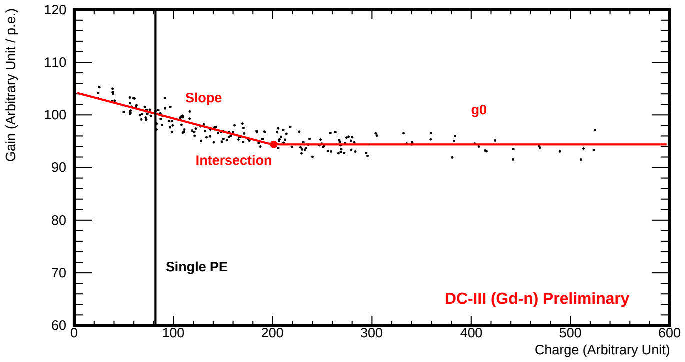


FIGURE 3.4: Gain vs. integrated charge for a typical readout channel. Black dots show the measurements of the gain, using calibration data from the IDLI system with different light injection positions and light intensities (i.e. different charges). Red line shows the gain function obtained from a fit, which is applied to convert integrated charge to number of photoelectrons. The gain is determined with the three parameters from EQ. 3.5. On this figure, g_0, *slope* and *intersection* refer, respectively, to the g_i^m, the l_i^m, and the c_i^m of EQ. 3.5. The black line shows the position of one photoelectron.

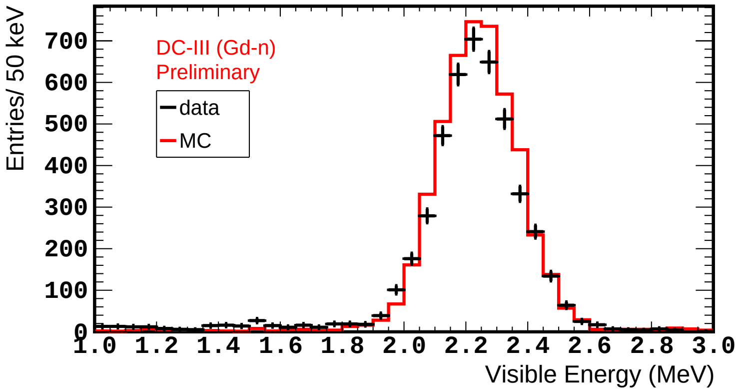


FIGURE 3.5: Visible energy of the Hydrogen neutron capture peak from ^{252}Cf fissions.

Therefore, we cannot use the spallation neutron captures on Gadolinium to build the map.

This method needs a high number of events in order to get a correct fit of the Hydrogen capture peak. The detector volume was divided in several sub-volumes and a spallation neutron selection was done for each sub-volumes. The detector was taken as a two dimensional projection of a cylinder over the (ρ, z) plane. The detector is divided in 10 subdivisions over ρ and in 14 subdivisions over Z, leading to a minimal number of events in a given sub-volume of ~ 1000. The detector response was approximated to be uniform over the azimuthal angle ϕ. This assumption is supported by several studies which shown that, for a given (ρ_i, z_i) division, the deviation over ϕ, from the mean value of the detector response, is lower than 2%.

From the selection sample of 2.22 MeV spallation neutron captures on Hydrogen, the peak mean position is extracted with the fit (cf. section 3.2.1.2). The detector center is taken as reference, then, the correction values for a (ρ, z) position in the detector is computed with the following formula:

$$f_{corr}(\rho, z) = Peak\,mean(\rho, z)/Peak\,mean(0, 0) \tag{3.9}$$

With $Peak\,mean(0, 0)$ is the fitted peak position for the detector center. $f_{corr}(\rho, z)$ is determined from the fit for the center of each subdivisions, it is then linearized using interpolation between the known values. The resulted correction factor is shown as a 2D figure: FIG. 3.6. The Uniformity correction map used in the DC-III (Gd-n) analysis was tagged as *ESv5* map.

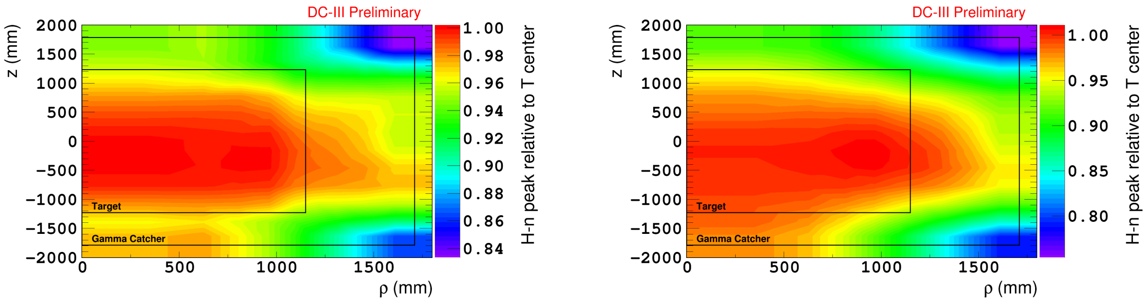


FIGURE 3.6: Uniformity correction maps for data (*left*) and MC (*right*). The color scale show the correction factor.

3.2.4.1 Uniformity uncertainty

As explain above, one of the energy calibration aims, and so of the Uniformity correction, is to have a MC as close as possible as the true DATA, in order to not bias the $\overline{\nu}_e$ spectrum and so θ_{13} measurement. Then we compute the uncertainty associated to the Uniformity correction by comparison between the MC reconstructed energy and the Data reconstructed energy. In order to do it, we need an independent sample, different than the one used for the Uniformity maps generation. For the Gadolinium analysis the independent samples are the fast neutrons captured on Gadolinium in Data and the IBDs with a Gadolinium neutron-capture in MC.

Then, for each (ρ, z) positions we compute the asymmetry between the E_{vis}^{MC} map and the E_{vis}^{DATA} map:

$$Asymmetry\,[\%] = 2 \times \frac{E_{vis}^{MC} - E_{vis}^{DATA}}{E_{vis}^{MC} + E_{vis}^{DATA}} \times 100 \tag{3.10}$$

The results are shown FIG. 3.7 for the Gadolinium analysis. The projection of this asymmetry over a one dimension plot allow us to have the uncertainty. The deviation of the asymmetry, i.e., the RMS, is taken as the Uniformity uncertainty:

$$\delta_{Uniformity} = 0.36\% \tag{3.11}$$

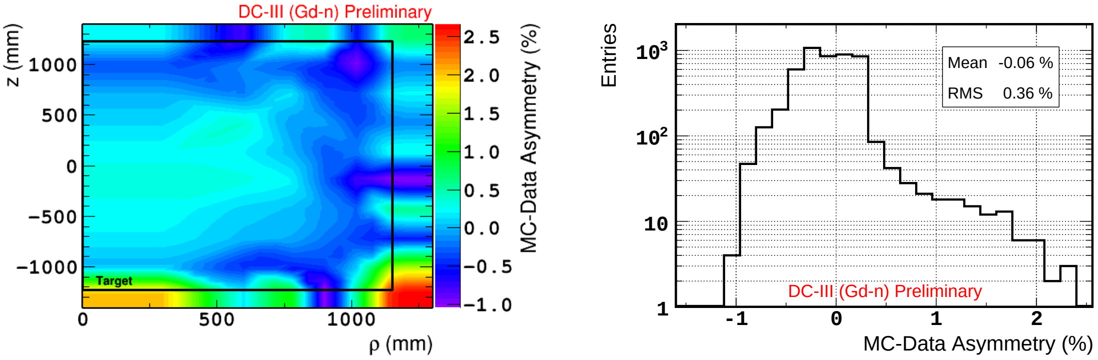


FIGURE 3.7: *Left*: 2D figure of the asymmetry between E_{vis}^{MC} map and E_{vis}^{DATA} map at the Gadolinium neutron-capture energy, the color scale show the asymmetry in percent. The black line shows the border of the ν-target. *Right*: Projection of the left plot in a one dimension figure.

This represents an improvement by $\sim$ 16% compared to the previous Uniformity uncertainty used in [147]. In addition, the mean of the distribution, with a value of -0.06%, demonstrates the good agreement between the DATA and the MC. This is discussed in section 3.2.7. Increases of the asymmetry can be seen of the 2D figure at the borders of the ν-target. This is exptected as the fit of the Gadolinium peak is more difficult for samples selected close to the ν-target borders. This is interpreted as due to the lack of statistics and to some geometric effects (like the difference of light yield between the ν-target and the γ-catcher liquids). A similar issue exists for the Hydrogen capture peak fit at the γ-catcher borders, however, here, the geometric effects should be mainly the energy lost in the buffer. The rectangular shape of the peak in the one dimension figure of FIG. 3.7 is not fully understood. Investigations were performed [150] and were not able to state if its come from a code artefact or a coincidence.

3.2.4.2 After Muon analysis

When a cosmic enters in the detector a high amount of energy, i.e. of light, is deposited in the detector, coming from the photoelectric showers, the Cerenkov effect and the nuclear reactions. Due to this high signal, the detector electronics are saturated in the first microseconds after the muon passage. This saturation biased the events reconstructions in this time window. Then, in order to avoid to use biased events, we do not use the events in the first 50 μs after a muon. However, as shown in Fig. 3.8 and in Fig. 3.9, some events 50 μs after a muon may suffer from this electronic bias, which leads to a trigger inefficiency. Due to the higher muon rate in the Near Detector, this point was a major concern for the Near Detector analysis. In order to prepare the Near Detector analysis, we investigate on the possible bias due to the electronic saturation inside the energy correction. I performed then the so-called "After Muon" analysis.

In reference [117], the author showed that the mean value of the baseline (cf. section 3.1.1) of events in the first microseconds after a muon deviates from the global mean value of the baseline. He showed that we could discriminate biased and unbiased events using a map of the baseline deviation as a function of the muon energy and of the time between the muon and the event.

In order to build this map, we selected all events with the same conditions as explained in section 3.2.1.2, except the condition on the time between the muon and the event which was modified as $\Delta t_\mu \in [0\,\mu s, 1\,ms]$. For each events we registered the baseline value, as well as the muon energy and the time between the event and the last muon. We reported these results inside Fig. 3.10.

We estimated the baseline range were the events are unbiased by using the events with a $\Delta t_\mu > 400\mu s$, as these events were expected to be unbiased, as shown in Fig. 3.8. The baseline of these events is shown in Fig. 3.11. The fit of the baseline distribution gives an estimation of the baseline range for the unbiased events of 210.020 ± 0.042. The range size was determined at 3σ in order to avoid the rejection of correct events.

We generated an Uniformity correction map as described in section 3.2.4, using a looser cut on Δt_μ and the After Muon Flag map to select Fast Neutrons:

Valid event	$PE_{ID} \times f_{MeV} > 0.7$ MeV **and** not a random trigger **and** not a light-noise event **and** $\Delta t_\mu \in [30\,\mu s, 3\,ms]$ **and** After Muon Flag value ∈ [209.078, 210.062]

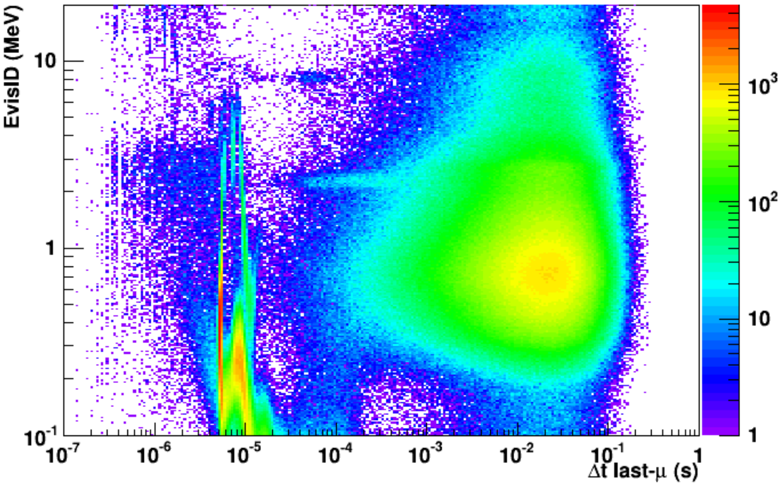


FIGURE 3.8: Energy deposition after a muon for the Far Detector I. It is possible to see Hydrogen neutron captures at 2.223 MeV, as well as a thin line at ~ 8 for Gadolinium neutron captures. The region of interest for the IBD analysis is the bunch of events occuring ~ 1 ms after the muon. After pulse events due to the muon occured mainly in the first $20\,\mu s$ after the muon. However, other after pulse events occured between $20\,\mu s$ and $100\mu s$ after the muon. The present analysis was done with the first configuration of Far Detector electronics. Note that with the Far Detector electronics upgrade (FDII), some of these last after pulse effects disappear.

FIG. 3.12 is the resulted Uniformity correction map. The comparison with the official Uniformity correction data map, FIG. 3.13, shows that the two maps are very similar with a deviation mainly below 0.1%, with some regions between $[0.1\%, 0.2\%]$ close to the γ-catcher borders. This increase of the asymmetry at these positions is due to the lack of statistics and to some geometric effects (as explain above). Then, this analysis demonstrates that the official method to compute the Uniformity correction of the energy scale is not affected by the electronic saturation due to the muon energy deposition. This was an important result in the perspective of the Near Detector analysis, meaning that the current method to compute the Uniformity correction for the Far Detector was still valid for the Near Detector. However, my studies for the Hydrogen analysis shown that the $50\,\mu s$ time cut was not enough to avoid the electronic overload effect due the muon passage, it is discussed in section 3.2.8.1.

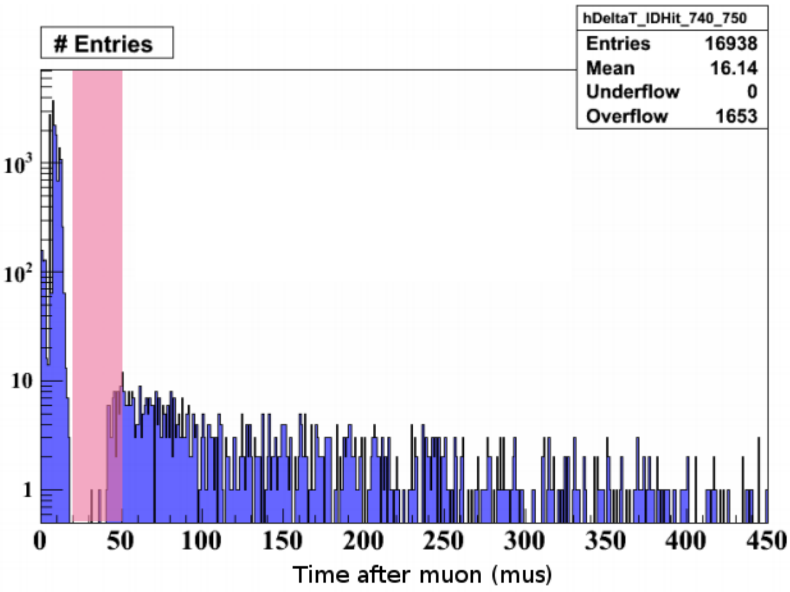


FIGURE 3.9: Number of events following a muon passage in the detector, here for muon depositing > 220 MeV in the ID. The red box show the time range of the possible trigger inefficiency.

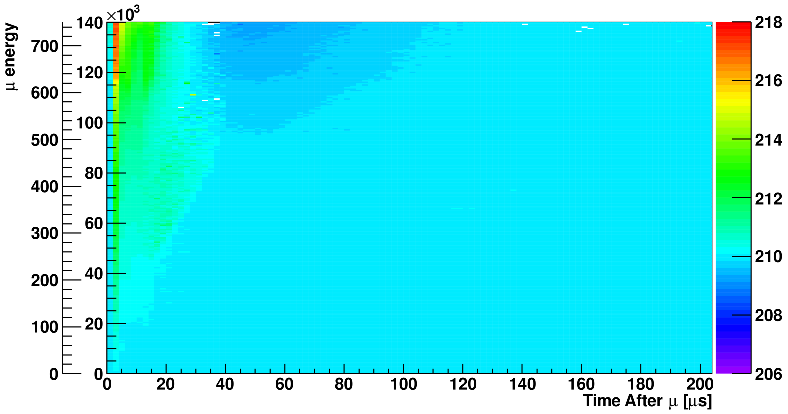


FIGURE 3.10: After Muon Flag map. The Y axes show the muon energy in PE (right axis) and a rough estimation of this energy in MeV (left axis) computed as $E_\mu = PE_\mu \times f_{MeV}$. The color scale shows the mean baseline value for a given (muon energy, time after the muon) position. We did not see any structures after $200\mu s$.

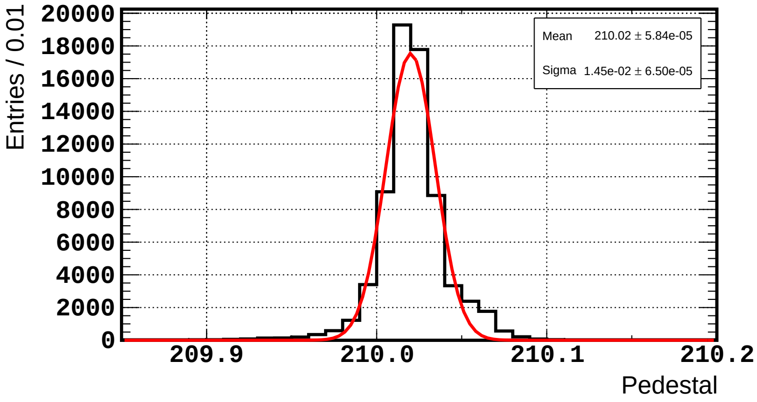


FIGURE 3.11: Baseline (as called pedestal) value from the events in FIG. 3.10 with a $\Delta t_{AfterMuon} > 400\mu s$. A fit with a Gaussian was performed, giving a baseline range for the unbiased events of 210.020 ± 0.044 (3σ).

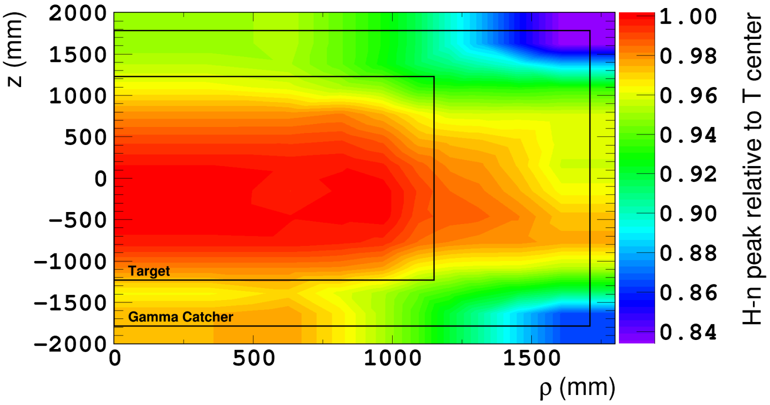


FIGURE 3.12: Uniformity map for data with After Muon Flag cut used. The color axis shows the ratio of the Hydrogen capture peak position relative to the one at the detector's center.

3.2.5 Stability calibration

The stability calibration, f_s^{data} was introduced to correct for the time dependence of the gain and detector response in data, i.e. the time dependence of the reconstructed number of PE, PE^{data}. The stability is monitored using the spallation neutron captures on Gadolinium and on Hydrogen, as well as the α decay of ^{212}Po, collected by using the Bi-Po coincidence signal. The ^{212}Po is produced by the ^{212}Bi β decay, which can be use as a prompt signal. The energy of the ^{212}Po α decays

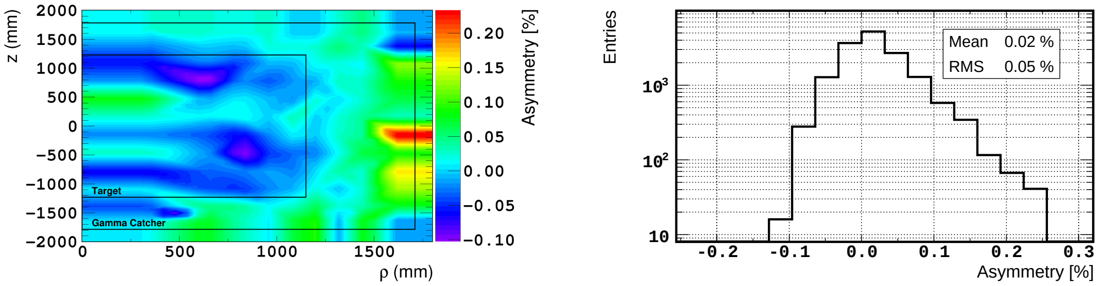


FIGURE 3.13: *Left*: 2D figure of the asymmetry between After Muon Uniformity map and the classic one, the color scale show the asymmetry in percent. *Right*: Projection of the left plot in a one dimension figure.

is 8.8 MeV, however, due to quenching effect, the visible energy is about 1 MeV, which can be use as a delayed signal. The the ^{212}Po half life is 443 ns allowing to select both signals in a short time windows. The reference time of this calibration is the day when the Double Chooz "candle" run was taken, so day 398 since the beginning of data taking. We express this reference time as $t_{ref} = 398$ days.

The DC-III (Gd-n) stability correction is applied as a function of time and energy as described by the following equation:

$$\begin{aligned} f^{data}(E_{vis}^{0,data}, t) &= 1.0 + \delta\alpha(t) \times c_{\alpha\,time\,dep}(E_{vis}^{0,data}) \\ &-c_{stability}(t \text{ days} - t_{ref} \text{ days}) \end{aligned} \tag{3.12}$$

Where $\delta\alpha(t)$ is the time variation of the mean gain of all channels, it is defined as $\alpha = n/PE^{data}$ where n is the Poisson corrected number of hit PMTs from EQ. 3.7. $\delta\alpha(t)$ has been measured using the spallation neutron capture on Hydrogen sample and is shown in FIG. 3.14. However, $\delta\alpha(t)$ has an energy dependence, which is larger at low energy, due to gain non-linearity associated with the bias in the baseline estimation. To correct this energy dependence a scaling factor $c_{\alpha\,time\,dep}(E_{vis}^{0,data})$ is applied. This energy dependence was measured using three samples at different energies: ^{212}Po α decay and spallation neutron captures on Hydrogen and Gadolinium. The scaling factor was determined to minimize the time variation of the peak energy for all samples. It was extract from the fit of the left figure in FIG. 3.14, giving:

$$c_{\alpha\,time\,dep}(E_{vis}^{0,data}) = 0.7866 - 0.07101 \times E_{vis}^{0,data} \tag{3.13}$$

The gain stability calibration is complementary to the linearized PE calibration and corrects effects from electronics, including the single PE inefficiency, the remaining gain non-linearity and so on.

The second term in Eq. 3.12 was implemented to correct the remaining time variation of the detector response, due to the scintillator light yield and the readout response after application of the gain calibration. This time variation of the detector response was measured using the peak energy of the spallation neutron captures on Hydrogen, distributed in the whole Inner Detector (Target and γ-catcher volumes). A fit was performed to measure the slope of this variation over time, giving the following result:

$$c_{stability} = 8.24 \times 10^{-6}\ \text{day}^{-1} \tag{3.14}$$

The Fig. 3.15 shows the stability of the peak energies of the ^{212}Po α decays and of spallation neutron captures on Hydrogen and Gadolinium, before and after the application of the stability correction. The Hydrogen capture peak is the most stable as the time variation was extracted using spallation neutron captures on Hydrogen.

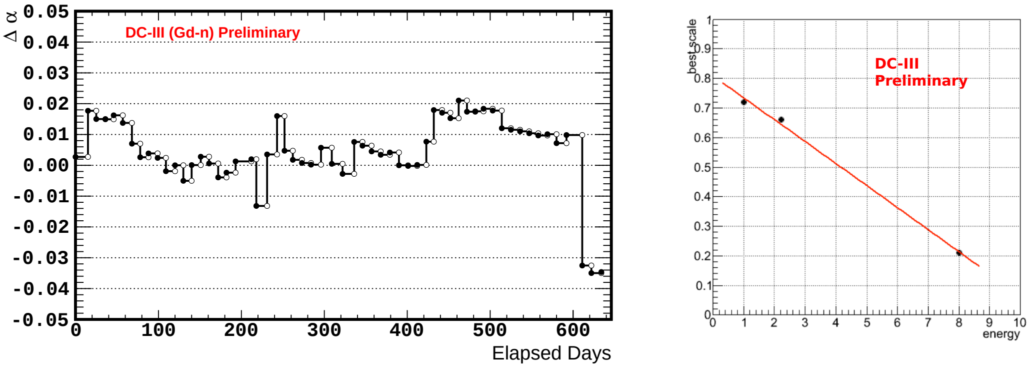


Figure 3.14: *Left*: Time variation of mean gain of all channels ($\delta\alpha(t)$). *Right*: Best scaling factor as a function of the energy. The black dots show the best scaling minimizing the time variation of, from the left to the right, the peak energy for the BiPo sample, the spallation neutron captures on Hydrogen sample, and the spallation neutron captures on Gadolinium sample. The red line shows the fit performed on the data points.

For the stability uncertainty, the RMS of the stability distribution of ^{212}Po α decays and of Spallation Neutrons capturing on Gd from Fig. 3.15 were used. These RMS were linearly interpolated and weighted by the positron energy spectrum from MC. It led to the following Stability uncertainty:

$$\delta_{Stability} = 0.50\% \tag{3.15}$$

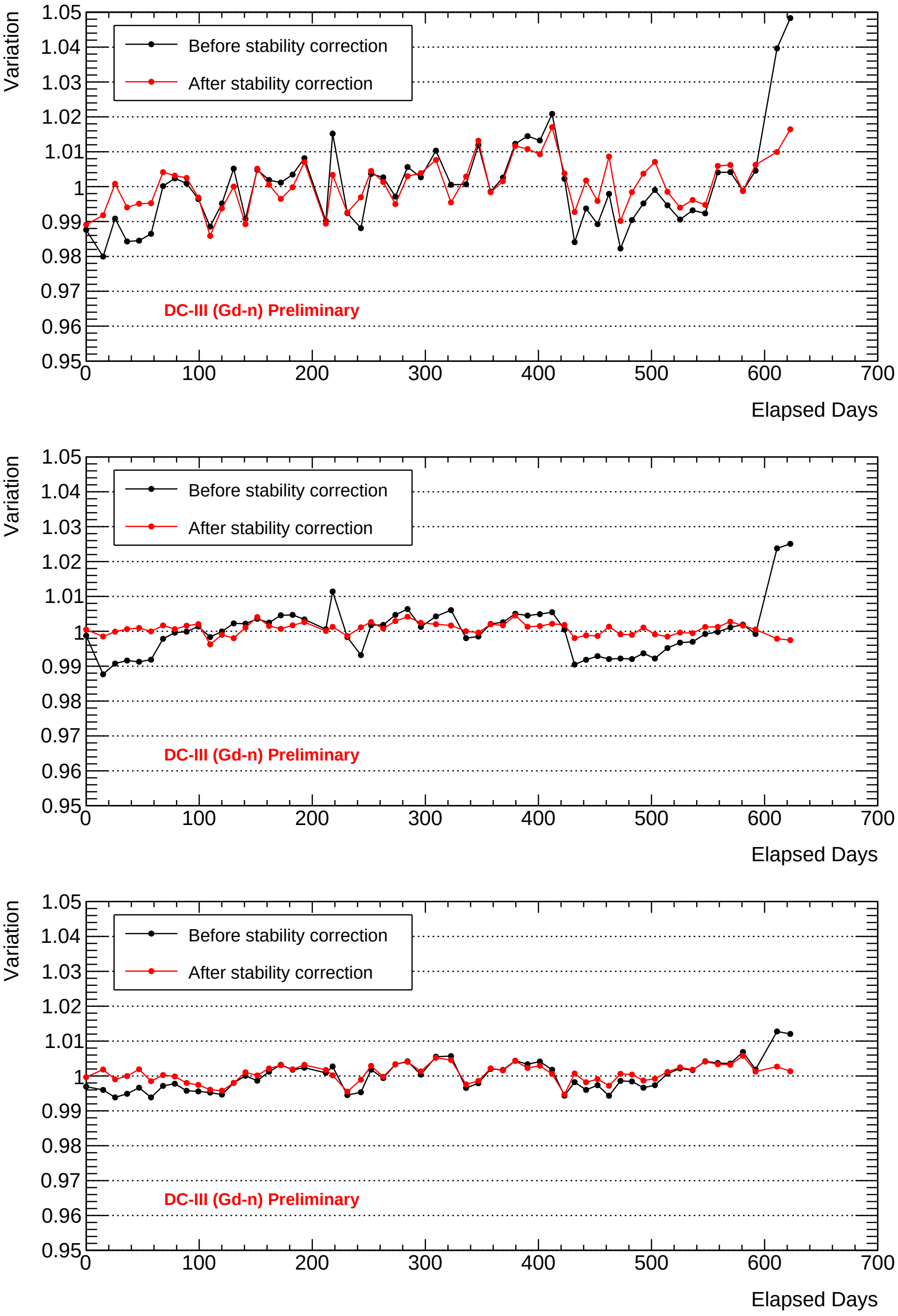


FIGURE 3.15: Stability of peak energy of ^{212}Po α decays (top) and spallation neutron captures on Hydrogen (center) and Gadolinium (bottom).

3.2.6 Non-linearity

The Monte-Carlo 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 (Eq. 3.3), only QNL is applied. The non-linearity calibration of the DC energy reconstruction is then express as:

$$f_{nl}^{MC} = f_{QNL}^{MC} \tag{3.16}$$

The LNL is included in the energy model used in the θ_{13} fit (see section 3.2.9).

3.2.6.1 Charge non-linearity (QNL)

A systematic bias in the modeling of the readout system of Double Chooz detectors, as well as inside the charge integration algorithm RecoPulse (cf. sec. 3.1.1), causes different charge non-linearity between data and MC. A part of this discrepancy is corrected by the Linearized PE calibration, however, an energy non-linearity calibration is needed to correct the residual discrepancy between data and MC. I was in charge of this correction.

The charge non-linearity correction is determined by the $data/MC$ ratio for the Gadolinium neutron capture peak and the Hydrogen neutron capture peak from the calibration runs with a ^{252}Cf neutron source deployed on the DCZAxis calibration system (cf. sec. 2.3). This ratio are plotted in function of the true energy of the neutron capture peak and a linear fit is performed between the Hydrogen point and the Gadolinium one, as shown in Fig. 3.16. Tab. 3.3 shows the mean values and the spreadings of the distributions of the fitted slopes and intercepts for different set of calibration runs.

We took as QNL correction the results from the ZA-II (center) runs[2], since they shown the higher scattering. The intercept was corrected to have a null correction at the Hydrogen peak position. In addition, as we cannot probe the charge non-linearity at high energy we decided to apply a limit at 10 MeV for the correction. Therefore, the correction applied is the following:

$$\begin{aligned} f_{QNL}(E_{vis}^0) &= 0.0023 \times E_{vis}^0\,[MeV] + 0.9949 & (E_{vis}^0 < 10\,MeV) \\ &= 1.0179 & (E_{vis}^0 \geq 10\,MeV) \end{aligned} \tag{3.17}$$

We compute the remaining charge non-linearity after application of this correction. We decide to compute it for the whole 2nd calibration ZAxis runs. The

---

[2] This stands for the calibration runs taken during the second calibration campaign, with source deployed in the center of the detector.

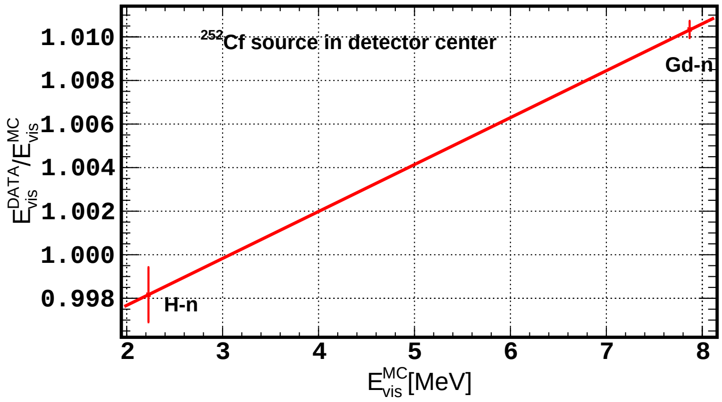


FIGURE 3.16: $E_{vis}^{DATA}/E_{vis}^{MC}$ for ^{252}Cf calibration candle run taken in detector center during the second calibration campaign. A linear fit with a slope was performed between the H-n point and the Gd-n one.

results of this analysis is shown in FIG. 3.17. The remaining charge non-linearity is then:

$$f_{QNL}(E_{vis}^{0}) = -0.0001 \times E_{vis}^{0}\,[MeV] + 1.004 \tag{3.18}$$

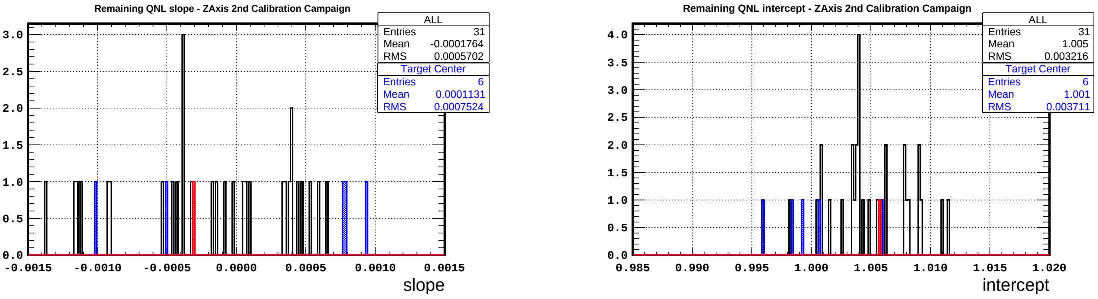


FIGURE 3.17: Results of the QNL fit after application of the QNL correction. *Left*: Fit slope and *right*: Fit intercept

For the QNL uncertainty we decided to take the maximum deviation and then took the RMS of the remaining QNL intercept distribution:

$$\delta_{QNL} = 0.4\% \tag{3.19}$$

Sample	*slope* (%/*MeV*)	σ(*slope*) (%/*MeV*)	*intercept* (%)	σ(*intercept*) (%)
ZA-II (center)	0.23	0.03	98.97	0.16
ZA-II (all)	0.21	0.01	99.20	0.07
ZA-I (center)	0.19	0.01	98.70	0.08
ZA-I (all)	0.24	0.01	98.49	0.10
average	0.22	0.03	98.84	0.16

TABLE 3.3: Results of the QNL fits. ZA means ZAxis and the number is the calibration campaign number. ZA-x (all) consists in all the ^{252}Cf ZAxis calibration runs taken during the calibration campaign x, whereas ZA-x (center) consists in only the ^{252}Cf ZAxis calibration runs where the source was deployed in the center of the detector. The average values are computed using all ZAxis runs.

3.2.6.2 Light non-linearity (LNL)

In DC-III (Gd-n), a discrepancy of the energy scale between data and MC remains after QNL calibration. As show in FIG. 3.18, this discrepancy shows dependence on the single γ energy and not on the total visible energy, shown in FIG. 3.19. Then, the cause of this discrepancy is not the charge reconstruction but comes from the scintillator modeling (quenching and Cerenkov), which has a particle dependent effect on energy. In the MC simulation the scintillator's light yield is expressed by the Birks' law:

$$\frac{dL}{dx} = L_0 \frac{\frac{dE}{dx}}{1 + k_B \frac{dE}{dx}} \tag{3.20}$$

where L_0 is a constant light yield, here sets to be 8152 photon/MeV, and k_B is the Birks' constant which represents the quenching effect, here sets to be $k_B = 0.202$. The Birks' law is an empirical formula which is valid for organic scintillator [184]

The correction of the light non-linearity has been determined thanks to several MC simulations of γ and of neutron calibration sources. These MC simulations were generated with several different combinations of Birk's constant k_B (representing the quenching effect) and of scintillator's light yield. The variation of the scintillator's light yield changes the ratio of scintillation light to Cerenkov light, which can be a potential of data-MC discrepancy. From these simulations, four (k_B, LY) combinations were found to have a good agreement with data, as shown in FIG. 3.18. Some positron MC simulations were generated with these four combinations of parameters. The ratios between the visible energy of the positron

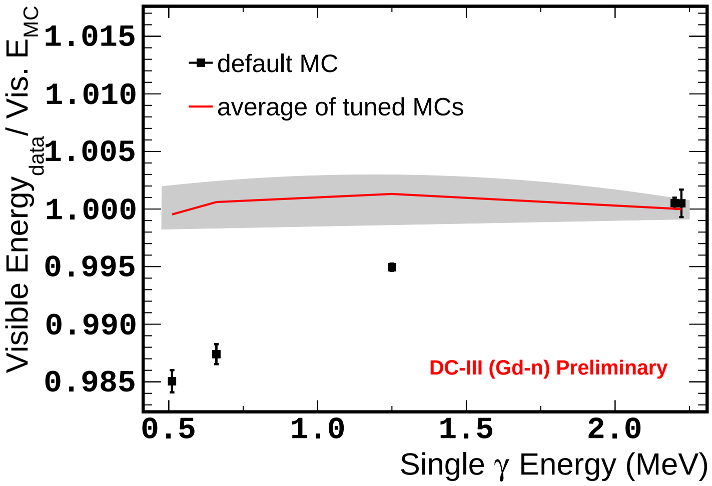


FIGURE 3.18: Remaining discrepancy of the energy scale between the data and MC after QNL correction is applied. Horizontal axis is the averaged single γ energy. Black dots show the ratio of data to the MC with default scintillator parameters and red line shows that the average values from the different combinations of Birk's constant k_B and light yield of liquid scintillator which give a good agreement with the data.

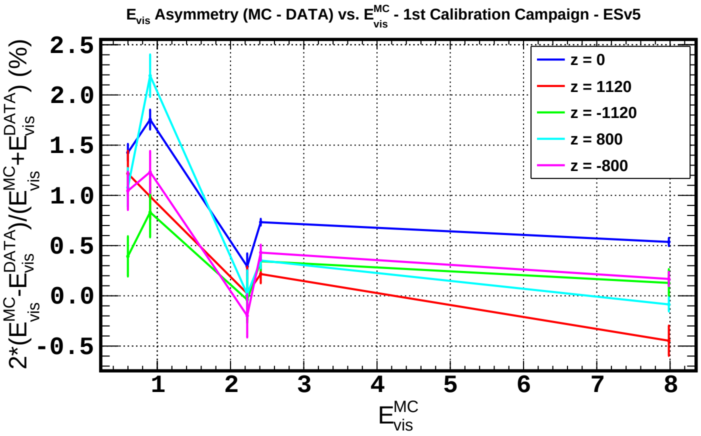


FIGURE 3.19: Remaining discrepancy of the energy scale between the data and MC after QNL correction is applied. The different lines show the data-MC asymmetry at different position in the Target.

MC and the one generated with the default values in the $\overline{\nu}_e$ MC were fit by a $1/x$-model.

The source of the bias remains indistinguishable between the modeling of Cerenkov or quenching effects, as they have correlations. However, the application of the best-fit (with the $1/x$ functions) allows to fully correct for this bias. The fit is applied independently on Target without the γ-catcher, as both volumes do not share the same scintillator liquids. The best-fit for the Target volume was computed as follows:

$$f_{LNL} = \frac{a_{LNL}}{E_{vis}^{0,MC}} + b_{LNL} \tag{3.21}$$

where $a_{LNL} = -0.027 \pm 0.0062$ MeV and $b_{LNL} = 1.008 \pm 0.0026$.

This LNL correction is not included in the visible energy computation (cf. Eq. 3.3). Instead, this correction is used in the rate+shape fit as part of the energy scale model (cf. section 3.2.9) and is applied to $\overline{\nu}_e$ MC. The fit parameters errors are taken into account considering their correlations.

The LNL comes from scintillation parameters not well simulated. When this was fully understood, it would have caused a significant delay in reproducing all MC simulations ($\overline{\nu}_e$ and calibrations). Instead, the Double Chooz collaboration decided to use the computed correction in the Final Fit. The LNL correction only addresses to discrepancies as observed in calibration data and "translated" to positrons. This correction does not correct these discrepancies entirely, it is specific to positron. For other particles a different LNL correction has to be computed.

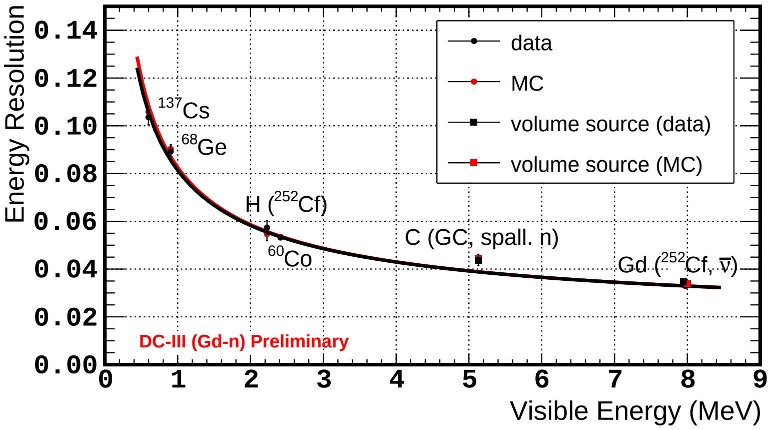
3.2.7 Validation of the energy reconstruction

In order to demonstrate the consistency of our energy reconstruction, I studied the variation of the Uniformity uncertainty and of the mean MC-data asymmetry in the different stages of the visible energy reconstruction. The results of these studies are shown in the following table:

ESv5 configuration	MC-data asymmetry
Without Stability and QNL corrections	$-1.23 \pm 0.29\%$
Without QNL correction	$-1.38 \pm 0.35\%$
Without Stability correction	$+0.08 \pm 0.30\%$
Default	$-0.06 \pm 0.36\%$

Table 3.4: Results of the Uniformity uncertainty study. The visible energy was computed with and without Stability and QNL correction and we took the mean and the RMS of the data-MC asymmetry on H-capture peak for each configuration.

Results showed that the spread of the data-MC asymmetry over the whole Target volume does not have any variations depending on the Stability and the QNL corrections. The RMS deviation is below 0.07%. However, this study shows the major improvement coming from the QNL correction. QNL correction reduces the bias (i.e. the mean) of the data-MC asymmetry to a value consistent to 0%.



data:
$a = 0.0773 \pm 0.0025$
$b = 0.0182 \pm 0.0014$
$c = 0.0174 \pm 0.0107$

MC:
$a = 0.0770 \pm 0.0018$
$b = 0.0183 \pm 0.0011$
$c = 0.0235 \pm 0.0061$

FIGURE 3.20: Energy resolution of Double Chooz. The fits were performed with $\frac{\sigma}{E_{vis}} = \sqrt{\frac{a^2}{E_{vis}} + b^2 + \frac{c^2}{E_{vis}^2}}$ where a is the stochastic term of the energy resolution, b the constant term and c is linked to the electronics noise. The fitted value of these parameters is shown on the right.

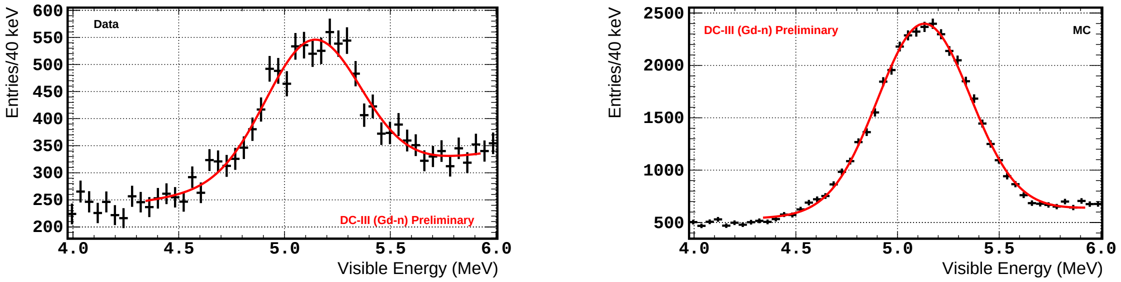


FIGURE 3.21: Fast neutrons capture peak on Carbon for data (*left*) and MC (*right*). Red line shows the fit function, a convolution of a Gaussian and of a degree 1 polynomial. The fit results are for the mean 5.135 ± 0.010 MeV for data and 5.130 ± 0.001 MeV for MC.

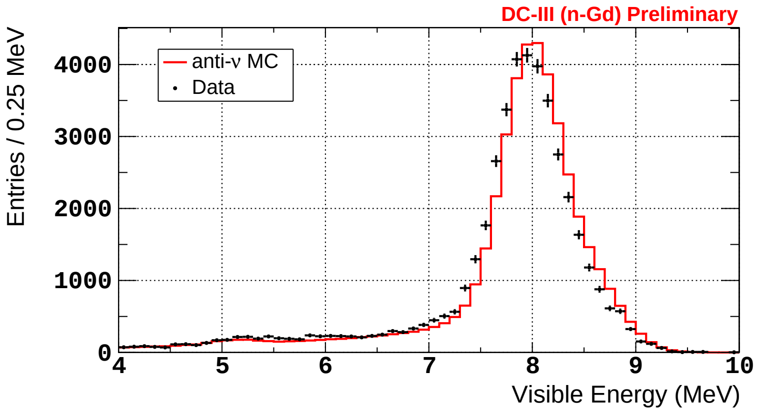


FIGURE 3.22: Gadolinium neutron capture peak for IBD from Gd-III selection

FIG. 3.20 shows the resolution of our energy reconstruction as a function of the visible energy. It shows a very good agreement between data and MC in the Target center with calibration data, in the Target volume with $\overline{\nu}_e$ IBD and in the γ-catcher volume with fast neutron capture on ^{12}C. The Carbon capture peak of fast neutrons was studied to illustrate the validity of our energy reconstruction, as we calibrated it with Hydrogen and Gadolinium peak. I was in charge of this study.

We used the fast neutron capture events with an energy around 5 MeV and fit it with a convolution of a Gaussian function and of a degree 1 polynomial function. The result of its analysis is shown in FIG. 3.21 and illustrates a very good agreement between data and MC with a deviation lower than 1%.

As a final illustration of the validity of our energy reconstruction, FIG. 3.22 shows the delayed neutron-Gadolinium peak of our Gadolinium IBD signal. It agrees within 1σ of our energy scale uncertainty.

Then, thanks to our uniformity and stability calibration, $\overline{\nu}_e$ interactions behave in the full target region over a period of $\sim$ 600 days as if interacting in the target center during the same day.

3.2.8 Energy Reconstruction for DC-III (H-n)

The energy reconstruction described in the previous sections was done for the third Double Chooz analysis using IBD neutron capture on Gadolinium (DC-III (Gd-n)). For the Double Chooz analysis using IBD neutron capture on Hydrogen, it was mandatory to recompute some uncertainties, as the previous uncertainties were computed with only the Target as fiducial volume, whereas the fiducial volume for Hydrogen analysis consists in Target and γ-catcher. Both Uniformity and

Stability uncertainties needed to be recomputed. For the same reasons, it was also mandatory to recompute the Light Non-Linearity (LNL).

3.2.8.1 Uniformity uncertainty

In order to compute the Uniformity uncertainty for the DC-III (H-n) selection, we could not use the same method than for the Gadolinium selection. In the Hydrogen selection we needed to control the γ-catcher in addition to the Target. Then, the Uniformity uncertainty computation from the Gadolinium capture visible energy peak was not possible. We had to use the Hydrogen capture visible energy peak. In order to have two independent samples, the generation of a new correction map was mandatory. Therefore, in order to have two independent samples representative for the full data taking period, we generate two run lists taking one out of two runs from the full run list of the analysis.

One of the run list has been used to generate the new Uniformity correction map, with the same method detailled in section 3.2.4. It was tagged as *ESv6* map. The other run list has been used to compute the Uniformity uncertainty with the *ESv6* Uniformity correction map. We did not use the new Uniformity correction map to compute the visible energy in the $\overline{\nu}_e$ selection. Instead, we used the same Uniformity correction map than for the DC-III (Gd-n), presented in section 3.2.4. The only purpose of the *ESv6* Uniformity correction map, was to allow the computation of the Uniformity uncertainty at Hydrogen neutron capture energy.

As we selected one run out of two for each run list, the possible variations of the detector response over time are taken into account. Therefore, both run list should be equivalent. For MC Uniformity correction map, we kept the one use in the Gadolinium analysis. However, we used another MC sample to compute the Uniformity uncertainty, in order to ensure an independent sample.

The resulted Uniformity correction map for data is shown in Fig. 3.23. The compatibility between this new correction map and the one use for the $\overline{\nu}_e$ selection has been done by computing the asymmetry between the *ESv6* map and the *ESv5* map. The resulted asymmetry is shown in Fig. 3.24. With a mean asymmetry and an asymmetry deviation below 0.1% (−0.07% and 0.08% respectively) the *ESv6* and *ESv5* data map are almost identical. This demonstrates the validity of the method used to compute the uncertainty for the DC-III (H-n) analysis.

The uncertainty has been computed using this *ESv6* map with the same method than for the Gadolinium analysis. Fig. 3.25 shows the resulted asymmetry map and the asymmetry projection. As for the Gadolinium selection, the RMS of the projection is used as uncertainty, so 0.25%, this represent an improvement by a factor ~ 6 compared to the previous Hydrogen analysis [55]. However, the high bias, −0.84%, was not expected. We expected a value close to the MC-

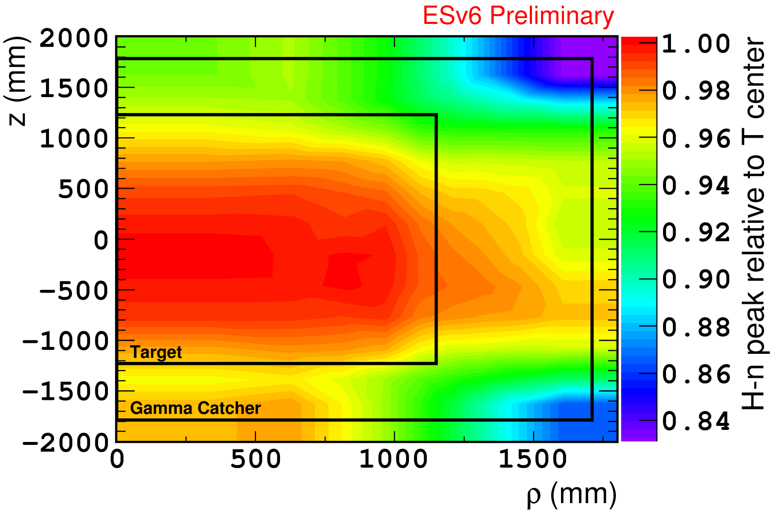


FIGURE 3.23: Uniformity correction map for data generated with half statistics for Hydrogen analysis.

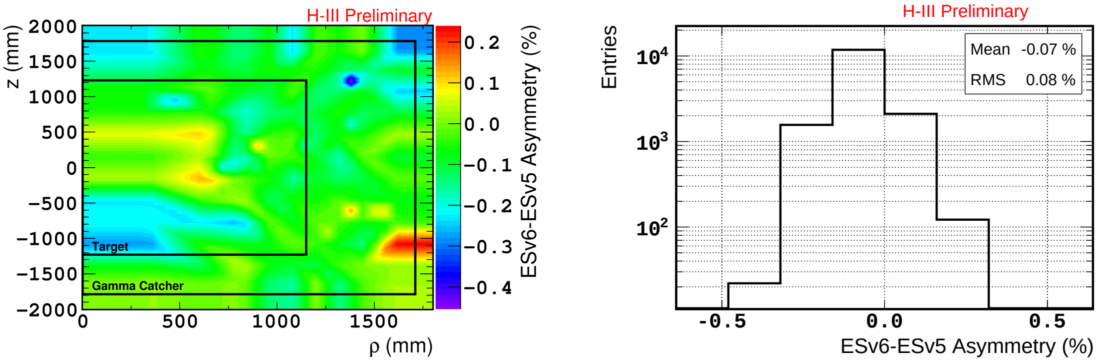


FIGURE 3.24: *Left*: 2D figure of the asymmetry between *ESv5* data map and *ESv6* data map, the color scale show the asymmetry in percent. *Right*: Projection of the left plot in a one dimension figure.

Data asymmetry for the Hydrogen neutron capture peak with the ^{252}Cf reference calibration run. However, this ^{252}Cf MC-Data asymmetry was about +0.18%, which indicates a 4σ deviation from the expected value, whereas the deviation from the expected value was 1σ in the Gadolinium selection. We investigated in order to figure out why.

In order to probe a potential issue in the energy reconstruction method, we reproduced the uncertainty computation with different configurations of the energy scale. We also compute the uncertainty at the Gadolinium neutron capture energy

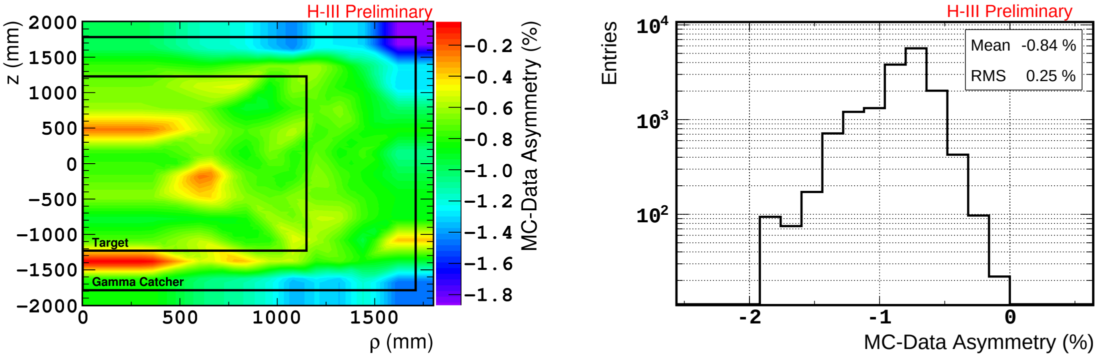


FIGURE 3.25: *Left*: 2D figure of the asymmetry between E_{vis}^{MC} map and E_{vis}^{DATA} map at the Hydrogen neutron-capture energy, the color scale show the asymmetry in percent. *Right*: Projection of the left plot in a one dimension figure.

for comparison. The results of these investigations are shown in the TAB. 3.5.

ESv6 configuration	H-n asymmetry	Gd-n asymmetry
Default	$-0.84 \pm 0.25\%$	$-0.16 \pm 0.29\%$
Without Stability correction	$-0.52 \pm 0.24\%$	$-0.02 \pm 0.31\%$
Without QNL	$-0.84 \pm 0.25\%$	$-1.47 \pm 0.28\%$
Without Stability correction and QNL	$-0.53 \pm 0.23\%$	$-1.32 \pm 0.30\%$
Expected asymmetry from ^{252}Cf	$+0.18 \pm 0.13\%$	$+0.28 \pm 0.04\%$

TABLE 3.5: Results from investigation in order to understand the high bias in the *ESv6* asymmetry plot. Several energy reconstruction configuration were tested. The expected asymmetry was measured using the ^{252}Cf candle run.

Results for the Gadolinium selection asymmetry can be compared to the ones from TAB. 3.4. Accounting a small variation ($< 0.1\%$) the Gadolinium selection results with *ESv6* are consistent with the ones with *ESv5*. The Gadolinium selection mean asymmetry, in the default configuration, has the deviation from the ^{252}Cf value of $\sim 1.5\sigma$. For the Hydrogen selection asymmetry, as it was designed to, QNL has no impact. However, the stability correction increases the absolute MC-data asymmetry by $\sim 0.3\%$, this was not understood.

As the issue explanation for the mean asymmetry could not be found from the different configurations of the energy reconstruction, we investigated on the possibility of a bias due to the muon passage in the detector. In section 3.2.4.2 we showed that the bias due to the muon high energy depositions should be negligible at the Gadolinium energy. This demonstrated that the future Double Chooz Near

Detector analysis could safely use an energy reconstruction built with a dead time of $50\mu s$ after a muon. However, even if most effects due to the muon energy depositions are removed by the $50\mu s$ cut, some low energy after pulse effect can still be present around $50\mu s$ as it is possible to see in FIG. 3.8.

It is possible that some PMT baselines are still biased due to the high energy deposition, without being flagged by the After Muon map (see section 3.2.4.2). This can be due either to a small deviation of the baseline, below the threshold of the After Muon analysis, or to a distortion of the baseline such that the mean of the baseline is not affected. If it happens in coincidence with a Hydrogen neutron capture, these baseline distortions could cause a small shift of the Hydrogen capture peak. Thanks to the higher energy released by the Gadolinium neutron-captures, the impact of such baseline distortions should have a negligible impact on the Gadolinium capture peak. This could be an explanation of the high negative MC-data asymmetry for the Hydrogen selection, whereas the MC-data asymmetry is close to 0% for the Gadolinium selection.

In order to test it, we selected a new Fast Neutron sample, using a time cut of $100\mu s$ after a muon. This cut value was chosen to be totally safe from potential bias due to the muon energy depositions. It was not optimized. From this new Fast Neutron sample we generated a new Uniformity correction map for data. This new Uniformity correction map is shown in FIG. 3.26.

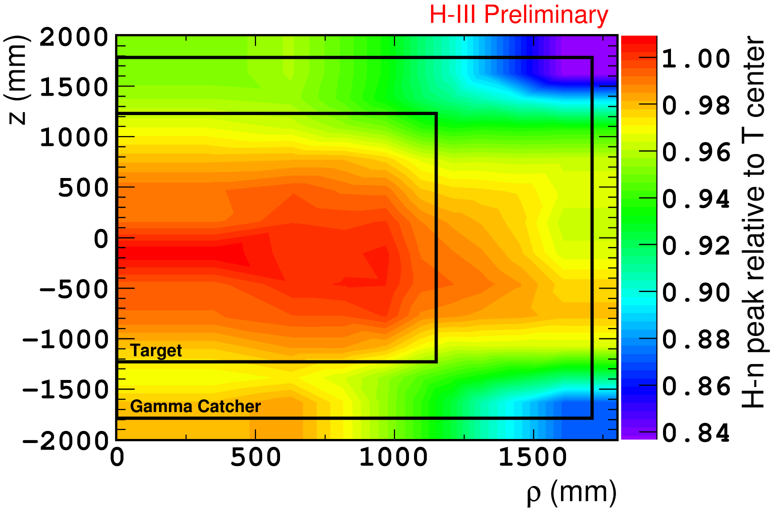


FIGURE 3.26: Uniformity correction map for data generated with half statistics. Fast Neutron sample used for this map was selected using a time cut of $100\mu s$ after a muon instead of the default one ($50\mu s$).

We applied this new Uniformity correction map in order to compute the MC-data asymmetry. The resulted asymmetry map and asymmetry projection are shown in FIG. 3.27. The mean asymmetry dropped to +0.14% using this new Uniformity correction map. The expected value from ^{252}Cf H-n was also shifted by 0.05%. The modification in the expected value can be explain by the fact that the neutron capture does not really occur in the exact center of the detector. It can occur as far as 1 meter from the source position (due to the spatial cut). This induces a small dependency to the Uniformity correction map. Due to this shift of the ^{252}Cf neutron capture peaks, we needed to compute a new QNL correction. Which was estimated, with the method described in section 3.2.6.1, to be:

$$f_{QNL}^{100\mu s}(E_{vis}^{0}) = 0.0014 \times E_{vis}^{0}\,[MeV] + 0.9968 \tag{3.22}$$

The asymmetries computed with the new $100\mu s$ correction map are summarized in the TAB. 3.6. The high shift of the ^{252}Cf Gd-n asymmetry is mainly due to the modification of the QNL correction. With the previous QNL correction, the ^{252}Cf Gd-n asymmetry was +0.33% and the Fast Neutron Gd-n asymmetry was +0.34%. For the first time in the Double Chooz analysis, the MC-data asymmetry from fast neutron and from ^{252}Cf were consistent within 0.01%.

As the $100\mu s$ cut reduces our statistical power by 80%, we needed to confirm that this *good* result was not due to a statistical bias. We compute the asymmetry between the new $100\mu s$ correction map and the *ESv6* correction map. The asymmetry map is shown in FIG. 3.28. This asymmetry map shows some structure well defined and this not consistent with a statistical fluctuation. We interpreted these structures are likely to be due to the effects from the muon high energy deposition. We considered that this result confirm that the $50\mu s$ cut after a muon is not enough for the fast neutron analysis. However, $100\mu s$ should not be the best cut value as it removed ~ 80% of our statistical power. For the future analysis, the Double Chooz collaboration will have to fine tune this cut, in order to both be safe from the bias due the muon high energy depositions and save the statistical power.

Sample	H-n asymmetry	Gd-n asymmetry
Fast Neutron sample	$+0.14 \pm 0.24\%$	$-0.12 \pm 0.31\%$
^{252}Cf neutron	$+0.23 \pm 0.13\%$	$-0.12 \pm 0.04\%$

TABLE 3.6: MC-data asymmetry results with the Fast Neutron sample and the ^{252}Cf neutron sample

This new map was not applied for the DC-III (H-n) IBD selection. We think that the bias due the muon can be taken into account by the LNL correction, and we think that it was the case for the DC-III (Gd-n) analysis. The LNL

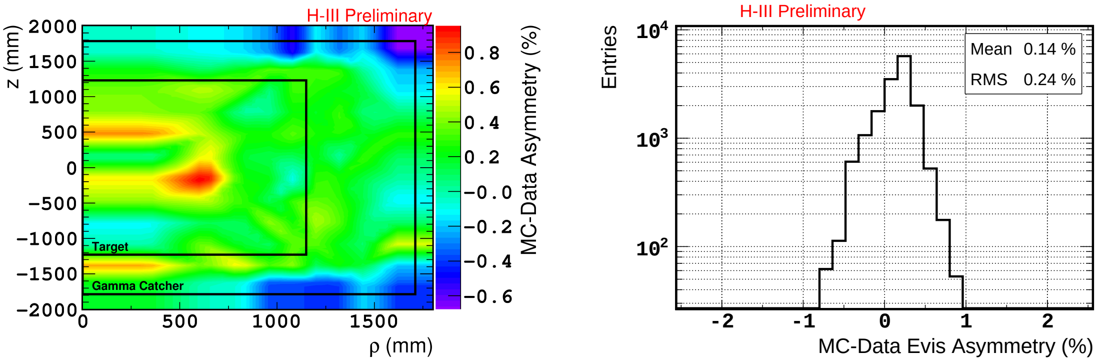


FIGURE 3.27: *Left*: 2D figure of the asymmetry between E_{vis}^{MC} map and E_{vis}^{DATA} map at the Hydrogen neutron-capture energy, the color scale show the asymmetry in percent. *Right*: Projection of the left plot in a one dimension figure. Fast Neutron sample used was selected using a time cut of $100\mu s$ after a muon instead of the default one $(50\mu s)$.

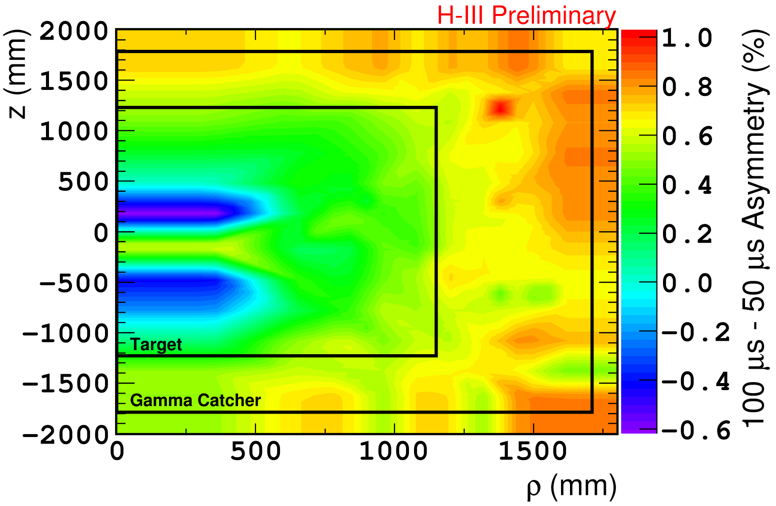


FIGURE 3.28: 2D figure of the asymmetry between $100\mu s$ correction data map and *ESv6* correction data map, the color scale show the asymmetry in percent.

correction computes the remaining discrepancies between the data and the MC and then could correct more than the true "Light" non-linearity. The DC-III (H-n) LNL study is described in the following section. Even if not applied for the DC-III (H-n) analysis, this study will have to be applied for the future analysis of Double Chooz. We demonstrated the existence of this issue due to the high energy deposition following cosmic muons. The light non-linearity does not aim to the correction of this issue and is itself a temporary correction, it should be included directly in the Monte Carlo simulation for the future analysis.

As the RMS of all H-n asymmetry projections was stable, we took the RMS of the asymmetry projection computed with the *ESv6* correction map as DC-III (H-n) Uniformity uncertainty:

$$\delta^{H}_{Uniformity} = 0.25\% \tag{3.23}$$

This represents an improvement by $\sim$ 81% with respect to the Uniformity uncertainty used in the previous Hydrogen analysis [55].

3.2.8.2 LNL correction

As the ν-target LNL was known and well computed, in order to estimate the γ-catcher light non-linearity, we decided to compute it using comparison between the ν-target and the γ-catcher. In order to perform it, we studied the single γ visible energy data-MC ratio from the calibration radioactive sources (see section 3.2.1.1). The single γ visible energy is the mean energy per γ, it is defined in Eq. 3.24. The mean number of γ for each source was defined using [146] and MC simulations, results can be find in Tab. 3.7.

$$E_{vis}^{Singe\,\gamma} = \frac{E_{vis}}{Number\,of\,\gamma} \tag{3.24}$$

Source	E_{vis} (MeV)	$< Number\,of\,\gamma >$	$E_{vis}^{Singe\,\gamma}$ (MeV)
^{137}Cs	0.661	1	0.661
^{68}Ge	1.022	2	0.511
$^{2(2}Cf$ (H-n)	2.223	1	2.223
^{60}Co	2.505	2	1.252
^{252}Cf (Gd-n)	7.937	3.6	2.205

Table 3.7: Mean number of γ and mean visible energy per γ per source. The mean number of γ for Gd-n capture was estimated with several MC simulations.

The results of the analysis is shown in Fig. 3.29. The results in the γ-catcher showed a different shape than in the one in the ν-target. At low energy the DATA/MC ratio is higher than 1 for the γ-catcher and lower than 1 for the ν-target. We found a variation of the shape in the γ-catcher: for run taken close to the γ-catcher border the shape seems flatter than for run taken in the middle of the γ-catcher. In order to understand it, we applied harder cuts in the selection: selecting only events occurring at 10, 15 and 20 cm from the mean reconstructed source position. This study showed that the two γ-catcher shapes were closer when the spatial cut was harder.

A fit was performed to estimate the difference between the ν-target and the γ-catcher LNL, as shown in 3.29. The fit showed that the ν-target and the γ-catcher LNLs were almost the inverse of each other. We interpreted this as the following: the energy depositions from sources located close to the γ-catcher volume do not shown the γ-catcher LNL, instead they show a mixing of the ν-target and γ-catcher LNLs. As the ν-target and γ-catcher LNLs seems to be the inverse of the other, the resulted shape from the LNLs mixing is almost flat.

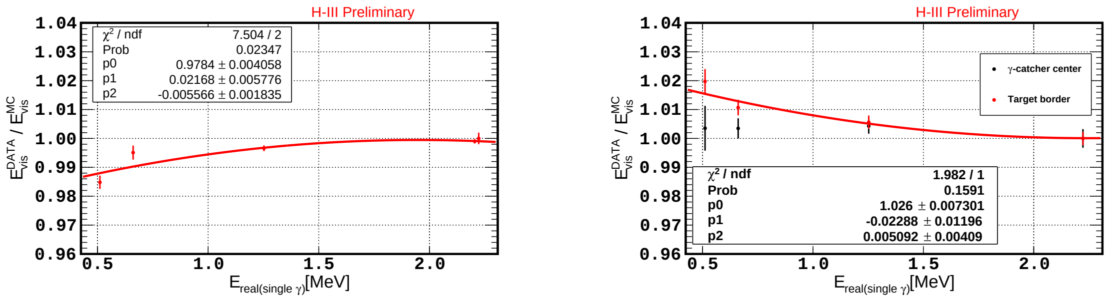


FIGURE 3.29: $\frac{E_{vis}^{DATA}}{E_{vis}^{MC}}$ as a function of the single γ visible energy for calibration sources. Calibration runs were taken during the second calibration campaign. Both were fitted with a 2-degrees polynomial function. *Left*: Singe γ E_{vis} for Target volume. *Right*: Singe γ E_{vis} for γ-catcher volume. The red dots show the $E_{vis}^{Singe\,\gamma}$ in the middle of the γ-catcher, black dots show the $E_{vis}^{Singe\,\gamma}$ at the border between Target and γ-catcher volume.

In order to integrate this LNL in the rate+shape fit (see section 3.2.9), two solutions were suggested: measuring the two LNLs in order to apply both of them in the fit, or increasing the LNL input error and let the fit figures out the effective LNL. The first solution was not practical, as the LNL has a position dependence. It would have been difficult or simply impossible to integrate it in the fit. Investigations were done to study the second possibility. We computed the error due to the LNL for the ν-target and for the γ-catcher. This was done with the fit results: we weighted the prompt spectrum with an error function deriving from the fit function, and took the integral of the resulted function as error. The *error* function was assumed to be flat for energies upper than the Hydrogen neutron capture energy. It can be computed as:

$$
\begin{aligned}
LNL\,error\,(E_{vis}) &= (a_{fit} \cdot E_{vis}^2 + b_{fit} \cdot E_{vis} + c_{fit}) + \sigma_{H-n} & if\; E_{vis} < 2.223\,MeV \\
&= \sigma_{H-n\,fit} & else
\end{aligned}
\tag{3.25}
$$

where a_{fit}, b_{fit} and c_{fit} are the 2-degrees polynomial function's parameters and σ_{H-n} the error on the Hydrogen peak position. The weighted spectrums are shown in FIG. 3.30. This shows that the LNL γ-catcher error is $\sim 1.13\times$ LNL Target error. Then, if we increase the LNL error, both ν-target and γ-catcher LNLs can be covered. Several rate+shape studies were also done with the DC-III (Gd-n) selection to ensure the fit stability with this method [151, 152] and found that we reach better results if we increase the LNL error from the Gadolinium analysis by a factor 2, leading to the following values for the $1/x$ LNL model parameters (cf. EQ.3.21):

$$a_{LNL} = 0.0000 \pm 2 \times (a_{LNL}^{Gd} + \sigma_{a_{LNL}^{Gd}}) = 0.0000 \pm 0.0664 \tag{3.26}$$

$$b_{LNL} = 1.0000 \pm 2 \times ((1 - b_{LNL}^{Gd}) + \sigma_{b_{LNL}^{Gd}}) = 1.0000 \pm 0.0219 \tag{3.27}$$

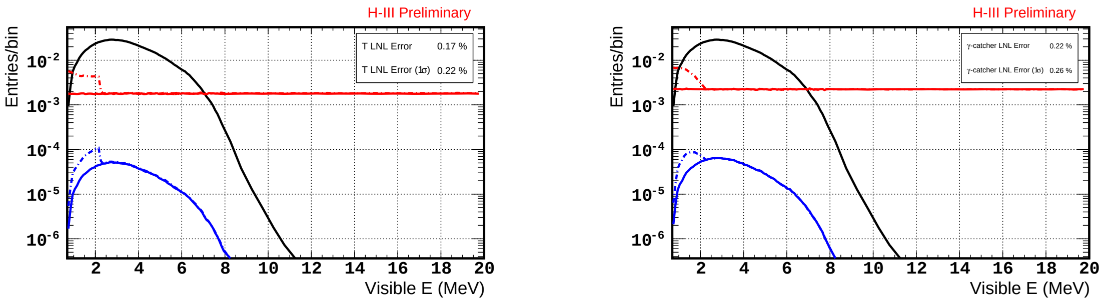


FIGURE 3.30: IBD positron spectrum weighted by the LNL error function (EQ. 3.25). The black line is the IBD prompt spectrum, red line shows the LNL error function and blue line shows the weighted spectrum. Dash lines show the LNL error function and the weighted spectrum when taking into account the error coming from energy peak positions. *Left*: In the ν-target volume: $Integral = 0.17\%$ and 0.22% when taking into account error from energy peak positions. *Right*: In the γ-catcher volume: $Integral = 0.22\%$ and 0.26% when taking into account error from energy peak positions.

3.2.9 Energy model in final fit

Energy reconstruction uncertainties are taking into account in the rate+shape fit via an energy model. Since we apply different corrections to the energy, these corrections have to be taken into account in the energy mode. Therefore, we express it as follows:

$$E_{vis}^{MC} \rightarrow E_{vis}^{MC} \cdot LNL \cdot stability/uniformity \cdot QNL \tag{3.28}$$

where, LNL address for the light non-linearity, $stability/uniformity$ address for the possible biases from Uniformity and Stability and QNL address for the charge non-linearity.

As explained in section 3.2.6.2, the LNL is described by a $1/x$ model. We already expressed LNL as follows (Eq. 3.21):

$$LNL = \frac{a_{LNL}}{E_{vis}^{MC}} + b_{LNL} \tag{3.29}$$

for notation purpose a_{LNL} is expressed as a in the energy model. Note that for the Hydrogen analysis, the result of this formula is 1, only the uncertainties on a_{LNL} and b_{LNL} are considered. We express the $stability/uniformity$ part as follow:

$$stability/uniformity = b_{st/u} \tag{3.30}$$

From the studies done for the Uniformity and Stability corrections, estimated that there is no bias due to Stability or Uniformity and decided to set $b_{st/u} = 1$. The uncertainty linked to $b_{st/u}$ is the quadratic sum of the Stability and Uniformity uncertainties: $\sigma_{b_{st/u}} = \sqrt{\delta^2_{Uniformity} + \delta^2_{Stability}}$. We assumed no correlations between the Uniformity and Stability corrections.

Finally, as explained in section 3.2.6.1, the QNL correction is expressed, below 10 MeV, as the following model:

$$QNL = b_{QNL} + c_{QNL} \times E_{vis}^{MC} \tag{3.31}$$

where $b_{QNL} = 0.9949$ and $c_{QNL} = 0.0023$ MeV^{-1}. Above 10 MeV, we have: $b_{QNL} = 1.0179$ and $c_{QNL} = 0$ MeV^{-1}. For the energy model, we took the values of the remaining charge non-linearity (i.e. after application of the QNL correction):

$$b_{QNL} = 1.004 \pm 0.004 \tag{3.32}$$

$$c_{QNL} = c = -0.0001 \pm 0.0006 \tag{3.33}$$

These values are the same for both Hydrogen and Gadolinium analysis. For notation purposes we note c_{QNL} as c in the energy model.

The energy model can be then expressed as:

$$E_{vis}^{MC} \rightarrow E_{vis}^{MC} \cdot (\frac{a}{E_{vis}^{MC}} + b_{LNL}) \cdot b_{st/u} \cdot (b_{QNL} + c \cdot E_{vis}^{MC}) \tag{3.34}$$

With simple math, we can have the simplified Eq. 3.35 as energy model. This last equation is how the energy model is integrated in the rate+shape fit.

$$E_{vis}^{MC} \rightarrow a' + b' \cdot E_{vis}^{MC} + c' \cdot (E_{vis}^{MC})^2 \tag{3.35}$$

where $a' = a \cdot b_{st/u} \cdot b_{QNL}$ (MeV), $b' = a \cdot b_{st/u} \cdot c + b_{LNL} \cdot b_{st/u} \cdot b_{QNL}$ (dimensionless) and $c' = b_{LNL} \cdot b_{st/u} \cdot c$ (MeV^{-1}).

The final fit itself is detailed in section 5.5.

3.3 Summary

The Double Chooz energy reconstruction has been described in details. The energy reconstruction is a critical part in the understanding of our detector. It is divided in several parts: the pulse reconstruction which allows to extract the charge from the waveform; the linearized PE calibration which converts the charge into PE, correcting for the gain non-linearity; the absolute energy scale in order to provide the conversion factor from PE to MeV; and the detector response uniformity and stability correction which allow to consider the detector response as the stable over time and position.

The QNL correction allows to correct the MC in order to reproduce the charge non-linearity in data. And LNL takes into account the last MC-data asymmetries and propagate them for final fit.

The new methods developed for the DC-III (H-n) analysis are going to be propagated and used for the next Double Chooz analysis, for both Hydrogen and Gadolinium analysis. The Double Chooz collaboration will have to tune the time after muon cut for fast neutron analysis. This is already ongoing.

However, RecoZoR development can lead to a totally new energy reconstruction, this new pulse reconstruction method is described in the chapter 4.

Chapter 4

A new pulse reconstruction method: RecoZoR

As already introduced in the previous chapters, Double Chooz is looking for IBD interactions inside the Inner Detector. The main analysis of Double Chooz is a rate+shape analysis and therefore, the energy reconstruction is critical for the success of this analysis.

The energy reconstruction of Double Chooz has been already presented and described in chapter 3. It has been shown in section 3.2.2, that the energy reconstruction suffers from a gain non-linearity at low charge, where the calibration is harder, as shown in Fig. 3.4 (revisited below). This non-linearity is due to the limited sampling of the waveform digitizer, which can bias the channel baseline estimation [116]. This non-linearity is the dominant systematic in the Double Chooz energy reconstruction process. In addition, this non-linearity is uncorrelated across detectors. This implies that there will be no cancellation between ND and FD in the second phase of the experiment.

The main region of interest for the θ_{13} measurement is between $[0.5, 3]$ MeV. A 3 MeV energy deposition correspond to ~ 570 PE (see section 3.2.3). If it occurs in the center of the detector, $1-2$ PEs are expected in each PMT. This PE per PMT should not change a lot if the energy deposition occurs in another point of the ν-target. Therefore, this main region of interest is where the impact of the non-linearity is the largest.

The non-linearity is also linked to the pulse reconstruction method,the output signal of the PMT is recorded by dedicated flash-ADC (see section 2.2.2) digitizing waveforms of 256 ns. The waveform can be interpreted as a “film” of what occurs in the PMT. Fig. 3.2, from chapter 3 (revisited below), is an example of such waveform.

Since the waveforms are the sum of the single-PE pulses in the output signal of the PMT, it should be possible to decompose the waveforms into single-PE

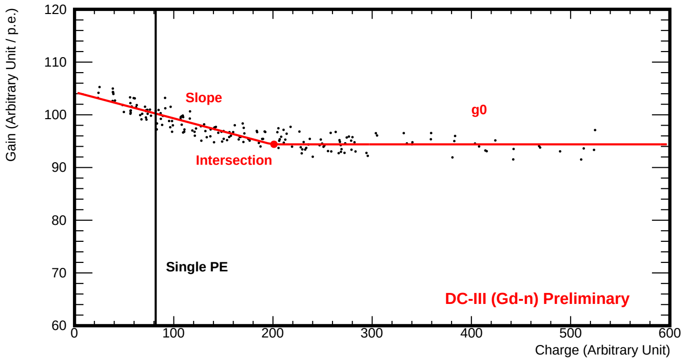


Figure 3.4 (revisited): Gain vs. integrated charge for a typical readout channel. Black dots show the measurements of the gain, using calibration data from the IDLI system with different light injection positions and light intensities (i.e. different charges). Red line shows the gain function obtained from a fit, which is applied to convert integrated charge to number of photoelectrons. The gain is determined with the three parameters from Eq. 3.5. On this figure, g_0, *slope* and *intersection* refer, respectively, to the g_i^m, the l_i^m, and the c_i^m of Eq. 3.5. The black line shows the position of one photoelectron.

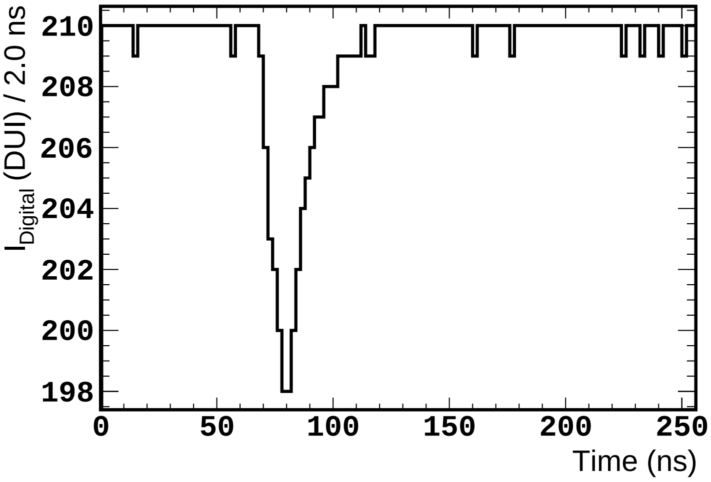


Figure 3.2 (revisited): Double Chooz digitalized waveform from one PMT. The negative pulse is the intensity of the electron cascade inside the PMT due to a PE. Time binning is 2 ns. DUI stands for *digital units of intensity* and is equivalent to ADC counts.

pulses. This would provide additional time information, and information on the fluctuations of the PMT response. However, in the present DC energy reconstruction this is not assumed. In fact, as already detailed in section 3.1.1, the pulse reconstruction method consists in the integration of the pulses in the waveforms without any assumption on the shape of the pulse. This absence of assumption provide a major robustness to the pulse reconstruction, since in the early stages of DC the response and its non-linearity were not fully understood. However, this underestimates the information in the waveforms, the "films" of the detector.

In this chapter, a totally new pulse reconstruction method will be described, using the knowledge of single-PE shape (as measured from data) in order to decompose the waveform in single-PE pulses, hence maximizing the information extracted from the waveform. This should allow to go beyond the main limitations of the current method of DC. Indeed, it can be expected that this new method will allow to significantly improve the non-linearity of Double Chooz and the time resolution of the experiment. This new algorithm is called "RecoZoR" and the preliminary results are promising. If these preliminary results, presented here, are validated, this new pulse reconstruction method, started as a part of this thesis, could become a major improvement of the full reconstruction of DC, since it will impact the ID vertex reconstruction (see section 2.7.1) and the overall energy reconstruction. Other experiments, using similar electronics than Double Chooz, could also take advantage of this method in order to improve their energy reconstruction.

The idea behind RecoZoR is to reverse the read-out simulation of Double Chooz (called RoSS), where the full knowledge on the PMT response is used, in order to extract the number of PE and the starting time of each PE in the waveform (hence the name: RoSS → SSoR → ZoR). Therefore, we will start this chapter by a detailed presentation of RoSS.

4.1 RoSS: a read-out system simulation

As already introduced in section 2.6.3, RoSS is the read-out system simulation used in the Double Chooz framework. RoSS performs the simulation of the PMT response by generating each photoelectron (PE) produced by the PMT. A PE generated by a PMT produces a charge deposition which is registered in the PMT output signal as a negative pulse, as shown in Fig. 3.2 (revisited above). RoSS simulates these waveforms for each PMT for a given energy deposition inside the detector.

Geant4 simulates the light path to the PMT, and therefore provides a simulated number of PE generated by the PMT and their incoming times. RoSS takes this information as an input for the waveform simulation, generating each PE separately.

In a perfect case, the peak of a PE starts at the PE arrival time to the PMT and is normalized at 1 PE. However, the amplitude of the peak depends on the secondary electrons cascade's amplitude which has to be taken into account. RoSS generates Gaussian variations around the peak theoretical time and its normalization. Due to these variations, some PEs can be removed from the waveform or have a normalization equivalent to several PEs. These simulated variations take into account the specific gain of each PMTs.

Dedicated analysis [145] showed that the shape of a PE peak inside a waveform can be described as a Landau distribution:

$$\varphi_{spe}(t;\mu,\sigma) = exp\left[it\mu - |ct|(1 + \frac{2i}{\pi}\log(|t|))\right] \tag{4.1}$$

where μ and σ are the Landau distribution parameters. The parameter μ can also be called the most probable value of the landau distribution. The shape of a PE peak is then expressed as:

$$I_{digital}(t) = \varphi^{n}_{spe}(t;\mu_{spe},\sigma_{spe}) \times Q^{f}_{PE} \tag{4.2}$$

where μ_{spe} and σ_{spe} are the Landau distribution parameters and Q^{f}_{PE} is the charge normalization factor, in arbitrary units, computed from the variated normalization factor, in PE. $\varphi^{n}_{spe}(t;\mu_{spe},\sigma_{spe})$ is a normalized Landau, computed as $\varphi^{n}_{spe}(t;\mu_{spe},\sigma_{spe}) = \frac{\varphi_{spe}(t;\mu_{spe},\sigma_{spe})}{\sigma_{spe}}$,. In RoSS simulation, μ_{spe} and σ_{spe} were tuned to be:

$$\mu_{spe} = 8 + t^{f}_{PE} \tag{4.3}$$

$$\sigma_{spe} = 2.5 \tag{4.4}$$

where t^{f}_{PE} is the starting time of the peak, after taking into account the random variations. Each PE peak for a given PMT is simulated by RoSS and summed into the waveform. This can generate complex waveforms with several PE peaks superimposed.

RoSS also has a standalone mode, called *injector* mode, which allows to generate a random number of PE in a given PMT without Geant4 simulation input. For most of the examples show in this chapter, RoSS was used in the *injector* mode.

4.2 RecoZoR

RecoZoR aims to decompose the waveform into single-PE pulses. Several possibilities to decompose the waveform have been investigated, like Continous Wavelet

Transform, but did not provide useful results. The method presented here tries to fit the waveform with one or more single-PE pulse shape, using the single-PE pulse shape information from RoSS.

Note that RecoZoR is not finished yet, the version presented here is the current development version, results are still preliminary.

RecoZoR algorithm works in two main steps. The first one, the *peak finding*, consists in a scan of the waveform in order to identify the different peaks within the waveform. It is used as an input for the second step, which consists in the extraction of the single-PE pulses from the waveform.

4.2.1 First step of the algorithm: Peak finding

As a first step, the number of separate peaks within the waveform need to be determined. This is performed with a simple method of *pulse counting*.

The *peak counting* requires to compute the waveform baseline in order to find the peaks. The baseline refers to the mean intensity of the PMT output signal, i.e. the flat line around 210 ADC counts (or DUI) in Fig. 3.2. As RecoZoR is made to be independent from the current pulse reconstruction, we did not use the baseline computed by it. Then, a very rough baseline estimation is performed. We performed a flat fit on the first 16 ns of the waveform. The 16 ns time window was chosen from the floating baseline estimation method (see section 3.1.1). This rough baseline estimation gives results which are consistent with the ones of the current estimation method for simulated waveforms.

Fig. 4.1 shows typical waveforms with 1 and 2 PEs. Using the rough baseline estimation, we can try to find the position of each peak within the waveform. However, in Fig. 4.1, fluctuations of 1 ADC count can be observed around the baseline. These fluctuations are due to the inherent electronic noise. In order to avoid the selection of such fluctuations as PE, only peaks with an amplitude > 1 ADC are selected.

We developed a method allowing to find each peak, and an estimation of its starting time (i.e. here, the time when the peak starts to be > 1 ADC) and of its ending time (i.e. here, the time when peak starts to be < 1 ADC or when another peak starts). This method allows to scan the waveform. Fig. 4.2 shows an example of application. Several tunnings were performed to avoid the selection of random fluctuations as peaks.

If the time resolution was ideal, this simple method could maybe be used in order to compute the number of PE. Unfortunately, PEs can enter in the PMT in a short time window (within few ns). Therefore, as the time resolution is ~ 2 ns, these "simultaneous" PEs will contribute to the same pulse within the waveform. Fig. 4.3 shows a typical example of such case, where only one peak is found whereas two PEs were simulated. This "pill-up" is expected to be dominant when

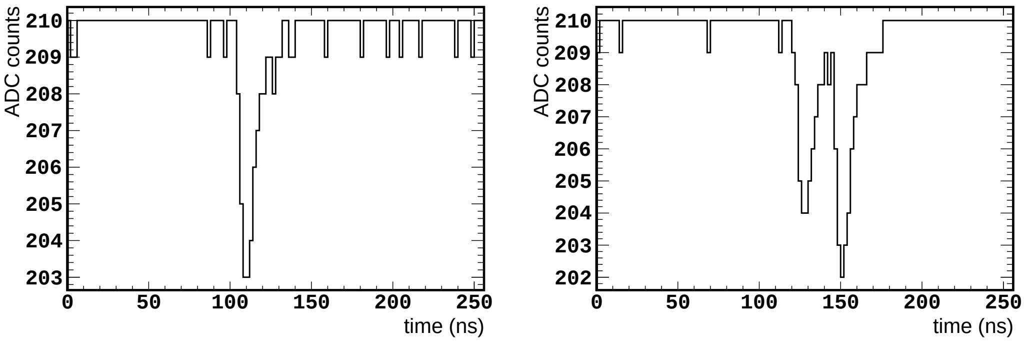


FIGURE 4.1: Waveform generated by RoSS with 1 PE (*left*) and 2 PEs (*right*).

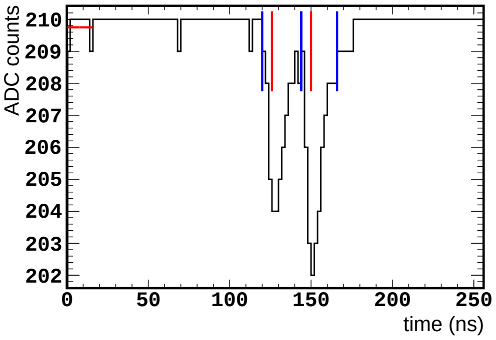


FIGURE 4.2: Waveform generated by RoSS with 2 PEs. Horizontal red line on the left is the fit performed to estimate the PMT baseline. The vertical blue lines are the limits of the found negative pulses, the vertical red lines are the pulse maximum positions.

the number of PE is high. Therefore, RecoZoR performance are expected to be lower than the one of the shape independent method in this case.

4.2.2 Second step of the algorithm: Single-PE pulse extraction

As already presented above, RoSS simulates single-PE using normalized Landau with fixed parameters. This normalized Landau is also used as the single-PE pulse shape by RecoZoR. A fit can be performed on each pulse of the waveform, found by the *peak finding* step, with the following formula:

$$I(t) = \varphi^{n}_{fit}(t; \mu_{fit}, \sigma_{fit}) \times Q_{fit} + B \tag{4.5}$$

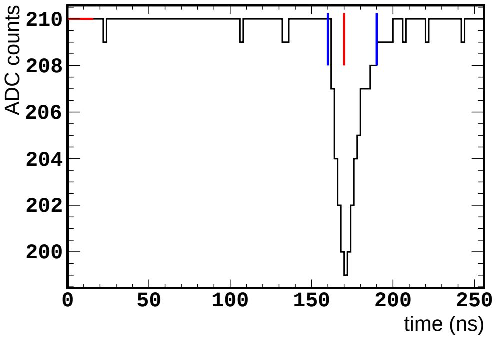


FIGURE 4.3: Waveform generated by RoSS with 2 PEs. Both PEs contribute to the same negative pulse within the waveform. The horizontal red line on the left is the fit performed to estimate the PMT baseline. The vertical blue lines are the limits of the found pulse, the vertical red line is the pulse maximum position.

where Q_{fit}, μ_{fit} and σ_{fit} are the fit parameters related to the parameters used by RoSS in EQ. 4.2, and B is the baseline of the waveform.

The position of the peak maximum and the rough baseline from the *peak finding* are used as starting values for the parameters μ_{fit} and B, respectively. σ_{fit} is initialized at 2.5 as in RoSS. We used the starting and the ending times of the pulse from the *peak finding* as fit range. FIG. 4.4 shows the result of the fit on two different peaks.

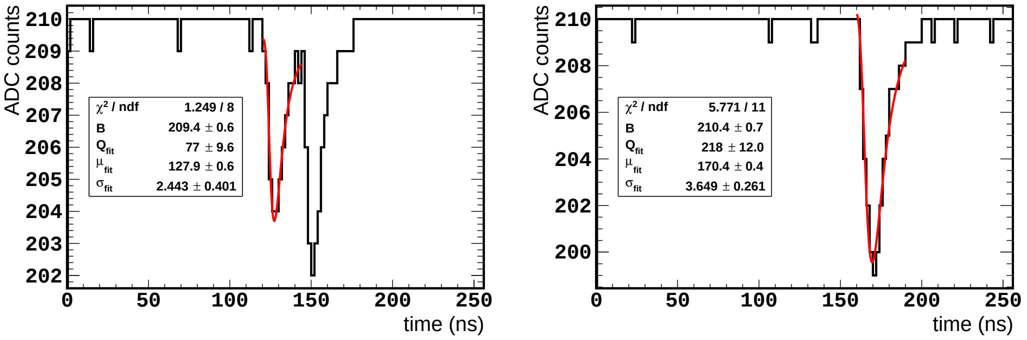


FIGURE 4.4: Waveform generated by RoSS with 2 PEs fitted with EQ. 4.5. $\chi^2/n.d.f.$ values are not meaningful as the fits were performed without taking into account any errors.

Our studies showed that the landau fit is distorted when there is a “pill-up”. Such distortion can be seen in the right figure of FIG. 4.4: σ_{fit} is far away from its definition of 2.5. The detection of such distortion could then allow RecoZoR to detect the presence of “pill-up” PEs in one peak. We investigated on several ways to detect such distortions, the one presented here is the only one which allows to find consistent results.

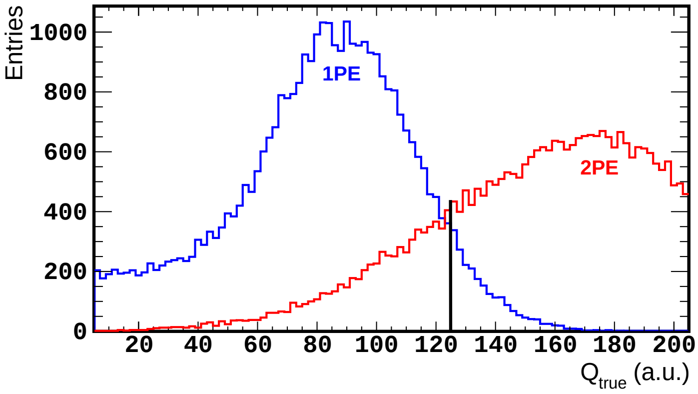


FIGURE 4.5: True charge normalization factor for waveform simulated with 1 PE (blue) and with 2 PEs (red). The black line shows the position of the 125 arbitrary units.

In order to take care of these distortions, we decided to add constraints in the fit. FIG. 4.5 shows the distribution of the charge normalization factor for one and two PEs. This distribution is expected to be the same for all PMTs and to be stable over time. From this figure, normalization charge minimum and maximum values for a single PE can be extracted. These values go from 0 to $\sim$ 170 arbitrary units. However, this distribution includes single PE with $<$ 1 ADC count amplitude, which we do not think to be able to reconstruct. Therefore, as we want only PEs with an amplitude $>$ 1 ADC count, a minimum value[1] of $\sim$ 20 a.u was used. The fact that we are not reconstructed PE with $<$ 1 ADC count could have an impact on the resolution of RecoZoR, as it is missing this PE. However, since these PE have an amplitude $<$ 1 ADC count, it is currently not possible to discriminate them from the noise.

The charge normalization factor distribution for two simulated PEs is also shown in FIG. 4.5. The integral of this distribution was normalized to the one of the 1 simulated PE distribution. The intersection between the 2 simulated PEs and the 1 simulated PE distributions is $\sim$ 125 a.u., therefore we decide to use it a the upper limit on the charge normalization factor, Q_{fit}, in the fit. For σ, we

---

[1] This minimum value should be $\sim$ 14 a.u. if the ADC count was not an integer. Here we took the value for 1.5 ADC count

arbitrary set a 10% variation window around its initial value. Tighter and wider variation windows were tested, the 10% window showed the best results.

Therefore, the fit parameters were tuned to be:

- $\mu_{fit} \in$ peak limits from the *peak finding*, starting at the position of the peak maximum
- $\sigma_{fit} \in [2.2, 2.8]$, starting at 2.5
- $Q_{fit} \in [20, 125]$ a.u., starting at 100 a.u.

Using these parameter limits, the distortion of the fit should be detected by looking at the parameter values after the fit. Fits distorted by the "pill-up" effect are expected to show parameter values equal to the parameter limits. Therefore, when a fit output presents parameter values close the limits, we interpret this as an indication of the presence of an additional PE. RecoZoR can then add another single-PE shape in the fit and repeat the operation (cf. Eq. 4.6). Some technical conditions were also applied to avoid the possibility to fit a single peak with an abnormal number of single-PE shapes. Fig. 4.6 shows the same waveforms as in Fig. 4.4 when RecoZoR uses parameter limits and multiple landau fit.

$$I(t) = \varphi^{n}_{fit,0}(t; \mu_{fit,0}, \sigma_{fit,0}) \times Q_{fit,0} + \varphi^{n}_{fit,1}(t; \mu_{fit,1}, \sigma_{fit,1}) \times Q_{fit,1} + ... + B \quad (4.6)$$

This operation is repeated until the parameter values become consistent with the parameter limit, or until the number of landau distributions become abnormal (i.e. number of degree of freedom in the fit < 1).

In the case of waveform with multiple peaks, like in the left plot in Fig. 4.6, there could also be some overlap between the peaks. In this case, the fit output would be biased. In order to avoid this issue, each peak is subtracted to the waveform after the fit. Fig. 4.7 illustrates this process.

Using the information from the different fits, RecoZoR is able to fit the full waveform. This "final" fit makes it possible to add additional constraints into the fit and to improve the estimation of the baseline, as well as the one of the time and charge information of each PE. Fig. 4.8 presents two different waveforms with the fit performed on the full waveform.

4.2.3 RecoZoR's performance

From the different fits and reconstruction methods applied, RecoZoR is then able to provide several information: the number of PE and, for each PE, its time and its charge normalization factor. The results presented here were obtained using the *injector* mode of RoSS.

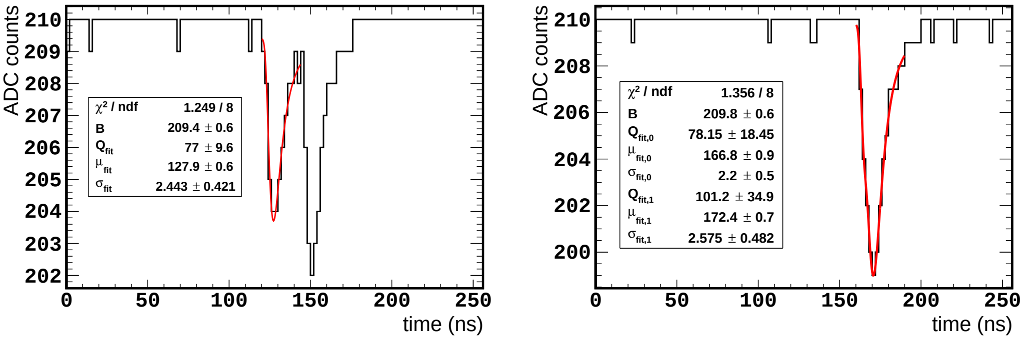


FIGURE 4.6: Waveform generated by RoSS with 2 PEs fitted with parameter limits and multiple landau fits. $\chi^2/n.d.f.$ values are not meaningful as the fits were performed without taking into account any error.

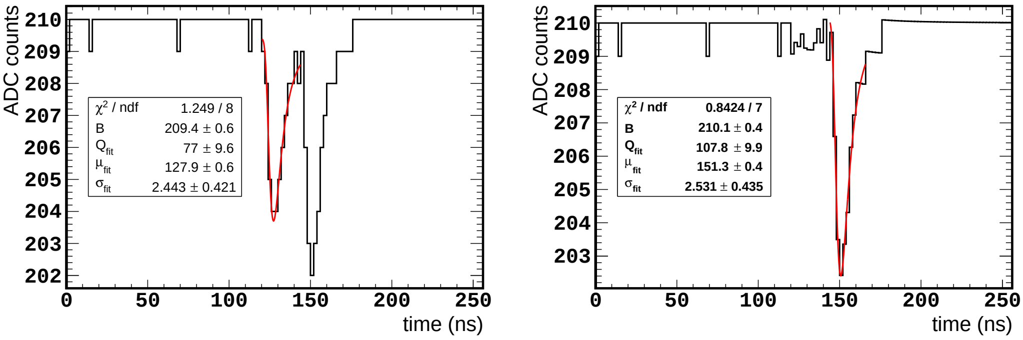


FIGURE 4.7: Separate fit of each peaks in a waveform generated with 2 PEs. The first peak (fit on *left*) was subtracted from the waveform before the fit of the second peak (*right*). $\chi^2/n.d.f.$ values are not meaningful as the fits were performed without taking into account any error. .

In this section, we will compare the results obtained with the shape independent method and with the shape dependent method. For simplicity, the algorithm used for the shape independent pulse reconstruction method is refered by its name within the Double Chooz framework: "RecoPulse". In section 3.1.1, the start time (t_{start}) computed by RecoPulse has already been introduced. RecoPulse provides a second time definition for each waveform, which is used in the document: the time of the maximum of the pulse (t_{max}).

The performance of RecoZoR on the PE counting are going to be shown using $\Delta NPE = NPE_{True} - NPE_{Reco}$ distributions, where NPE_{True} is the number of

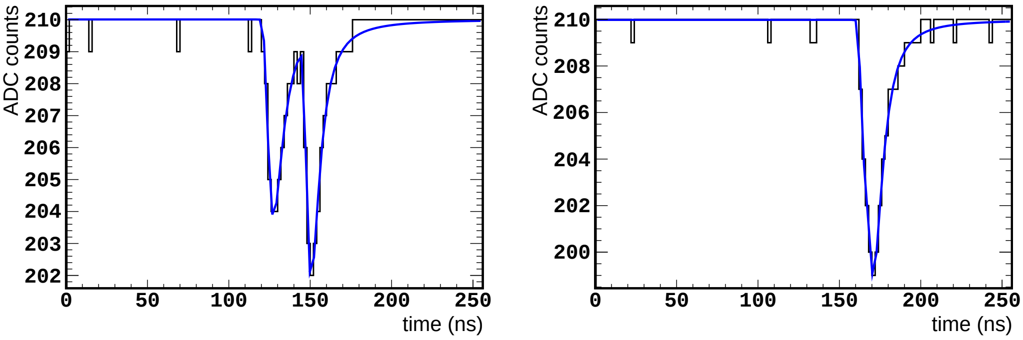


FIGURE 4.8: Waveform generated by RoSS with 2 PEs fitted with a combination of all peak's fits.

PE simulated by RoSS and NPE_{Reco} is the one reconstructed either by RecoZoR or RecoPulse. The number of PE for RecoZoR is the number of single-PE shape within the fit of the waveform. The performance on the time resolution are going to be shown using $\Delta t = t_{start\,first\,PE,True} - t_{Reco}$ distributions, where $t_{start\,first\,PE,True}$ is the simulated arrival time of the first PE, and t_{Reco} is a reconstructed time from RecoZoR or RecoPulse. RecoPulse reconstructed time definitions are t_{start} and t_{max} (see above). For RecoZoR, t_{Reco} refers to the μ of the first single-PE shape fit, since according to RoSS μ is directly linked to t_{start}. The resolution on the number of PE (on the time) is defined as the spreading of the ΔNPE (Δt) distribution.

Looking at waveforms simulated with only one PE, in FIG. 4.9, we can observe that RecoZoR and RecoPulse miss a certain amount of single PE. RecoZoR misses $\sim 5\%$ of the single PEs, whereas RecoPulse misses $\sim 6\%$ of the single PEs.

The reasons why RecoZoR misses these single PEs were investigated. FIG. 4.10 shows the true time and the normalization charge for the single PEs missed by RecoZoR. It can be seen that RecoZoR mainly misses PE with $Q_{true} < 20$ a.u. or with $t_{start,true} > 248$ ns (i.e. 256 ns -8 ns). PEs with $Q_{true} < 20$ a.u. have a < 1 ADC count amplitude, therefore RecoZoR is not able to reconstruct them. PEs with $t_{start,true} > 248$ ns are difficult to fit as their peak maximum is often outside the waveform range. In addition, the current *preliminary pulse counting* method has difficulties to identify then as a peak. The correction of this issue is ongoing. Currently, RecoZoR is able to reconstruct only $\sim 11\%$ of the PEs generated after 248 ns. Fortunately, the beginning of the peak of these PEs can be seen within the waveform, therefore issue can be flagged and reported by RecoZoR. Investigations are ongoing in order to reconstruct these PEs. These PEs, with $Q_{true} < 20$ a.u. or with $t_{start,true} > 248$ ns, will also not be considered in the following discussion.

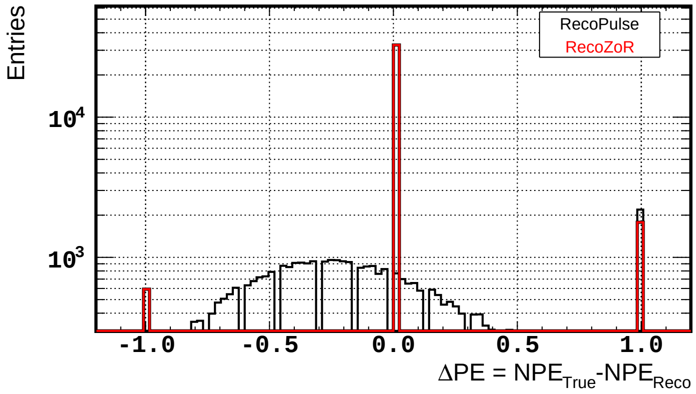


FIGURE 4.9: ΔNPE distribution for RecoZoR (red) and RecoPulse (black). Only waveform generated with 1 PE were used.

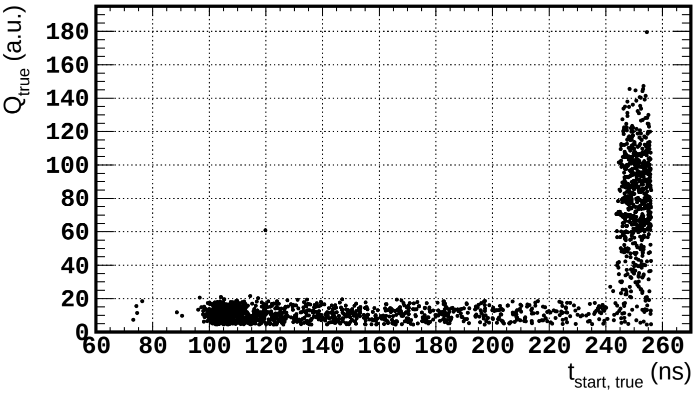


FIGURE 4.10: PE true charge versus true starting time for waveforms generated with 1 PE and not reconstructed by RecoZoR. Waveforms generated with a starting time > 256 ns are not physics and so not presented here.

Looking only at "reconstructible" PEs, the RecoZoR's resolution can be measured through the number of PE and through the PE time.

FIG. 4.11 shows the $\Delta NPE = NPE_{True} - NPE_{Reco}$ distribution and the $\Delta t = t_{start\,first\,PE,True} - t_{Reco}$ distribution for single PE waveforms reconstructed with RecoZoR and RecoPulse. We find a resolution on the number of PE of 0.135 PE for RecoZoR and 0.410 PE for RecoPulse. The resolution of RecoZoR is better than the one of RecoPulse by a factor ~ 3. The Δt figure shows that RecoPulse misses more single PE than RecoZoR (i.e. $\Delta t > 100$ ns means that $t_{Reco} \lll t_{True}$ so $t_{Reco} \sim 0$ ns). A zoom was performed on the peak for both algorithms in Fig. 4.12. We find here a time resolution of 0.369 ns for RecoZoR and of 0.801 ns (t_{max}) and 0.840 ns (t_{start}) for RecoPulse time definitions. The time resolution is clearly better with RecoZoR, with an improvement of a factor $\sim$ 2. In addition, the $\sim$ 8 ns shift of the distribution mean value is expected, since, in RoSS, the relation between the starting time of a PE and the landau μ is: $\mu = t_{start} + 8$.

These preliminary good results are a demonstration of the shape dependent method capabilities. RecoZoR is better than RecoPulse on the single PE level.

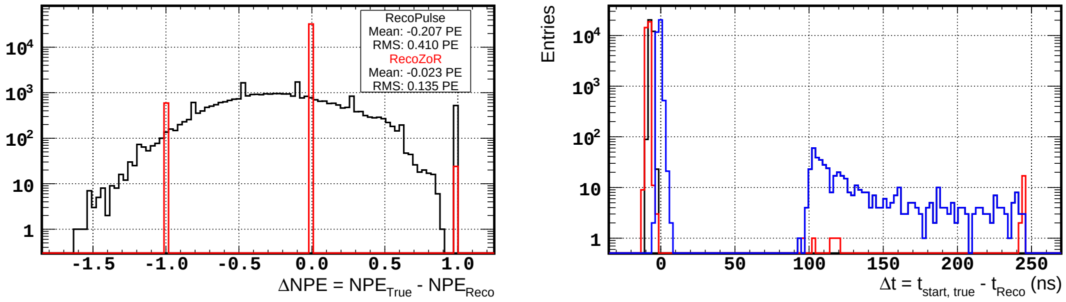


Figure 4.11: ΔNPE and Δt distributions for RecoZoR (red) and RecoPulse (black and blue). RecoPulse provides two time definitions, the time of the peak's maximum (black) and the peak's starting time (blue). RecoZoR time is the μ of the landau fit. Only waveforms generated with 1 PE and with $Q_{true} > 20$ a.u. and $t_{start,true} < 248$ ns are shown. .

Fig. 4.13 shows the RecoZoR and RecoPulse ΔNPE and Δt distributions when waveforms generated with 2 PEs are considered. On the time distribution figure, a second peak with a long tail is visible for RecoPulse start time (blue). This suggests that RecoPulse tends to give as starting time the time of the second PE. We suspect it to happen mainly in case of "pill-up". Thanks to the fact that it is a shape dependent method, RecoZoR does not present the same behavior. The time resolution for RecoZoR is 1.495 ns, whereas the resolutions for RecoPulse are biased due to this effect. When zooming on the peak, the spreading are similar for RecoZoR and RecoPulse. We find a resolution on the number of PE of 0.335 PE for RecoZoR and of 0.542 PE for RecoPulse. Therefore, RecoZoR also better than RecoPulse on the reconstructed number of PE at 2 PEs.

Fig. 4.14 shows the ΔNPE distribution as a function of the true number of PE for RecoZoR. RecoZoR misses at least one PE for the main part of the waveforms with $NPE_{True} > 4$ PEs. This is due to the "pill-up" of PE with which RecoZoR has difficulties to deal with.

Fig. 4.15 shows the resolution on the number of PE of RecoZoR and RecoPulse.

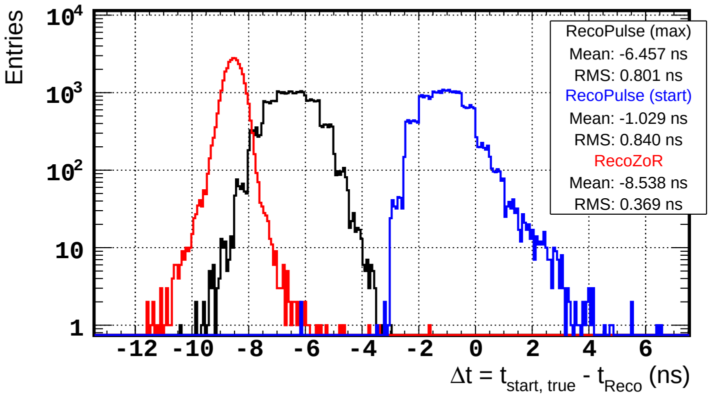


FIGURE 4.12: Δt distribution for RecoZoR and RecoPulse (black and blue). RecoPulse provides two time definitions, the time of the peak's maximum (black) and the peak's starting time (blue). RecoZoR time is the μ of the landau fit. Only waveforms generated with 1 PE and with $Q_{true} > 20$ a.u. and $t_{start,true} < 248$ ns are shown. .

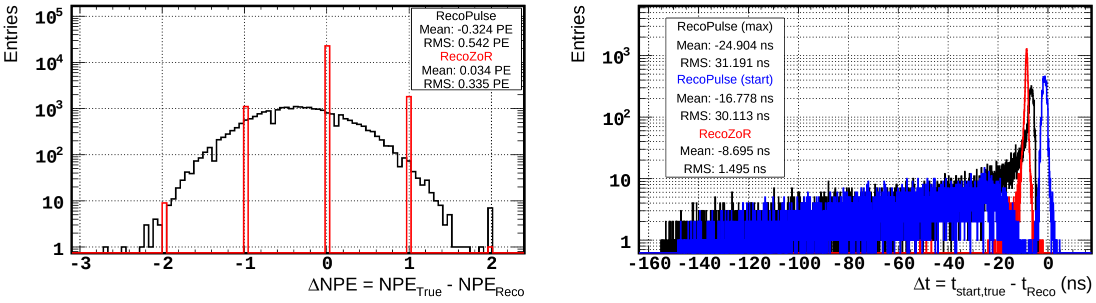


FIGURE 4.13: ΔNPE and $\Delta t_{first\,PE}$ distributions for RecoZoR (red) and RecoPulse (black and blue). RecoPulse provides two time definitions, the time of the peak's maximum (black) and the peak's starting time (blue). RecoZoR time is the μ of the first landau fit. Note that the spreadings (RMS) are similar between RecoPulse and RecoZoR if we only consider the peak. Only waveforms generated with 2 PEs and with $Q_{true} > 20$ a.u. and $t_{start,true} < 248$ ns are shown. .

It is visible that RecoZoR has better linearity and resolution than RecoPulse until $NPE_{True} = 6$ PEs. For $NPE_{True} > 6$ PEs, RecoPulse's performance are better than RecoZoR's ones. It was expected from RecoPulse to demonstrate better result for a high number of PE. This can be consider as a good result.

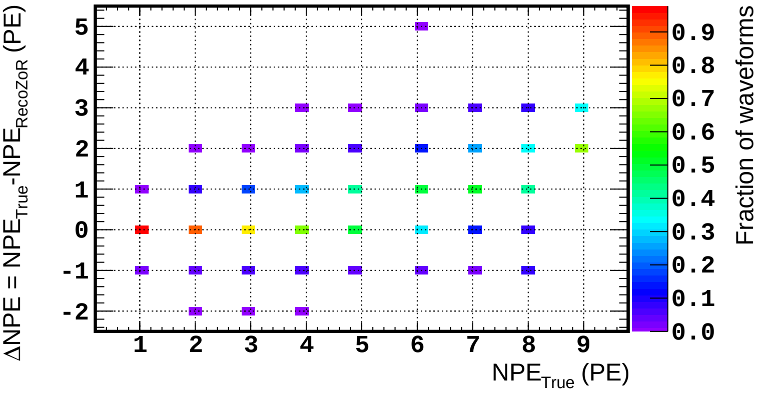


FIGURE 4.14: RecoZoR ΔNPE distribution as a function of the true number of PE. The color scale indicates the fraction of waveforms, generated with a given number of PE, for a given ΔNPE value. Only waveforms generated with PE validating $Q_{true} > 20$ a.u. and $t_{start,true} < 248$ ns conditions are shown. The statistic is very low (<100 waveforms) for $NPE_{True} > 6$ PEs .

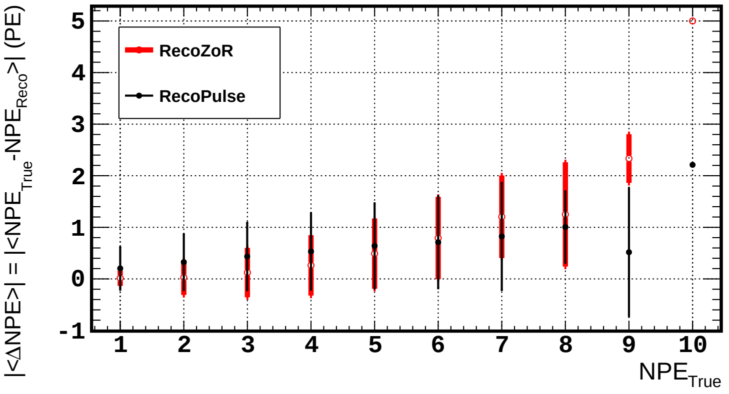


FIGURE 4.15: Absolute value of the mean value of the ΔNPE distribution as a function of the true number of PE for RecoZoR (*red*) and RecoPulse (*black*) (the error bars show the spreading of the ΔNPE distribution). Only waveforms generated with PE validating $Q_{true} > 20$ a.u. and $t_{start,true} < 248$ ns conditions are shown. The statistic is very low (<100 waveforms) for $NPE_{True} > 6$ PEs .

Applications of RecoZoR on data

In order to probe the capabilities of RecoZoR on real physics waveforms, the current version was applied on the physics data from the IBD selection. FIG. 4.16 shows an example of a complex waveform fitted by RecoZoR. It is a demonstration of RecoZoR capabilities to be adapted on complex cases, whereas it was developed only with simple waveforms of 1 to 3 PEs.

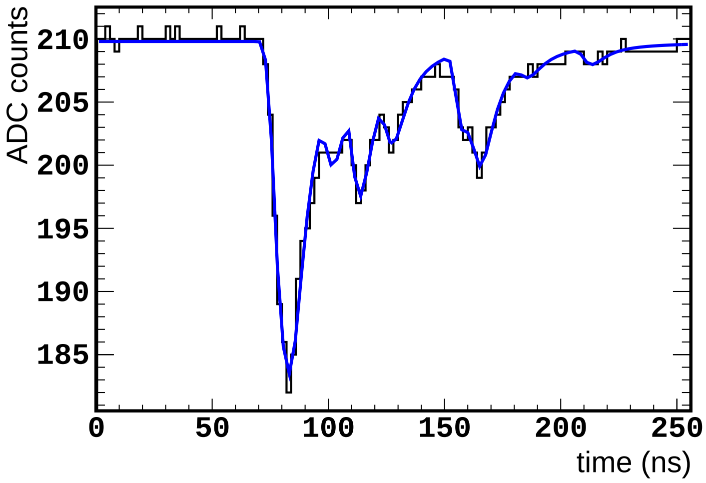


FIGURE 4.16: Waveform from IBD data fitted by RecoZoR. RecoZoR reconstructed 11 PEs and RecoPulse computed 12.59 PEs .

4.3 Summary and perspectives

We developed a new method for DC-like experiment calorimetry, trying to reconstruct each PE separately taking advantage of the waveform shape information. For a low number of PE (< 6) RecoZoR algorithm demonstrates, with this shape dependent method, better results than the shape independent calorimetry method using the integration of the waveforms (RecoPulse algorithm). The results of RecoZoR are better on the time resolution, on the resolution of the number of PE and on the linearity. The main difficulty for the shape dependent method is the “pill-up” affecting PEs. This effect increases as a function of the number of PE reaching the PMT. Then, the standard method is expected to show better results at high energy.

The algorithm presented here is still in development. On the current level we have already succeeded in having good resolution at 1 and 2 PEs for number of PE and the time, with an improvement of $\sim 2\times$ with respect to the current pulse reconstruction method. In addition, the linearity of the number of PE is also better than with RecoPulse when the number of PE is low (< 6). Investigations are ongoing to correct the (software) issue occuring when the PE interacts at the end of the waveform ($t > 248$ ns) and to improve results with higher number of PEs.

Futur developments of RecoZoR have been investigated. One is to use RecoZoR in order to correct for the non-uniformity of the detector response (cf. section 3.2.4): The scintillation light coming from the energy deposition within the liquid scintillator propagates itself until it reaches the PMTs. In a perfect case with a spherical PMT distribution, when an energy deposition occurs in the detector center, all PMTs will be hit at the same time. It means that all the first PEs of all PMTs will have the same starting time. However, if the energy deposition does not occur in the detector center PMTs will be hit at different times, depending on their position. With the current pulse reconstruction, using charge integration, it means that the first hit PMT will have a larger integration window, and the last hit PMT will have a shorter integration window. In other words, a part of the charge deposited in the last hit PMTs is missed. This causes a non-linearity of the charge integration, which is related to the non-uniformity of the detector response. Using RecoZoR computation of the waveform starting time it is possible to determine the best time window to be applied on all PMTs. This time window could be use for a charge integration using RecoZoR fit function, or as a time window to count the number of PE reconstructed by RecoZoR. This possibility is currently investigated, first results showed that it is a complex study due to the fluctuations of single-PE shape. Indeed, in the cases where the PE misses the first dynode, the pulse will have a reduced amplitude, but will reach the anode earlier than expected, about 30 ns before for Double Chooz PMTs. [154, 155]

The current good results, yet preliminary, allow us to consider using such a method in a close future. A combination between the single PE reconstruction method and the charge integration method could be considered for Double Chooz analysis. The collaboration plans to run in the future two parallel analysis, one standard and the other one using RecoZoR, in order to cross-check the consistency of RecoZoR results.

If validated, this method could be used in other experiments using similar electronics and PMT to Double Chooz. Investigations about it are ongoing within the JUNO collaboration.

Chapter 5

θ_{13} measurement in Double Chooz experiment

Double Chooz goal is the measurement of θ_{13}. In order to reach it, it is necessary to measure the oscillations of $\overline{\nu}_e$. Therefore, the selection of $\overline{\nu}_e$ within the data sample is needed. Due to the low IBD cross-section, IBD events can be considered as rare and the sample is expected to be dominated by background events. Consequently, a precise analysis has to be performed to select the IBD events and to reject the background events.

In this chapter, we cover the Double Chooz offline analysis performed to select $\overline{\nu}_e$ events. The two last analysis are covered: the last Gadolinium analysis, DC-III (G-n), and the last Hydrogen analysis, DC-III (H-n). Both analyses were performed on a sample of data taken since April 2011 to March 2012.

The different cuts applied to select IBD signal are described. Since the Double Chooz analyses suffers from several different backgrounds, several methods were developed to reject them, these different types of background and the methods used to reject them are also described. Among them, the Inner Veto veto method (IV-veto) is particulary detailed. This method, allowing to reject both accidental and correlated backgrounds (see section 5.1.1) was one of my contributions to the Double Chooz analyses during this thesis.

The results of the $\overline{\nu}_e$ selection are also detailed, with the computation of the remaining backgrounds rates and shapes. The details on the computation of the remaining correlated background are not presented in this chapter. It is detailed in the following chapter (chapter 6) since it was one of my contributions to the Double Chooz analyses. The different systematics which are present in the Double Chooz analyses, as well as the methods used to compute them, are described in this chapter. Finally, the extraction of θ_{13}, via a rate+shape fit or a Reactor Rate Modulation (RRM) fit of the $\overline{\nu}_e$ selection, is presented. The rate+shape fit is based on a combination of a fit of the rate and of the spectral shape in order to extract θ_{13}. A RRM fit allows to extract θ_{13} from observed and predicted rates at different reactor powers.

5.1 $\overline{\nu}_e$ events selection

As explain in section 2.1, Double Chooz use the Inverse β Decay in order to detect the reactor $\overline{\nu}_e$:

$$\overline{\nu}_e + p \rightarrow e^+ + n \tag{5.1}$$

The IBD signature is the coincidence between a prompt signal and a delayed delayed signal. The prompt signal is the energy released by the positron energy loss and by its annihilation, and the delayed signal is the energy released by the neutron capture on a Gadolinium or a Hydrogen nucleus. The neutron needs to be thermalized before its capture on Gadolinium or Hydrogen nuclei, thanks to the presence of Gadolinium in the ν-target, the mean thermalization time in this volume has been estimated to be $\sim 30\ \mu s$, in the γ-catcher, the mean thermalization time has been estimated to be $\sim 200\ \mu s$. More details on the fraction of neutron capture on Gadolinium or Hydrogen nuclei in the different volumes are given in section 5.3.2.2 and 5.3.3.2. FIG. 5.1 shows the expected spectrum of neutrino interacting in the detector.

Some particles can have a similar signature within the Double Chooz detectors. Therefore, the main difficulty of the analyses is to reject this background. Two analyses, or selections, were performed by the Double Chooz collaboration: the Gadolinium analysis, where neutrons from the IBD reactions are captured on Gadolinium nuclei, and the Hydrogen analysis, where neutrons from the IBD reactions are captured on Hydrogen nuclei. The main analysis of Double Chooz is the Gadolinium analysis. Each analysis has to deal with a different difficulties and different backgrounds. These different backgrounds are described in the following section.

5.1.1 Backgrounds

In both selections, the Double Chooz analyses suffer from several backgrounds. Three categories of backgrounds can be separed: the background induced by muons, the accidental background and the light noise.

The main background in the Gd-n analysis is coming from cosmogenic radioisotope produced by muon nuclear interactions in the detector, whereas the main background in the H-n analysis is the accidental background. The so-called light noise background was not expected at the beginning of the experiment, is the third background.

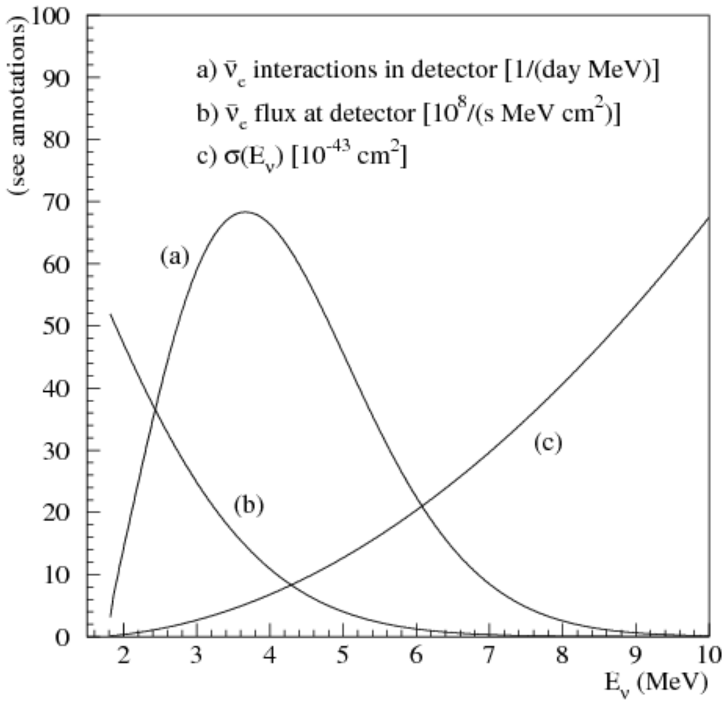


FIGURE 5.1: (a) $\overline{\nu}_e$ interaction in the detector, (b) $\overline{\nu}_e$ flux and (c) IBD cross section. The Y-axis is in arbitrary unit. Figure from [156]

5.1.1.1 Muon-induced background

Muons are one of the main sources of background in the Double Chooz $\overline{\nu}_e$ selections. They produce several particles which can have a signature similar to the IBD reaction in the liquid scintillator: fast neutrons, stopping muons and cosmogenic radio-isotopes like ^{9}Li, ^{8}He or ^{12}B.

Therefore, the Double Chooz analyses try to reject muon events by a muon tag and the application of a dead time after muons. The same method is applied by other reactor experiments, Daya Bay [46] and RENO [49].

Cosmogenic nuclei background

Cosmogenic nuclei background comes from radio-isotope produced by cosmic muon nuclear interactions within the detector. These interactions can produced several isotopes, however ^{9}Li and ^{8}He are the most dangerous for the Double Chooz analyses. Both decay via β^- decay and have of a probability to emit a neutron after their decay [146]:

$$\begin{aligned} {}^{9}Li &\xrightarrow{50.80\%} e^{-} + \overline{\nu}_{e} +^{8} Be + n \\ &\xrightarrow{49.20\%} e^{-} + \overline{\nu}_{e} +^{9} Be \end{aligned} \tag{5.2}$$

$$\begin{aligned} {}^{8}He &\xrightarrow{16.00\%} e^{-} + \overline{\nu}_{e} +^{7} Li + n \\ &\xrightarrow{84.00\%} e^{-} + \overline{\nu}_{e} +^{8} Li \end{aligned} \tag{5.3}$$

In addition, the maximum energy released by the electron in β^{-} decay is above 10 MeV (10.6 MeV for ${}^{8}He$ and 13.6 MeV for ${}^{9}Li$ [146]). The energy range of the electron is very similar to that of the positron from IBD reactions. Therefore, since the Double Chooz detectors do not have a magnetic field allowing to discriminate electron and positron, it is very difficult to distinguish between these decays and an IBD reaction. FIG. 5.2 is an illustration of this background.

Since the long half-life time of ${}^{9}Li$ and ${}^{8}He$ (119.1 ms for ${}^{8}He$ and 178.3 ms for ${}^{9}Li$ [146]), it is not possible to reject them by applying a dead time after the detection of a muon. Such huge dead time would have a high cost on the detection efficiency. The previous Double Chooz analysis applied a similar method applying a high dead time (0.5 sec) after the passage of cosmic muons depositing more than 600 MeV in the detector, and applying a shorter dead time (1 ms) after cosmic muons depositing less energy. This method increased the dead time by a factor ~ 2, reaching 9.2% of the total run time [147]. In the current analysis, only the 1 ms dead time after the passage of a muon is applied (see section 5.1.2.1). The Daya Bay collaboration applies a method similar to [147], rejecting all events occuring 1000 μs after the passage of a muon, 1000 μs if the energy deposited by the cosmic muon is more than 20 MeV, and 1 s if the energy if more than 2.5 GeV, respectively [48]. The RENO collaboration also applies a 1 ms dead time after the passage of cosmic muons, and increases this dead time to 10 ms if the cosmic muon deposits an energy higher than 1.5 GeV [49].

Correlated background

The so-call correlated background consists in fast neutron and stopping muon interactions within the detector. Thanks to the dead time applied after the passage of cosmic muons, most of the fast neutrons produced inside the detector are rejected, therefore, this background mostly comes from the outside of the detector. Fast neutrons can cross the Inner Veto and the buffer, and be thermalized within the liquid and captured on Gadolinium or Hydrogen nuclei. The thermalization is done via proton recoils, i.e. collisions with the Hydrogen nuclei within the liquid. Due to the large energy spectrum of the fast neutrons (from some keV to more than several GeV), these proton recoils can release an energy close to that of a prompt-like signal. Since the fast neutron thermalization, which is similar

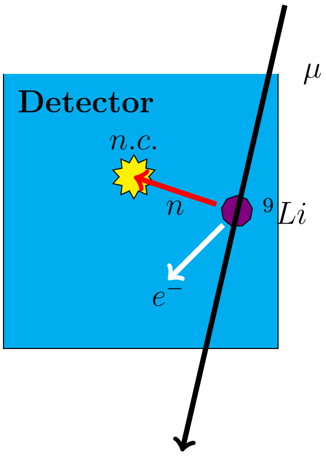


FIGURE 5.2: Illustration of the cosmogenic background with ^9Li isotope. Muon spallations on ^{12}C nuclei within detector liquid can produce ^9Li or ^8He radioisotope. These nuclei can decay with a neutron emission, and then mimic an IBD reaction. $n.c.$ is for neutron capture.

to the IBD neutron, the time coincidence between the prompt-like signal and the fast neutron capture is equivalent to the time coincidence of the IBD prompt and delayed signals. This produces an IBD-like signature.

Stopping muons are low energy muons which decay into the detector. Thanks to the Inner Veto of the detector which allows to reject almost all of the cosmic muons crossing it, the stopping muon background comes only from cosmic muons sneaking into the detector via the chimney (see section 2.2 for the design of the detector). Since 2012, the addition of the Upper Outer-Veto (see section 2.2.1.3) allows to detect these muons.

Inside the detector stopping muons lose all their remaining energy by ionization, which can produce a prompt-like signal. Their decay produces a Michel electron, which carries a high amount of enery (from almost 0 MeV to ~ 70 MeV [114, 115]). Its energy loss in the detector can mimic a delayed-signal.

FIG. 5.3 is an illustration of these backgrounds. More details on correlated background and about its analysis can be found in the next chapter (chapter 6).

5.1.1.2 Accidental background

Accidental background is due to a random coincidence between a prompt-like signal and a delayed-like signal.

In the main case, the prompt-like signal is due to γ from internal or external radioactivity, whereas the delayed-like signal is the capture of a fast neutron on a Gadolinium or a Hydrogen nucleus. The radioactivity γ spectrum is mostly a

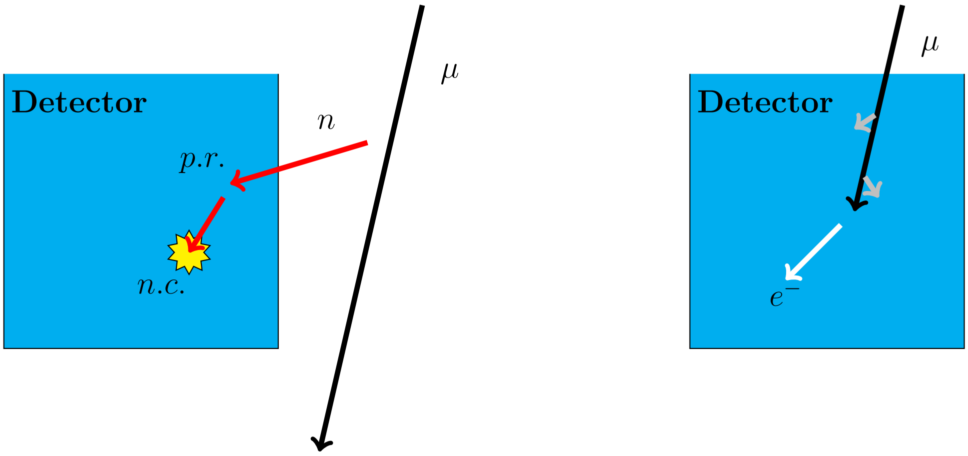


FIGURE 5.3: Illustration of the correlated backgrounds. On *left* is the schematization of the fast neutron background, and on *right* the stopping muon background. Gray arrows coming from stopping muon path are the schematization of the muon energy loss by ionization. *n.c.* is for neutron capture and *p.r.* for proton recoil.

low energy spectrum, going up to $4 - 5$ MeV [157]. In the Hydrogen analysis, the Accidental background is the main background due to the γs from ^{206}Tl and ^{40}K decays. ^{206}Tl emits a ~ 2.6 MeV γ which can mimics a delayed signal or a prompt signal.

Using the time and space correlations between the prompt-like and delayed-like signals, it is possible to reduce this background. This is used by the three reactor $\overline{\nu}_e$ experiment, the cuts are described in section 5.1.2.5 for the Double Chooz Gadolinium analysis, a different method was applied for the Double Chooz Hydrogen selection (see section 5.1.3.2). Details on the Daya Bay and RENO analyses can be found in [48, 49].

5.1.1.3 Light Noise

The so-call "Light noise" is an unexpected background discovers at the beginning of the data taking in the Far Detector. It is due to spontaneous flashes from some PMT bases. Several tests were performed, and demonstrated that the origin of the light noise was PMT glowing. It has been shown that the reduction of PMT HV could reduce the light noise. Investigations on a possible glowing from the epoxy used to cover the PMTs base have been also performed and demonstrated the dielectric properties of epoxy: the epoxy can produce light in an electric field, this process is favored by an increase of temperature [158].

5.1.2 Gadolinium selection (DC-III (Gd-n))

The Gadolinium selection is the main analysis of Double Chooz. Due to the Gd concentration and the number of proton within the Double Chooz liquids, $\sim 85\%$ of neutron capture within ν-Target volume happen on a Gd nucleus (see section 5.3.2.2). The mixing between Hydrogen and Gadolinium leads to a mean capture time for neutron on Gd of $\sim 30\,\mu s$.

Only the valid triggers within the data sample are used for the selection. The valid triggers are defined as:

- Not a random trigger
- Not a muon (see section 5.1.2.1)
- $\Delta T_\mu > 1$ ms
- $E_{vis} > 0.4$ MeV
- Not a light noise (see section 5.1.2.2)

where ΔT_μ is the time difference from the last muon tagged (see section 5.1.2.1). The 0.4 MeV cut was defined due to the trigger inefficiency below 0.4 MeV, as shown Fig. 5.4. Random triggers are the random activation of the trigger system, occuring at a fixed rate, and used to calibrate the baseline of the detector PMTs (see section 3.1).

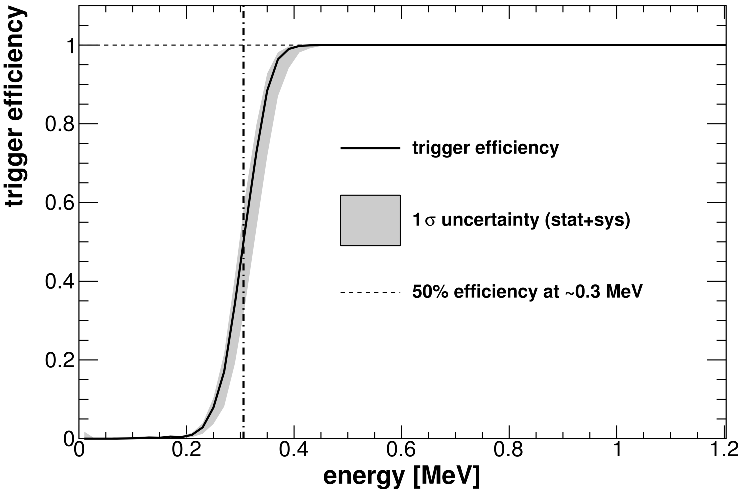


Figure 5.4: Trigger efficiency as a function of the energy [161]

5.1.2.1 Muon veto

In order to reject the main part of the muon-induced backgrounds, we tagged muon events with our Inner Veto and our Inner Detector. Indeed, due to the calibration hole in the Inner Veto, $\sim 2\%$ of the muons can sneak inside our detector without hitting the IV. The muon definition is the following:

- $E_{vis} > 20$ MeV
- $Q_{IV} > 30000$ a.u.

The cut on the energy deposited in the Inner Detector was defined to reduce the contamination of stopping muons. Muons not decaying inside the detector are expected to cross the full volume, and therefore should be rejected by the Inner Veto. The Q_{IV} definition corresponds to ~ 15 MeV.

Due to the high energy range of the fast neutrons (see section 5.1.1.1), both rejection definitions can reject high energy fast neutrons as muons. This is not an issue as the presence of fast neutrons can be the signature of a nearby muon. We defined a time window after muon of $1\,ms$ where all events are rejected. This dead time was defined using time correlation between IBD candidates and muons [179].

Muons crossing rocks near the detector produce a lot of fast neutrons, however most of them do not deposit energy in the IV or in the ID. In order to reject the remaining muon background some additional cuts were defined and are described in this chapter.

5.1.2.2 Light noise rejection

Several studies were performed to reject the light noise background. In DC-III (Gd-n) analysis, the following cuts are applied:

- $Q_{max}/Q_{tot} < 0.12$
- $rms(t_{start}) < 36$ ns or $rms(Q) < 464 - 8 \cdot rms(t_{start})$
- $Q_{diff} < 30000$ a.u.

where Q_{max} is the maximum charge reconstructed by a PMT and Q_{tot} is the total charge reconstructed in the ID. t_{start} is the start time of the first pulse, and $rms(t_{start})$ is the standard deviation of the t_{start} distribution for all PMTs. $rms(Q)$ is the standard deviation of the reconstructed charge distribution for all PMTs.

A two dimensional cut is made in order to minimize the IBD inefficiency, allowing to reject more precisely the light noise events, as show in FIG. 5.5.

Q_{diff} is defined as follows:

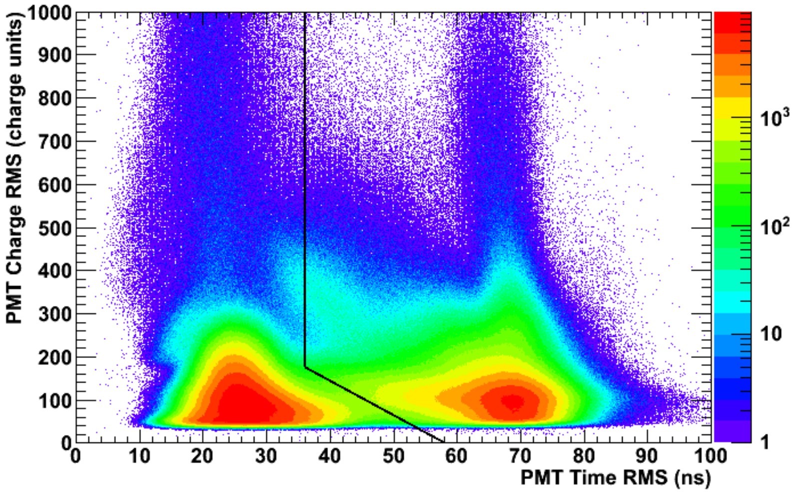


FIGURE 5.5: PMT Charge RMS vs PMT Time RMS spectra. Q_{max}/Q_{tot} and Q_{diff} cuts have been applied. The black line represents the 2D cut on the PMT Charge RMS / PMT Time RMS plan. Physics data belong to the left side of this line whereas light noise data belong to the right side of this line.

$$Q_{diff} = \frac{1}{N}\sum_{i}^{N}\frac{(Q_{max} - Q_i)^2)}{Q_i} \tag{5.4}$$

where N is the number of PMTs within a 1 m radius sphere centered on the position of the PMT which received the maximum charge. Q_i is the charge reconstructed for each PMT within the sphere.

This variable is used to represent the non-uniformity of the observed charge for neighboring PMTs around the one with the maximum charge. Light noise events are likely to be non-uniform since the light noise come from few PMTs. The threshold on Q_{diff} was set to avoid IBD candidate rejection while rejecting light noise events.

The inefficiency of light noise cuts was estimated to be 0.012% for DC-III (Gd-n) analysis, using MC simulations. It was noticed that the MC variable distributions were shifted with respect to the data variable distributions. Two analyses were performed to compute this inefficiency, one using MC variables shifted to match with the data ones, and one using non-shifted variables. Some additional light noise events are also rejected by FV veto (see section 5.1.2.7).

5.1.2.3 Prompt energy window

In the DC-III (Gd-n) analysis, the prompt energy window was defined as:

- $E_{vis} \in [0.5,\ 20]$ MeV

It was demonstrated that the trigger efficiency reaches 100% between 0.4 and 0.5 MeV, as shown in Fig. 5.4. The lower cut on prompt energy was then put at 0.5 MeV in order to take advantage of this trigger efficiency to get more informations on the spectrum shape. As the background is dominant below 0.7 MeV, shape information between 0.5 and 0.7 MeV can be used to validate the background estimation.

The positron energy deposition is expected to go up to about 12 MeV. The prompt energy window was extended to 20 MeV. Events between 12 and 20 MeV are expected to be background, and more precisely correlated background (see chapter 6). These events can be used to add a stronger constraint on the correlated background with the rate+shape fit (see section 5.5.1).

5.1.2.4 Delayed energy window

In the DC-III (Gd-n) analysis, the delayed energy window was defined as:

- $E_{vis} \in [4,\ 10]$ MeV

Neutron Gadolinium captures emit a several γs with a total energy of ~ 8 MeV. Since the IBD contribution above 10 MeV is negligible, the higher cut was placed at 10 MeV to reduce the background, mainly due to stopping muons with high energy Michel electrons, or high energy fast neutron proton recoils. The lower cut was tuned by comparing the selection efficiency and the accidental background contamination. The 4 MeV cut gives a signal inefficiency about 1% and reduce the accidental by a factor ~ 7, with respect to a lower cut at 3 MeV.

In DC-II selection, the delayed energy window was $E_{vis} \in [6,\ 12]\,MeV$ [147]. In DC-III (Gd-n), thanks to the wider energy window at low energy, the neutron detection efficiency is higher, now close to 100% (see section 5.3.2.2). Therefore, the energy reconstruction dependence of the detection efficiency is reduced. This mean that the systematic uncertainty due to this delayed energy window cut is smaller for the same error on the energy reconstruction. However, the accidental background rate is higher.

5.1.2.5 Prompt and delayed coincidence

In DC-III (Gd-n) analysis, the prompt-delayed timing coincidence is defined as:

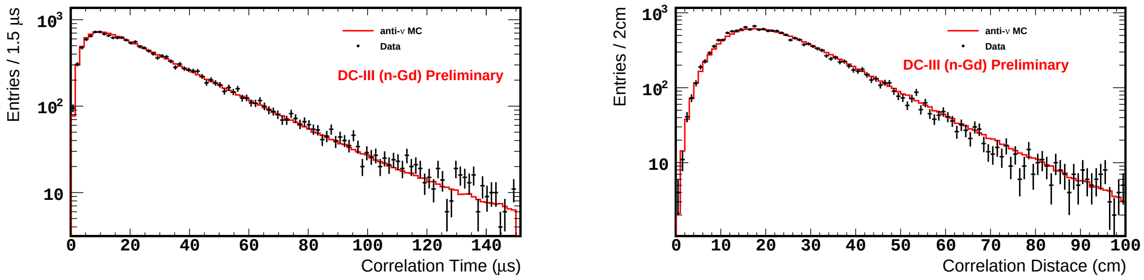


FIGURE 5.6: Prompt-Delayed time difference (*left*) and Prompt-Delayed vertex distance (*right*). The black points represent the events selected by the Gadolinium selection. The red line shows the MC expectation.

- $\Delta T \in [0.5,\ 150]\ \mu s$
- $\Delta R < 100$ cm

where ΔT and ΔR are, respectively, the time difference and the distance in space between the prompt and the delayed events. FIG. 5.6 shows the time difference and the distance in space between the prompt and the delayed events for the Gadolinium selection.

The lower cut on ΔT was defined to avoid a bias due to the pile-up effect. This effect is due to the scintillation light from a previous energy deposition in the liquid scintillator which bias the following energy deposition. An effect due to the electronics is also possible. Another benefice of this lower cut is the reduction of the stopping muon background, since muon have a low lifetime ($\sim$ 2 μs), the timing coincidence between the muon decay and the Michel electron is expected to be low.

The upper cut on ΔT was tunned considering the model dependence of spill-in/out effect (see section 5.3.2.2) and the IBD signal inefficiency from MC. The signal loss with respect to an upper cut at 200 μs was estimated to be 0.8% at 150 μs. The relative difference of spill-in/out flux was computed using two neutron simulations, GEANT4 and Tripoli4, to be 0.27% at 150μs. This upper cut allows to keep a high efficiency close to 100%.

ΔR is computed using the reconstructed vertex positions from RecoBAMA. This cut had been tunned to maximize the accidental rejection while keeping an high IBD efficiency. The resulted inefficiency of IBD signal is 0.3%, while this cut reduce the accidental background rate by a factor 7.

5.1.2.6 Multiplicity cut

The multiplicity of events following the passage of a muon was studied. The high multiplicity is interpreted being due to the production of fast neutrons by muon spallation. Therefore, it is possible to take advantage of this multiplicity to reject fast neutrons. The following cut was defined for the Gadolinium analysis:

- No valid triggers allowed in $200\,\mu s$ before the prompt candidate.
- No valid triggers, except the delayed candidate, allowed in $600\,\mu s$ after the prompt candidate

This cut was employed to reject coincidence with more than two signals. It was optimized in DC-III (Gd-n) using the accidental coincidences between three or more events [159].

5.1.2.7 FV veto

RecoBAMA, the Double Chooz vertex reconstruction algorithm, defines a negative log-likelihood of the observed charge and hit timing distribution at the reconstructed vertex. This log-likelihood, called Functional Value (or FV), is expected to become larger when the light source is not point-like or when it is shifted from the true vertex.

Since the muon deposits energy over all its path, stopping muon interactions in the liquid scintillator are not expected to be point-like. The same applied for the Michel electron produced by the stopping muon decay, due to its high energy.

The FV can be used to separate stopping muon and IBD candidates. This was confirmed for delayed signals on the (FV, E_{vis}) plane and the cut was optimized to separate them. Fig. 5.7 shows the separation between IBD candidates and stopping muons. However, the separation is not evident for prompt signals. One possible interpretation of this difficulty, is the fact that positrons from IBD reations can also have non-point like interactions, mostly when forming a positronium. In this case, they also deposit energy over all their path before their annihilations.

Therefore, this cut is applied on the delayed candidates only, and it rejects all delayed-like events which correspond to the following definition:

- $E_{vis} > 0.068 \cdot exp(FV/1.23)$

Events rejected by FVV behave as expected for stopping muons, showing no peak in the delayed energy spectrum, and a tight exponential shape in ΔT distribution. In addition, their vertex positions are clustered below the chimney. This is interpreted as a demonstration of the stopping muons rejection by the FV veto. Studies also showed that FV veto reject some of the remaining light noise [160].

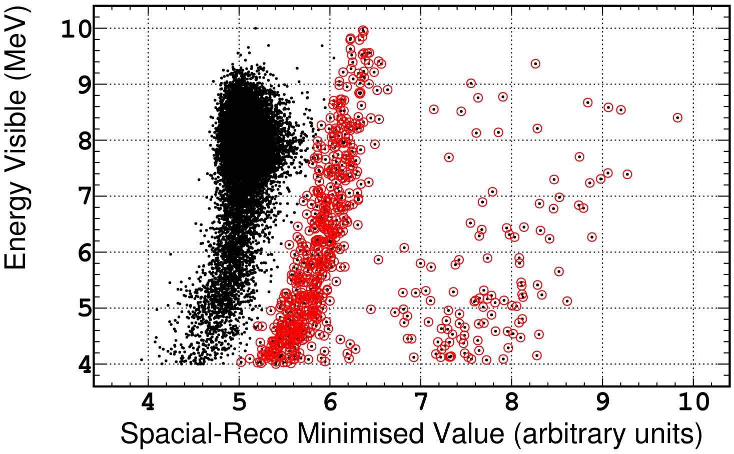


FIGURE 5.7: Correlation between delayed signal energy and the vertex reconstruction log-likelihood. In black the full data set, in red events rejected by the FV veto. Three population of events can be separate with this correlation: IBD-like events (left), stopping muons events (middle) and light noise events (right).

5.1.2.8 IV veto

The definition of IV veto is one of the result of my thesis. The motivation and the definition of this cut are detailed in section 5.1.4. In the Gadolinium analysis it rejects prompt signals which are selected by the following conditions:

- Minimal IV tagging conditions ($Q_{IV} > 400$ a.u. and IV PMTs multiplicity ≥ 2)
- $\Delta t^{QW}_{ID-IV} \in [-110,\, -10]$ ns
- $\Delta d_{ID-IV} < 3.7$ m

where Δt^{QW}_{ID-IV} is the time difference between the signal in the IV and in the ID, and Δd_{ID-IV} is the space distance between the IV interaction vertex and the ID interaction vertex.

5.1.2.9 OV veto

The detector OV is also used to tag muons which do not pass through the detector. Those muons often produce fast neutrons or decay within the detector as stopping muon. We reject IBD candidates whose the prompt signal is coincident with an OV trigger.

5.1.2.10 ^{9}Li reduction

In order to reduce the remaining ^{9}Li contamination, a likelihood based on the number of neutron candidates following a muon and on the distance between the neutron candidates vertex positions and the muon track.

The neutron candidates were selected in a sample of event occuring within the 1 ms time window after a muon. An energy window cut of $E_{vis} \in [4,\ 10]$ MeV was applied. For each muon, the number of neutron candidates and the distribution of their vertex position and the muon track were computed.

Several probability distribution functions (PDF) were generated using a sample of ^{12}B events and using a sample of ^{9}Li events: PDF of the number of neutron candidates following a muon and PDF of the distance between the neutron candidates vertex position and the muon track. PDFs from the ^{12}B sample were confirmed to be similar to the one from the ^{9}Li sample. In order to compute the likelihood, the PDFs from the ^{12}B sample were used, in order to have a better statistic.

The likelihood values are computed for each event in the Gadolinium dataset for all events tagged as muons within 700 ms before the event. Prompt candidates are tagged as ^{9}Li events and rejected if the following condition is satisfied:

- $\mathcal{L}_{^{9}\mathrm{Li}} > 0.4$

This likelihood rejects 1.1 $^{9}\mathrm{Li}\, events/day$, while the induced inefficiency has been computed to be $0.536 \pm 0.019\%$. Events passing all the selection cuts but this one can be use to measure the spectrum shape of cosmogenic background as detailed in section 5.1.1.1.

5.1.3 Hydrogen selection (DC-III (H-n))

Double Chooz was the first reactor experiment to perform a Hydrogen analysis, followed one year later by Daya Bay. A Hydrogen analysis can be useful for the Double Chooz experiment in order to cross-check the results of the Gadolinium analysis. The $[4, 6]$ MeV distortion, discovered by Double Chooz with its Gadolinium analysis can be confirmed, as well as the value of θ_{13}. The Hydrogen analysis can also be used for a combination with the Gadolinium analysis results, increasing the resolution on the precision result.

The advantage of the Hydrogen analysis compared to the Gadolinium analysis is the increase of the scatistics, thanks to the the wider fiducial volume (the ν-target and the γ-catcher will both capture the neutron) and to the larger number of Hydrogen nuclei in the liquids compared to the number of Gadolinium nuclei. However, the neutron capture cross-section on Hydrogen nuclei is smaller than the one on Gadolinium nuclei, this leads to a higher neutron thermalization time in the γ-catcher volume compared to the one in the ν-target volume and to a smaller statistics in the ν-target, due to the competition with Gadolinium nuclei. In addition, the Hydrogen neutron capture signal is less clear than the Gadolinium neutron capture signal. Hydrogen neutron capture releases only one γ of 2.223 MeV. This energy is lower than the one released by Gadolinium neutron captures but it is also in the energy spectrum contaminated by the accidental background. This leads to an increase of the accidental background.

Therefore, in order to carry out a Hydrogen analysis, it is mandatory to take into account the particularities of Hydrogen neutron capture. The backgrounds in the Hydrogen analysis are not expected to be the same than for the Gadolinium analysis. New methods were also developed to deal with the accidental background and allow a huge reduction of its rate. Compared to the previous Double Chooz Hydrogen analysis [55], the accidental background was reduced by a factor ~ 17 in the DC-III (H-n) analysis. I contributed to this reduction with the development of the Inner Veto veto method, which allows to reject $\sim 27\%$ of the accidental background (see section 5.1.4).

Since the Hydrogen selection has to be compared with the Gadolinium selection, most of the Gadolinium analysis cuts were kept or adapted for the Hydrogen selection. The valid triggers in the Hydrogen selection were defined as:

- Not a random trigger
- Not a muon
- $\Delta T_\mu > 1.250$ ms
- $E_{vis} > 0.4$ MeV
- Not a light noise

Random trigger, muon and light noise definitions were kept the same as in Gadolinium selection. However, the dead time after muon was modified and tunned for the Hydrogen analysis to 1.250 ms.

5.1.3.1 Prompt energy window

In DC-III (H-n) analysis, the prompt energy window was defined as:

- $E_{vis} \in [1, 20]$ MeV

Since the γ-catcher is surrounded by a non-active volume, the buffer, γ(s) from the positron annihilation can escape γ-catcher to the buffer. It is also possible to have IBD reactions occuring within the buffer with one γ from the positron annihilation entering in the γ-catcher. In this case, if the neutron from the IBD

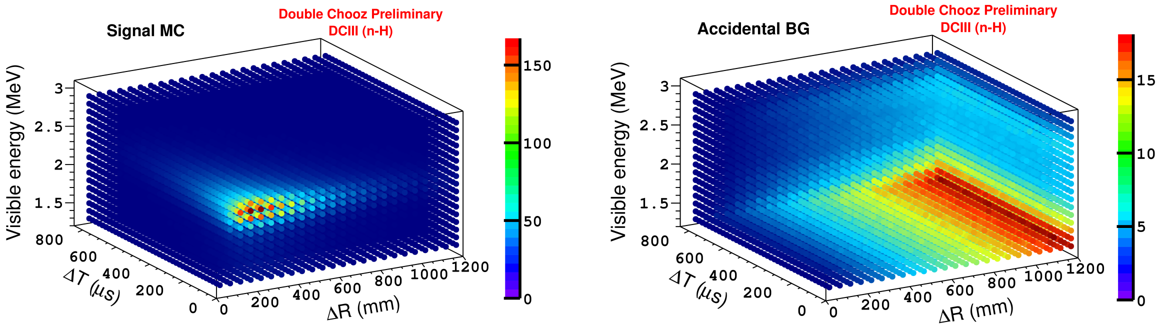


FIGURE 5.8: 3D plots showing the relations between the delayed signal energy, the prompt-delayed space correlation and the prompt-delayed time correlation, for IBD MC (*Left*) and accidental background (*Right*).

reaction is captured on a Hydrogen nucleus in the γ-catcher, the coincidence of the two signals, and therefore the IBD reaction, can be detected. However, these kind of prompt event could induce a bias in the prompt energy spectrum as these spill effects (see section 5.3.2.2) are not correctly simulated by the MC. In order to avoid them, the lower energy cut was increased to 1 MeV.

5.1.3.2 Articifial Neural Network cut and delayed signal

In the Gadolinium analysis, we selected the IBD signals using three separate cuts on the delayed signal energy (section 5.1.2.4), and the prompt-delayed space and time correlations (section 5.1.2.5). However, during the development of the Hydrogen analysis, studies showed interesting relations between these three variables. As shown, in FIG. 5.8, the relations between these three variables for IBD signals are very different than the one for the accidental background. Therefore, they can be use to separate the IBD signal from the accidental background.

The Double Chooz collaboration developed a multivariate analysis using an Artificial Neural Network (ANN) (cf. 2.7.2.1 for more details about Artificial Neural Networks). This ANN was applied on IBD candidates and took as input the delayed event energy and the space and timing correlation between prompt and delayed events [162]. FIG. 5.9 shows the distributions of these three variables for the IBD signal from MC simulations, for the accidental background and for the data.

The ANN allowed to use the relation between the three variables and to compute a value which can be interpreted as a probability to be an accidental event. The ANN was configured to give an output value of -1 for accidental events and of 1 for IBD events. Thanks to the use of this ANN relaxed cuts on the delayed en-

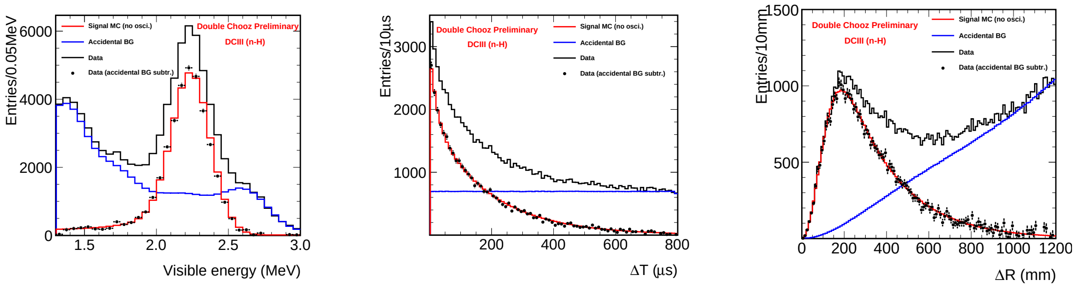


FIGURE 5.9: Delayed signal energy spectrum (*left*), prompt-delayed time correlation distribution (*center*) and prompt-delayed space correlation distribution (*right*. The red line is the MC simulation, the blue line is the accidental background, the black line is the data before background subtraction and the black points are the data after application of all selection cuts.

ergy window and on the prompt-delayed space and time correlations were defined. These cuts were tunned as:

- ANN output > -0.23
- E_{vis} (delayed) $\in [1.3,\ 3]$ MeV
- $\Delta T \in [0.5,\ 800]\ \mu s$
- $\Delta R < 120$ cm

As the Hydrogen neutron captures emit one γ of 2.223 MeV, both low and high energy borders were redefined to this energy, and tunned in order to keep a high IBD efficiency. The mean neutron capture time in γ-catcher is about 200 μs whereas it is 30 μs in the ν-target, as a consequence, the mean distance traveled by neutron is expected to be longer too. Therefore, both ΔT and ΔR cuts were adapted for Hydrogen selection.

With these cuts, and mostly thanks to the ANN cut, the ratio S/N reaches 11 with a signal efficiency of 80% [162]. In H-II selection, the ratio S/N was 1.1 with a signal efficiency of 85% [55]. The loss of efficiency is compensated by the higher S/N.

5.1.3.3 Multiplicity cut

In order to be compatible with the higher prompt-delayed time coincidence in the γ-catcher, the isolation cut needed to be modified in the DC-III (H-n) analysis, . Whereas the time coincidence cut in the Gadolinium selection was 150 μs, the time coincidence in the Hydrogen selection is 800 μ s. The multiplicity cut window needs to be larger than the prompt-delayed time coincidence cut. This would not be true if the Gadolinium selection multiplicity cut was used.

Therefore, several analyses were performed and tunned the multiplicity cut values for the Hydrogen analysis to be:

- No valid triggers allowed in 800 μs before the prompt candidate.
- No valid triggers, expect the delayed candidate, allowed in 900 μs after the prompt candidate

5.1.3.4 FV veto

The FV veto has been tuned for the Hydrogen selection. This cut was applied only on delayed candidates as follows:

- $E_{vis} > 0.2755 \cdot exp(FV/2.0125)$

FIG. 5.10 shows the separation between IBD candidates and stopping muons.

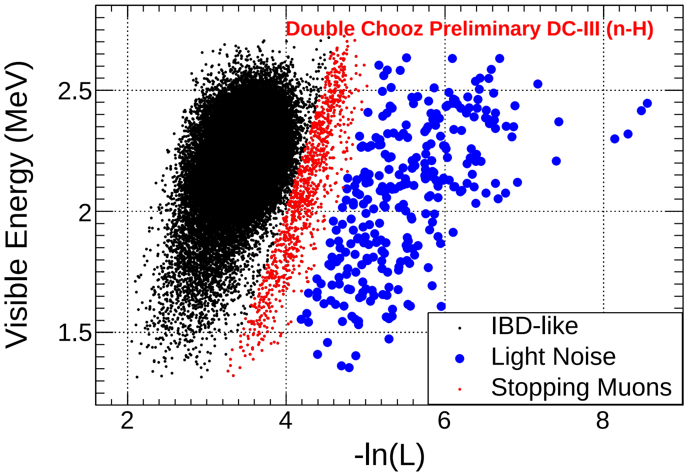


FIGURE 5.10: Correlation between delayed signal energy and the vertex reconstruction log-likelihood. Three populations of events can be identified : Light Noise events in blue, Stopping Muons in red, and IBD-like events in black. The IBD-like events include accidentals, ^{9}Li, fast neutrons... as they are physical events with a point-like light deposition (Only ANN cut was applied).

5.1.3.5 IV veto

The IV veto was tunned for the Hydrogen selection. In this analysis, the IV veto aims to reject mainly accidental background, taking advantage of the fact that γs from radioactivity from the external rocks can cross through IV to ID. This development is one of the contributions of my thesis. The IV veto is detailed in section 5.1.4.

The IV veto reject prompt and delayed signals in coincidence with an activation of the IV, which satisfy with the following conditions:

- Minimal IV tagging conditions ($Q_{IV} > 400$ a.u. and IV PMTs multiplicity ≥ 2)
- $\Delta t^{QW}_{ID-IV} \in [-110, -20]$ ns
- $\Delta d_{ID-IV} < 4$ m

5.1.3.6 Multiplicity pulse-shape cut

The Multiplicity pulse-shape cut (MPS) is a new method aiming to reduce the fast neutron background in the Hydrogen selection. It was developed after the Gadolinium selection and therefore not included in the DC-III (Gd-n) analysis.

MPS is based on the pulses start time distribution of the PMTs. A fit is performed to extract the start time of the pulses in each PMT waveforms (cf. section 3.1). This method to compute the start time is not the one applied by RecoPulse, neither than the one which will be applied by RecoZoR. Instead, it is simplified method which allow to compute a start time for all pulses in the waveform, whereas RecoPulse computes only one start time.

The time of flight between the event reconstructed position and the PMT is substracted to the computed start time for each pulse. Then, a pulses start time distribution is build. This distribution is shifted to start at 0 ns. Clear peaks in the pulses start time distribution are expected for significant energy depositions in the detector.

It was demonstrated, using a ^{60}Co source, that for electromagnetic interactions, the main peak is close to the start of the time distribution, as shown on the left figure in Fig. 5.11. However, using a sample of events tagged by the OV (i.e. correlated background), the main peak is shifted from the start of time distribution, as shown on the right figure in Fig. 5.11. This shift was interpreted as due to the multiple proton recoils from the fast neutron thermalizations. The rejection of events presenting a shift in the pulses start time distribution can allow to reject fast neutron events. Therefore, after some studies to tune the shift value, the MPS cut reject prompt signal with a $Shift > 5$ ns.

However, note that the pulses start time distribution of ortho-positronium events presents two peaks: the first one due to the positron ionization and the second one due to the 2 γs emission of the ortho-positronium decay. As the MPS algorithm computes the shift of the highest peak within the pulses start time distribution, a shift can appear in the cases where the ortho-positronium decay is more energetic than the positron ionization. Therefore, the MPS cut should not be applied on the ortho-positronium events.

The Double Chooz collaboration published an analysis allowing to identify the ortho-positronium events [170]. In order to avoid the rejection of IBD with ortho-positronium, the MPS cut was not applied for low energy events identified as ortho-positronium events with this analysis. However, as the ortho-positronium idenfication algorithm does not work below 1.2 MeV, the MPS cut was not applied for such low energy events.

Consequently, the MPS cut reject prompt candidates which satisfied all the following conditions:

- $Shift > 5$ ns
- $E_{vis} > 1.2$ MeV
- If $E_{vis} < 3$ MeV: Not an ortho-positronium event

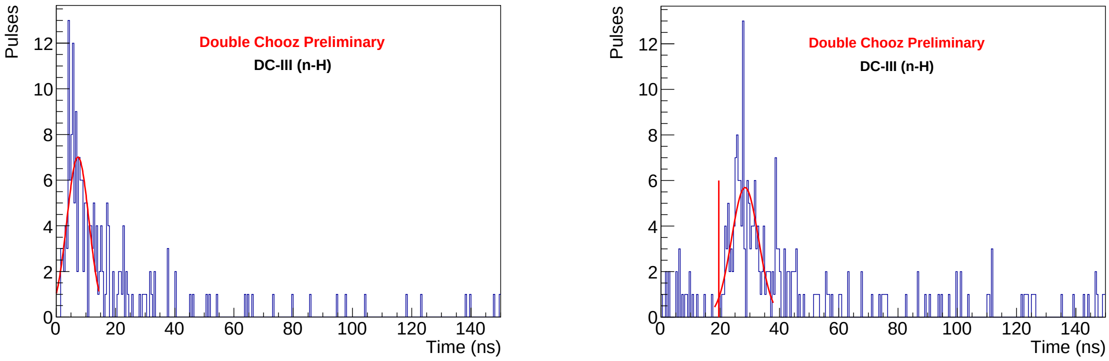


FIGURE 5.11: Pulse time distribution for an IBD event (*Left*) and a fast neutron event, tagged by the OV (*Right*). Fast neutron event demonstrates a shifted peak compared to the IBD event.

5.1.3.7 Other cuts

OV veto and ${}^{9}Li$ reduction cuts are applied in DC-III (H-n) selection as they are in the Gadolinium selection.

5.1.4 Inner Veto Veto

Several backgrounds in the Double Chooz analyses originate from the outside of the detector. As a Double Chooz event consists in an energy deposition occuring within a 256 ns time window (cf. section 2.2.2), it is expected that a part of this background, coming from outside, occurs simultaneously with an energy deposition in the Inner Veto. If this energy deposition is lower than the muon threshold in the IV, this event will not be tagged as a muon.

Then, the aim of the Inner Veto veto is to tag and reject background events in coincidence with an energy deposition within the IV. Three different background events can be tagged by IV: fast neutron, accidental and, to a lesser extent, stopping muons. FIG. 5.12 schematizes these three kinds of IV events.

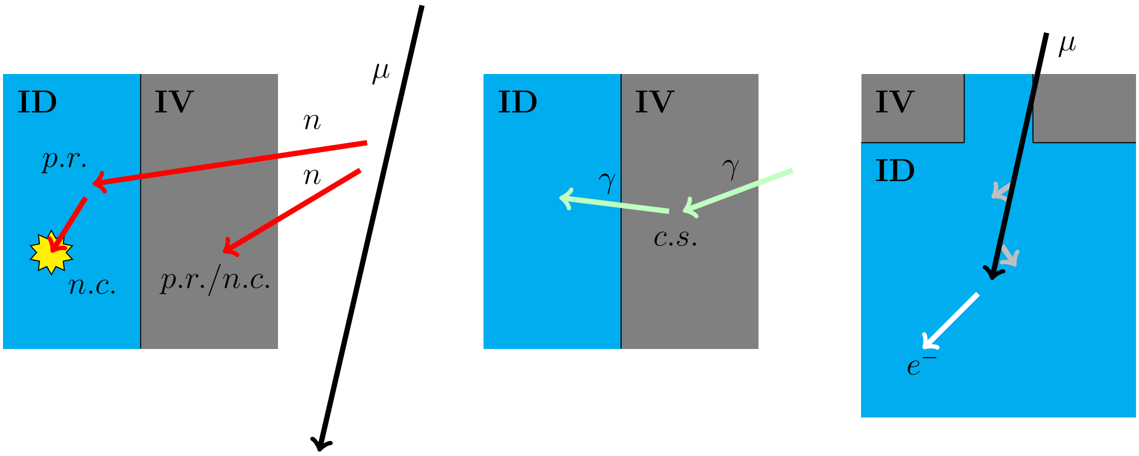


FIGURE 5.12: Illustration of the IV events backgrounds. On left is the schematization of the fast neutron events, on the center is the $\gamma - \gamma$ events, and on right the stopping muon events. Gray arrows coming from stopping muon path are the schematization of the muon energy loss by ionization. *n.c.* stands for neutron capture, *c.s.* for compton scattering, and *p.r.* for proton recoil.

Due to the high fast neutron yield per spallation, multiple fast neutrons can cross the detector simultaneously. Despite the low interaction cross-section, when one fast neutron interacts within the ID and causes a background event, it is possible to have another fast neutron, from the same muon, interacting within the IV via a proton recoil or a capture on a Hydrogen nucleus. The rejection of this kind of IV event is the main goal of the IV veto in the Gadolinium selection.

The accidental background is mainly due to γs from natural radioactivity. It is expected that high energy γs, from Thallium or Potassium decays, can reach into the ID from the IV, with a compton scattering. This kind of event is called

$\gamma - \gamma$. As accidental background is the main background in Hydrogen selection (cf. section 5.1.1.2), the rejection of $\gamma - \gamma$ events is the main goal of the IV veto in the Hydrogen selection.

The rejection of cosmic muons is the first goal of the Inner Veto detector. Most of the muons, including several stopping muons, are rejected by the IV (see section 5.1.2.1). However, in the Double Chooz analyses, the stopping muon background comes from the fact that there is a hole within the IV: the chimney hole. Muons can cross into the ID by this hole and not be tagged by the IV. Some of these stopping muons cross the corners of the chimney hole, depositing some energy within the IV. This low energy deposition, below the muon rejection threshold, can be used to tag the stopping muons, but the IV efficiency for these events is expected to be low.

FIGURE 5.13: Illustration of IBD in coincidence with IV events. On left is the schematization of accidental coincidence, and on right the correlated coincidence.

In addition to these different background events, the IV could also tag true IBDs occuring in coincidence with an energy deposition in the IV. This IV energy deposition can be accidental or correlated to the IBD. Accidental IV energy deposition can be due to fast neutron, γ interaction or PMTs radioactivity. Correlated IV energy deposition occurs when a γ from the positron annihilation or from the neutron capture escapes from the ID into the IV. This second kind of IV energy deposition is more likely to occur in the Hydrogen selection since the IBD reactions can occur closer to the IV, in the γ-catcher. FIG. 5.13 schematizes these two possibilities.

The IV veto cut was also designed to avoid the rejection of true IBD events. Its development for the Gadolinium and the Hydrogen analyses was one of my contributions during this thesis. In order to avoid the rejection of true IBD events, and to maximize the background rejection, we used the correlations between ID events and IV events. We used the timing correlation as well as the spatial correlation. This last feature was available through the NNIVReco algorithm (cf. section 2.7.2), which was developed before.

5.1.4.1 Rejection of IV internal radioactivity

The main source of accidental correlations between the ID and the IV was expected to come from the IV internal radioactivity due to the PMTs. This internal radioactivity is expected to induce low energy depositions in the IV. It should be possible to avoid ID-IV accidental correlations by requiring a minimal energy deposition in the IV. Two variables can be used to define a minimal energy deposition: the charge deposited in the IV and the number of hit IV PMTs. Both variables are expected to be correlated. The definition of the minimal energy deposition conditions was made with the Gadolinium selection.

In order to select internal radioactivity events, we search for events which can be interpreted as two simultaneous radioactive decays. Due to the low energy deposition, one can expect that internal radioactivity interactions will not trigger many PMTs. Therefore, we look for events where all the energy deposition is located in the top and bottom parts of the IV, without any hit lateral PMT.

In order to study this, we looked at the sum of the fractional charge in the top and bottom parts of the IV, as defined in Eq. 5.5. This equation returns 1 when there is no charge deposition in lateral IV PMTs. This allows to probe events which can be interpreted as two simultaneous radioactive decays, but it also probes events for which only PMT in the top or bottom parts of the IV are hit.

$$Q_{IV,Fractionnal}^{Top+Bottom} = \frac{Q_{IV}^{Top} + Q_{IV}^{Bottom}}{Q_{IV}^{Total}} \tag{5.5}$$

Fig. 5.14 shows the sum of the top and bottom fractional charges as a function of the total charge deposited in IV. Since we are looking for low energy events, this figure is showing only events with a total charge deposition in IV lower than 2000 a.u. ($\sim$ 1 MeV). A sub-sample of event with $Q_{IV}^{Top} > 0$ and $Q_{IV}^{Bottom} > 0$, is also shown in this figure. This analysis showed that most of the events with $Q_{IV,Fractionnal}^{Top+Bottom} = 1$ are low energy events ($Q_{IV} < 500$ a.u.).

In order to tune a cut on the IV charge, we studied the Inner Detector prompt energy spectrum of prompt signals in coincidence with an energy deposition in the IV. Different cut values were applied on the total charge deposited in the IV, from 100 to 500 arbitrary units. The different spectra are shown in Fig. 5.15. A peak between 2 and 3 MeV is visible on these spectra, it was interpreted as compatible with an IBD prompt signal. Since we aimed to avoid the rejection of true IBD events, the cut on the total charge deposited in the IV was tunned to make this peak disappear. The studies pointed out a cut value at 400 arbitrary units.

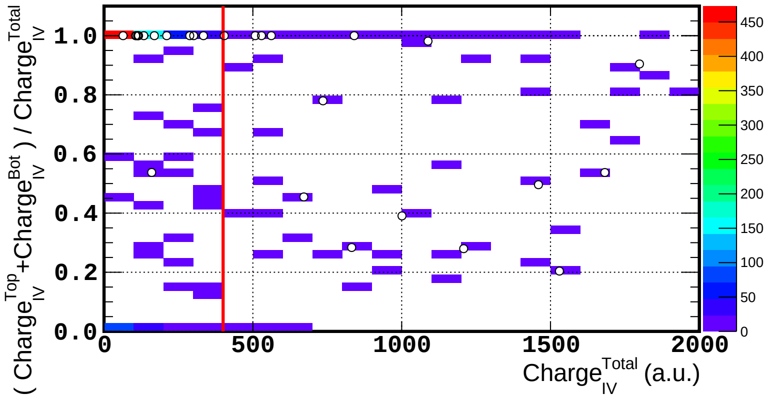


FIGURE 5.14: Correlation between the sum of IV top and IV bottom fractional charges as a function of the IV total charge. White dots are for events with $Q_{IV}^{Top} > 0$ and $Q_{IV}^{Bottom} > 0$.

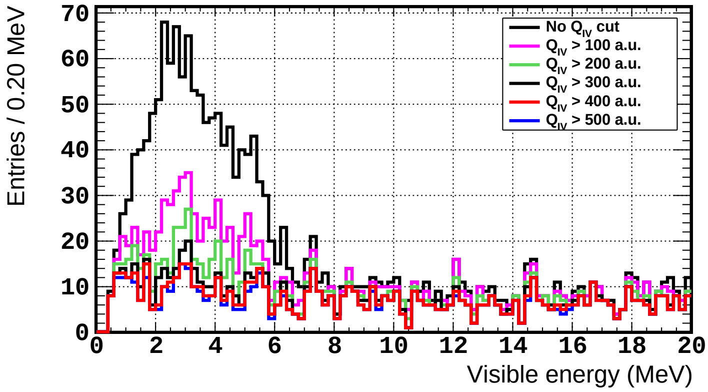


FIGURE 5.15: Energy spectrum of prompt signal occuring in coincidence with an energy deposition in the IV for different cut values on the total IV charge. Only prompt signals in space and time correlations with a delayed signal were selected.

FIG. 5.16 shows the number of hit PMTs for prompt signals in coincidence with an energy deposition in IV. From this figure, it is possible to see that there is only one hit PMT in the IV for most of the events. This figure also shows that the 400 a.u. cut rejects only events with less than 4 hit PMTs. Consequently, we demonstrate that making a cut on the number of hit PMTs is similar to the application of a cut on the charge deposition. Thus, we decided to apply only a minimal cut on the number of hit PMTs. As NNIVReco cannot reconstruct events with only one hit PMT in the IV, the minimal condition was decided as to have, at least, 2 hit PMTs in the IV.

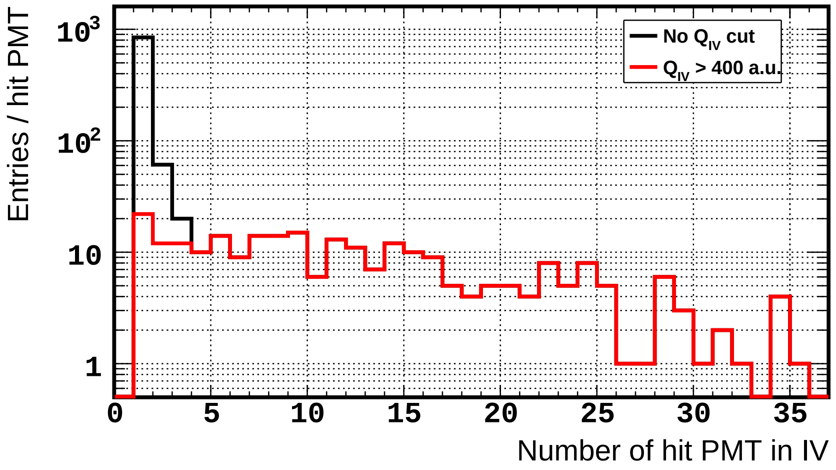


FIGURE 5.16: Distribution of the number of hit PMTs in the IV for prompt signals in coincidence with an energy deposition in IV. Only prompt signals in space and time correlations with a delayed signal were selected.

Therefore, in order to avoid the application of the IV veto on the IV internal radioactivity events. The IV veto was applied only on events which selected by the two following conditions:

- $Q_{IV} > 400$ a.u.
- Number of hit IV PMTs > 1

These two conditions are called "minimal IV tagging conditions".

5.1.4.2 The IV veto in the Gadolinium analysis

In the Gadolinium analysis, the IV veto cut aims to reject the correlated background. Accidental background tagged by the IV is expected to be low. In order to avoid $\overline{\nu}_e$ rejection, the timing and spatial correlations are used.

To determine the timing and spatial correlations, a sample of prompt signals satisfying the minimal IV tag condition was selected. Only prompt signals in time correlation with a delayed signal were selected. The prompt-delayed space correlation cut was defined after the definition of the IV veto. For prompt signals with ID-IV correlations, a peak is expected in the ID-IV timing and spatial correlation

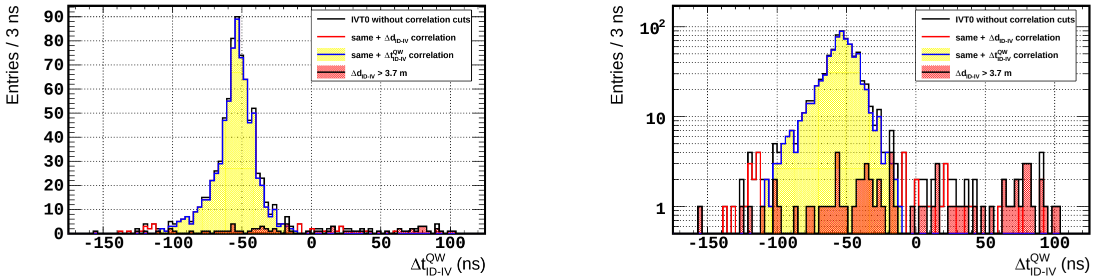


FIGURE 5.17: Charge-weighted time difference between ID energy deposition and IV energy deposition for prompt signal, in linear scale (*left*) and in logarithmic scale (*right*). Only prompt signals reaching the minimal IV tag conditions and in time correlation with a delayed signal were selected. The black distribution is the distribution of all prompt events, in red, the ID-IV space correlation cut, defined below, is applied and, in blue filled in yellow, both ID-IV space and time correlation cuts are applied. The black distribution filled in red shows the events rejected by the ID-IV space correlation cut, defined below.

distributions, whereas a flat behavior is expected for prompt signals in accidental correlations (i.e. without ID-IV correlations).

FIG. 5.17 shows the difference between the pulse time in the IV and the ID for the selected sample. The time definition used is the charge-weighted time, which is computed as shown in the following equation:

$$t_{QW} = \frac{\sum_i^{IV\,PMT} Q_i \times t_{start,\,i}}{Q_{IV}^{Total}} \tag{5.6}$$

where $t_{start,\,i}$ is the pulse start time of the PMT i provided by RecoPulse (see section 3.1). This pulse time definition was shown to be the most robust [171]. From this figure, we observed the expected peak, which was interpreted as due to the simultaneous fast neutron interactions in the IV and in the ID. The $\gamma - \gamma$ and the stopping muon contributions were expected to be outclassed by the fast neutron contribution for the Gadolinium selection. A flat component can also be observed at the extremums of the distribution, with an amplitude of ~ 2 events per bin. In order to determine the ID-IV timing correlation cut, we looked at the intersection between the peak and the flat component. These intersection was found to be about $[-114, -13]$ ns. We also performed a fit of the distribution, with $gaussian + flat\,line$ function, which gave a time range of $[-100.53, -6.21]$ ns at 4σ. However, the result of the fit cannot be used directly: due to the different ID-IV correlations the peak is not expected to be a gaussian, therefore the gaussian fit cannot provide a fully correct estimation. The ID-IV time correlation cut was determined to be $[-110, -10]$ ns. A cross-check was performed after the determination of the ID-IV space correlation cut and confirmed this value.

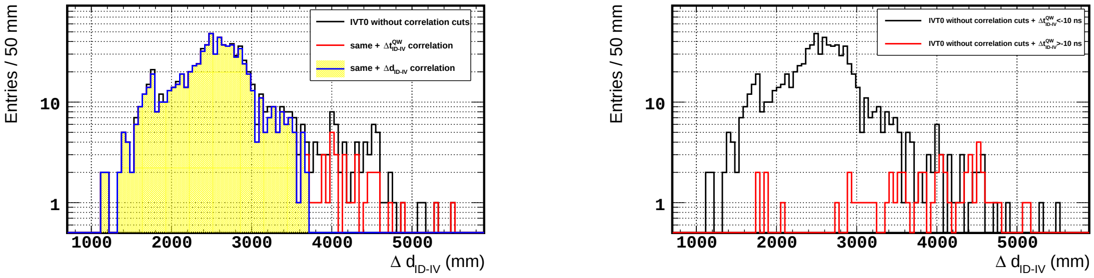


FIGURE 5.18: Distance between ID energy deposition and IV energy deposition for prompt signal in logarithmic scale. Only prompt signals satisfying the minimal IV tag conditions and in time correlation with a delayed signal were selected. On the *left* figure, the black distribution is the distribution of all prompt events, in red, the ID-IV time correlation cut is applied and, in blue filled in yellow, both ID-IV space and time correlation cuts are applied. On the *right* figure, the black distribution is the one of all prompt events with $\Delta t^{QW}_{ID-IV} < -10$ ns. The red distribution is the one of all prompt events with $\Delta t^{QW}_{ID-IV} > -10$ ns.

FIG. 5.18 shows the distance between the reconstructed interaction vertex in ID (reconstructed by RecoBAMA) and the one in IV (reconstructed by NNIVReco) for the selected sample. A structure, which was interpreted as a combination of several distorted landau distributions, is observed. We interpreted the distortion of the landau distributions as coming from geometric effects due to the ID and IV cylindric shapes, and/or to the different contribution to the ID-IV correlations (i.e. fast neutrons, stopping muons and $\gamma - \gamma$ events). On the right figure in FIG. 5.18, the selected sample is separated between ID-IV time correlated and time uncorrelated events, using the ID-IV time correlation defined above. The uncorrelated sample is consistent with a flat component. We used it to determine the ID-IV space correlation cut with a similar method than the one used for the for the ID-IV time correlation cut. We looked at the intersection between the flat component and the peak. A intersection point at ~ 3.7 m was found.

Here, a fit was not possible due to the complex shape of the distribution. However, using this value as a cut, the Δt^{QW}_{ID-IV} distribution of events with $\Delta d_{ID-IV} > 3.7$ m is consistent with a flat component as it can be see in FIG. 5.17. This was expected for events without correlation between ID and IV.

Therefore, we defined the IV veto to reject the prompt signals in coincidence with an activation of the IV, which are selected by the following conditions:

- Minimal IV tagging condition

- $\Delta t^{QW}_{ID-IV} \in [-110, -10]$ ns
- $\Delta d_{ID-IV} < 3.7$ m

Preliminary investigations were performed to apply the IV veto on delayed signals. However they suggested that it was likely to reject $\overline{\nu}_e$ events due to γ from the Gadolinium neutron captures escaping from the ID into the IV. Consequently, the IV veto was not applied on the delayed signals.

Inefficiency of the IV veto

The **inefficiency of the IV veto** is defined as the fraction of true IBD events rejected by the cut:

$$\overline{\varepsilon}_{IVV} = \frac{N_{IVV\,rejectedIBD}}{N_{IBD}} \tag{5.7}$$

where $N_{IVV\,rejectedIBD}$ is the number of true IBD events rejected by the IV veto and N_{IBD} is the number of IBD events selected in the final version of the Gadolinium analysis. For the DC-III (Gd-n) analysis $N_{IBD} = 17358$ events.

Two methods were developed to computed the IV veto inefficiency. The first one consists in the selection of a sample with only accidental correlations between the ID and the IV energy deposition. The second one consists in the study of the shape of the Δt^{QW}_{ID-IV} distribution to probe hints of the IV veto inefficiency.

For the first method, in order to build a sample with only accidental correlations between the ID and the IV, we decided to use the sample of random events. These events are registered by the detector electronics at a fixed rate without considering the amount of energy deposited in the detector. It is possible to have an energy deposition in the Inner Veto in coincidence with the random trigger. In order to build the accidental correlations, we took these IV signals from the random triggers sample. For the ID signal, we took the Gadolinium analysis sample of IBD, without IV veto applied.

The number of ID signal selected was of 17 525 events. The number of IV signal selected was of 824 188 events. For technical reason, it was not possible to associate each IV signal to each ID signal. Then, in order to have a significant statistics, we decided to buid the accidental correlations sample by associating to each ID signal 100 IV signals selected randomly in the IV signal sample.

FIG. 5.19 shows the ID-IV time and space correlations between ID and IV signals of the accidental correlations sample. It is visible that these correlations are very different than the correlations seen for the events rejected by IV veto in the Gadolinium selection. A flat behavior can be observed on both figures, it was interpreted as a demonstration of the fact that the ID-IV correlations were accidentals in this sample.

The IV veto rejects $5\,476 \pm 42$ events in the accidental correlations sample. We can determined the following ratio:

$$\frac{N_{IVV\,rejected\,AccCorr}}{N_{AccCorr}} = \frac{5476}{17525 \times 100} \tag{5.8}$$

where $N_{IVV\,rejected\,AccCorr}$ is the number of event rejected by the IV veto in the sample and $N_{AccCorr}$ is the total number of event in the sample. Since we selected only random IV signal with $Q_{IV} > 0$ a.u., this ratio can be interpreted as the probability of the IV veto to accidentaly reject an event with $Q_{IV} > 0$ a.u. (i.e. to reject an event which has accidental ID-IV correlations and $Q_{IV} > 0$ a.u.). Therefore, knowing the ratio of event in the IBD selection with $Q_{IV} > 0$ a.u., we were able to determine the probability to accidentaly reject an IBD event with the IV veto:

$$P = \frac{N_{IVV\,rejected\,AccCorr}}{N_{AccCorr}} \times \frac{N_{IBD\,with\,Q_{IV}>0}}{N_{IBD}} \tag{5.9}$$

This probability is interpreted as the IVV inefficiency, with:

$$N_{IVV\,rejectedIBD} = \frac{N_{IVV\,rejected\,AccCorr}}{N_{AccCorr}} \times N_{IBD\,with\,Q_{IV}>0} \tag{5.10}$$

Therefore, we have:

$$\overline{\varepsilon}_{IVV} = \frac{N_{IVV\,rejected\,AccCorr}}{N_{AccCorr}} \times \frac{N_{IBD\,with\,Q_{IV}>0}}{N_{IBD}} \tag{5.11}$$

We measured $N_{IBD\,with\,Q_{IV}>0} = 1140$, then the inefficiency value was then computed as:

$$\overline{\varepsilon}_{IVV}^{AccCorr} = 0.0205 \pm 0.0002\% \tag{5.12}$$

The uncertainty represents only the statistical uncertainty. The process was repeated several times to avoid a bias due to thee random selection of the IV signals, consistent results, within error bars, were found.

The second method which is to compute the IV veto inefficiency is a shape method. We used the Δt_{ID-IV} distribution of the IBD candidates selected by the Gadolinium selection (with no additionnal cut, and including the IV veto). Then, we probed the excesses and the holes in the distribution. We interpreted the presence of holes as the signature of an inefficiency: the IV veto rejects to many events and therefore rejects true IBDs. The presence of excesses was interpreted as the signature of a remaining background not rejected by the IV veto. Therefore, the hypothesis is that a flat behavior is expected if the inefficiency is ~ 0% and all the background events with ID-IV correlations are rejected. Note that this is not

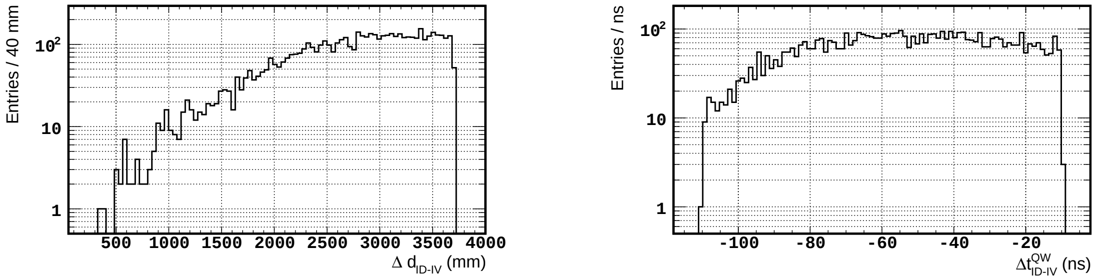
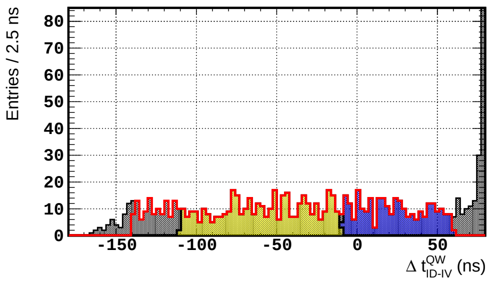


FIGURE 5.19: Space (*left*) and time (*right* correlations between ID and IV signals for events rejected by IV veto within the accidental correlations sample. IV veto rejects 5476 ± 42 events.

fully accurate due to the possibility to have IBD reactions with ID-IV correlations (as explained above in section 5.1.4).

FIG. 5.20 shows the Δt_{ID-IV} distribution of the Gadolinium selection IBD candidates.



Sample	Entries	Ratio
Uncorrelated $[-10, 60]$ ns	273	3.90 ± 0.24
IV veto $[-110, -10]$ ns	394	3.94 ± 0.20
Total $[-140, 60]$ ns	784	3.92 ± 0.14

FIGURE 5.20: *Left*: ID-IV time difference distribution for IBD candidates after application of all cuts. The different samples (red, blue and yellow) are used to compute the IV veto inefficiency. *Right*: Summary of the different sample from the time distribution.

In FIG. 5.20, no clear excess or hole was observed in the region where the IV veto is applied (i.e. $[-110, -10]$ ns). Therefore, three samples were studied in order to determine the inefficiency with a statistical method. The idea is to compare the number of remaining events in the region where the IV veto is applied to the number of events in a region where the IV veto is not applied and where no ID-IV correlations due to background are expected.

Note that there is time shift between the ID and the IV, if $\Delta t^{QW}_{ID-IV} = 0$, it does not mean that the ID signal is simultaneous with the IV signal. If $\Delta t^{QW}_{ID-IV} = 0$

the ID signal is likely to occur before the IV signal. However, the "simultaneous time" between the ID and the IV signals is difficult to set with precision as we use a charge-weighted time.

The first sample, called *uncorrelated sample*, is taken in the region where no ID-IV correlations due to background are expected. It was taken in $[-10, 60]$ ns. The second sample, called *IV veto sample*, is taken in the region the the IV veto is applied, therefore: $[-110, -10]$ ns. The third sample, called *total sample*, includes the two other samples. It was extended on the left to probe a part of the distribution where the IV veto is not applied. It was taken in $[-140, 60]$ ns. For each sample, the following ratio was computed:

$$ratio = \frac{N_{events}}{size_{sample}} \tag{5.13}$$

where N_{events} is the number of IBD candidates in the sample, and $size_{sample}$ is the size, in ns, of the sample. In our interpretation, if the IV veto is perfect (and the statistics are infinite) the ratios would be the same for all samples. Using the hypothesis formulated above we can have two other cases: If the *IV veto sample* ratio is higher than the two other ratios, it would be interpreted as consistent with a $\sim 0\%$ inefficiency but with remaining background events with ID-IV correlations. If the *IV veto sample* ratio is lower than the two other ratios it would be interpreted as consistent with a $> 0\%$ inefficiency but with no remaining background events with ID-IV correlations.

Here, the ratio of the *IV veto sample* is 3.94 ± 0.20, whereas it is 3.90 ± 0.24 for the *uncorrelated sample* and 3.92 ± 0.14 for the *total sample*. The three ratios are consistent within error bars. Therefore, we interpreted this as consistant with $\sim 0\%$ inefficiency. Consequently, the statistical uncertainty of the total sample was used as an upper limit for the IV veto inefficiency:

$$\overline{\varepsilon}_{IVV}^{Exces} < \frac{\left(\frac{N_{total\,sample}}{\sqrt{N_{total\,sample}}}\right)}{N_{IBD}} = \frac{\sqrt{N_{total\,sample}}}{N_{IBD}} \tag{5.14}$$

where $N_{total\,sample}$ is the number of events in the *full sample*. The table in Fig. 5.20 details the different values computed for this analysis. We computed the following limit of the IV veto inefficiency:

$$\overline{\varepsilon}_{IVV}^{Exces} < 0.16\% \tag{5.15}$$

This limit is consistent with the inefficiency value computed with the accidental correlations method. As the accidental correlations method may underestimate the inefficiency. The main method to compute the IV veto inefficiency was decided to be the the Δt_{ID-IV}^{QW} shape method.

Efficiency of the IV veto

The **efficiency of the IV veto** is defined as the fraction background events rejected by the IV veto:

$$\bar{\varepsilon}_{IVV}^{Excess} = \frac{N_{IVV\,rejectedBG}}{N_{BG}} \tag{5.16}$$

where $N_{IVV\,rejectedBG}$ is the number of background events rejected by the IV veto and N_{IBD} is the total number of background events. The number of background events depends of the background studied and of the method used to select them. For technical reasons, the efficiency is computed seperatly for each background rejected. The rejection efficiency of the correlated background was not estimated due to the method used to measure the correlated background (see chapter 6.

In order to validate that IV veto rejects correlated background, we studied the overlap between the sample of events rejected by IV veto and the one rejected by other vetos aiming to reject the correlated background. The OV veto and the FV veto also aim to reject correlated background and were used in this analysis. A sample of events was build, with all events rejected by at least of of the three vetos and with $E_{vis} > 12$ MeV or $\Delta t_{prompt-delayed} < 10\ \mu$s. These two cuts allows to select the correlated background and are explained in the chapter 6.

Tab. 5.1 shows the different number found by the analysis. The IV veto rejects 24% of all the events rejected by at least one of the three vetos. 7% of all the events rejected by at least one of the three vetos are only rejected by the IV veto. We interpreted this result as a demonstration that the IV veto rejects correlated background.

Veto	**Rejected events**	**Uncorrelated events**
IV veto	24%	7%
OV veto	62%	7%
FV veto	71%	19%
Total	90%	33%

Table 5.1: Check of the IV veto rejection consistency. The number of rejected events by a given veto is given as the fraction of the total number of events in the sample of event rejected by at least one of the three vetos, and which are selected by either $E_{vis} > 12$ MeV condition, either $\Delta t_{prompt-delayed} < 10\ \mu$s condition. The number of uncorrelated events are the fraction of events rejected only by the cut.

The IV veto is also able to reject accidental background, a study was performed to measure the IV veto rejection efficiency of the accidental background. FIG. 5.21 shows the prompt energy spectrum of the accidental background before and after the application of the IV veto and for accidental events rejected by the IV veto. The selection of the accidental background is detailed in section 5.1.1.2. We measured that before the application of the IV veto, there was 80 783 events in the accidental sample. The IV veto rejects 10 515±37 accidental events. Therefore, we computed the following IV veto rejection efficiency of the accidental background:

$$\varepsilon_{IVV}^{Acc} = 13.02 \pm 0.05(stat)\% \tag{5.17}$$

This high rejection efficiency was not expected when the IV veto was build. It was measured after the development of the IV veto for the Hydrogen analysis.

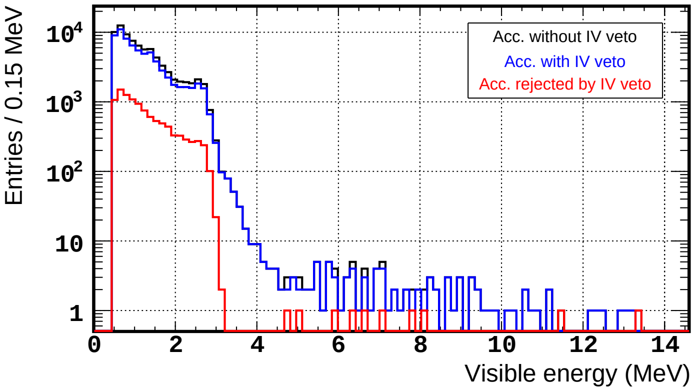


FIGURE 5.21: Accidental prompt energy spectrum before application of the IV veto (black), after application of the IV veto (blue) and for events rejected by the IV veto (red). All the Gadolinium IBD selection cuts were applied on the accidental sample.

Summary

The IV veto was developed for the Gadolinium selection in order to reject fast neutrons. It is applied on prompt candidates which satisfy some minimal conditions on the energy deposition in the IV. In order to reject only prompt candidates with correlations between ID and IV, the time and space correlations between the ID and the IV are used.

We developed two methods to measure the inefficiency of the IV veto, the first one gave an inefficiency of $0.0205 \pm 0.0002\%$, which was confirmed by second method which found an inefficiency $< 0.16\%$. The consistency of the IV veto rejection of correlated background was cross-checked using the FV veto and OV veto, and led to results consistent witht the hypothesis of correlated background rejection.

In addtion, the efficiency of the IV veto accidental background rejection was estimated, and found to be $13.02 \pm 0.05(stat)\%$. This result was not expected when we start the development of the IV veto for the Gadolinium selection.

The IV veto was not applied on the delayed signals due to the risk of true IBDs rejection. However, the developments of the Hydrogen selection version of the IV veto, detailed in the next section, allow to consider to apply the IV veto on the delayed signals in the future Gadolinium analysis.

5.1.4.3 The IV veto in the Hydrogen analysis

The Hydrogen analysis suffers from accidental background in the delayed signal, due to the lower delayed energy window. One of the most dangerous accidental background here is due ^{206}Tl, which decays with a 2.6 MeV γ emission. By compton scattering, a γ from ^{206}Tl can enter the detector through the IV and deposits an energy close to the H-n capture peak in the ID. Therefore, we needed to apply the IV veto to both prompt and delayed candidates. In addition, as many IBD candidates of the Hydrogen selection occur in γ-catcher, we cannot apply the same IV veto as in the Gadolinium analysis. As the time and space correlations between the ID and the IV could be different for events interacting in the γ-catcher. In addition, for the same reason, as IBD can occur closer to the IV, the risk to have IBD with correlations between the ID and the IV signals increase.

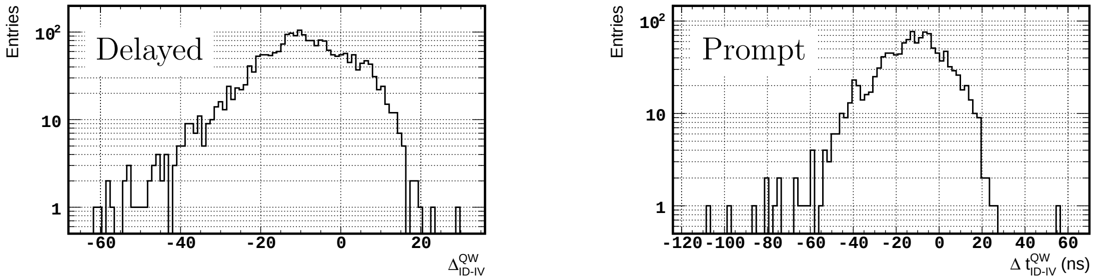


FIGURE 5.22: ID-IV time difference distribution for MC IBD delayed signal (*left*) and prompt signal (*right*) selected by the minimal IV tag conditions.

In order to measure and avoid the possibility to reject true IBD with a γ escaping from the ID to interact in the IV, we performed a study with the $\overline{\nu}_e$ MC. FIG. 5.22 (*left*) shows the Δt_{ID-IV} distribution for H-n MC delayed candidates with minimal IV tagging conditions. It highlights that γ from the neutron capture on Hydrogen can reach the IV. The main part of this kind of events is due to neutron captures occuring close to the external γ-catcher border (i.e. the border between the γ-catcher and the buffer). We also discovered that γ from positron annihilation can reach the IV, as shown on the right figure in FIG. 5.22. This figure shows the Δt_{ID-IV} distribution for H-n MC prompt candidates with minimal IV tagging conditions. This was not expected as γs from positron annihilation are low energy γs.

In order to avoid the $\overline{\nu}_e$ rejection, we tunned the Δt_{ID-IV} to reach an inefficiency of the IV veto of $\sim 0.01\%$ with the MC simulation. Such inefficiency takes only into account the IBDs with genuine correlations between the ID and the IV, it is only a component of the full IV veto inefficiency defined in the Gadolinium analysis (cf. EQ. 5.7). 0.01% was chosen as preliminary studies were indicated an IV veto inefficiency of $\sim 0.1\%$.

We found a 0.009% inefficiency for the delayed candidates if all events with $\Delta t_{ID-IV} < -21$ ns were rejected. For the prompt candidates, we found a 0.010% inefficiency for the same upper limit. For simplicity, we round it up to -20 ns. This leads to an inefficiency of 0.011% for the delayed candidate and 0.010% for the prompt candidate.

In order to define the ID-IV space and time correlation cuts, we applied the same method used for the Gadolinium analysis. We select a sample of prompt and delayed event selected by the minimal IV tag conditions, with only the ANN cut applied in the selection.

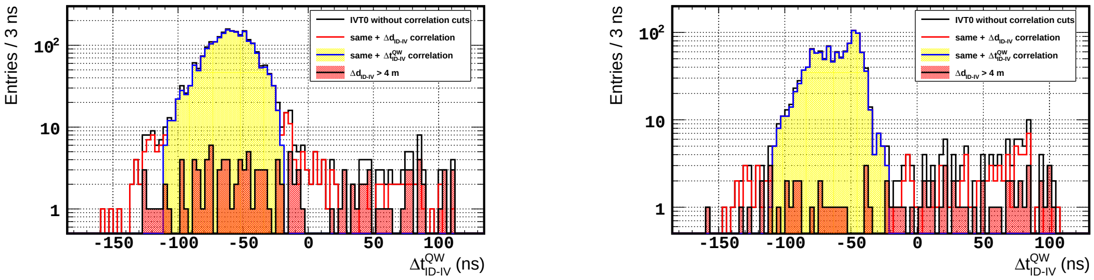


FIGURE 5.23: Charge-weighted time difference between ID energy deposition and IV energy deposition for prompt signal (*left*) and delayed signal (*right*). Only prompt and delayed signals satisfying the minimal IV tag conditions were selected. The ANN cut was also applied. The black distribution is the distribution of all prompt or delayed events, in red, the ID-IV space correlation cut, defined below, is applied and, in blue filled in yellow, both ID-IV space and time correlation cuts are applied. The black distribution filled in red shows the events rejected by the ID-IV space correlation cut, defined below.

FIG. 5.23 shows the difference between the pulse time in the IV and the ID for prompt candidate and delayed candidate. The charge-weighted time definition was used (as defined in EQ. 5.6). In these figures a peak, interpreted as due to the ID-IV correlation, is visible, as well as a flat component, which is expected to be due to the events without genuine correlations between the ID and the IV signals. Note that the distribution for the delayed signal presents two peaks, it is discussed briefly below.

From these figures, as for the Gadolinium analysis, we looked at the intersection between the flat component and the peak in order to define a range for the ID-IV time correlation. We found a time range of $[-120, -10]$ ns for the prompt signal and a time range of $[-110, -30]$ ns for the delayed signal. A $gaussian+flat\,line$ fit was also performed for the prompt signal, it led to a 3σ range of $[-119.0,\ -8.9]$ ns. Due to the distorted shape of the delayed signal δt^{QW}_{ID-IV} distribution, no fit was performed for the delayed signal. Pratically, we wanted to have the same cuts for both prompt and delayed IV vetos.

As we already defined a maximal upper of this time range to be at -20 ns, we investigated the impact of the upper limit with a value of -30 ns and of -20 ns. We found that the impact on the inefficiency was almost negligible between both configurations ($< 0.01\%$) [172, 173, 174]. Therefore, we decided to use the higher value: -20 ns. For the lower limit, we studied the impact of an lower limit at -120 ns or at -110 ns and found that the impact on the rejection efficiency of the IV veto was negligible (less than 5 more accidental events rejected). Therefore, we decided to keep -110 ns as for the Gadolinium definition, and we used the following time range: $[-110, -20]$ ns.

FIG. 5.24 shows the distance between the reconstructed interaction vertex in the ID (reconstructed by RecoBAMA) and the one in the IV (reconstructed by NNIVReco) for the prompt candidates. FIG. 5.25 is the same but for the delayed candidates. From these distributions, using the same method than in the Gadolinium analysis, we found an intersection between the flat component and the peak at 4 m for the prompt signals and at 3.8 m for the delayed signals. As for the determination of the ID-IV time correlation, we studied the impact of a cut at 3.8 m and at 4 m. The impact on the inefficiency was shown to be negligible ($< 0.01\%$). Therefore, we decide to use the higher cut.

Therefore, we defined the Hydrogen selection IV veto to reject prompt and delayed signals in coincidence with an activation of the IV, which satisfy with the following conditions:

- Minimal IV tagging conditions ($Q_{IV} > 400$ a.u. and IV PMTs multiplicity ≥ 2)
- $\Delta t^{QW}_{ID-IV} \in [-110,\ -20]$ ns

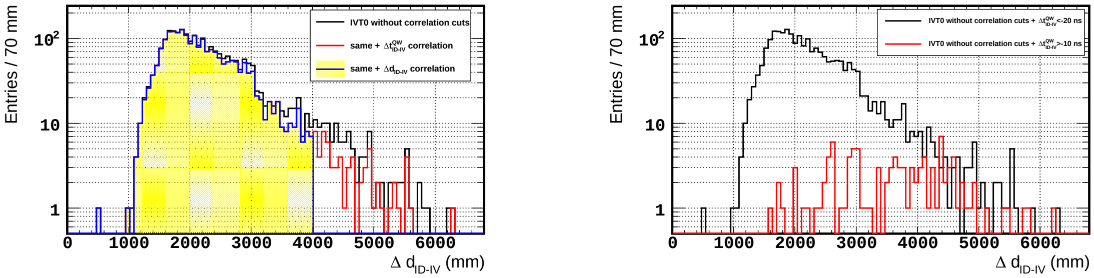


FIGURE 5.24: Distance between ID energy deposition and IV energy deposition for the prompt signals in logarithmic scale. Only prompt signals satisfying the minimal IV tag conditions were selected. The ANN cut was also applied. On the *left* figure, the black distribution is the distribution of all prompt events, in red, the ID-IV time correlation cut is applied and, in blue filled in yellow, both ID-IV space and time correlation cuts are applied. On the *right* figure, the black distribution is the one of all prompt events with $\Delta t^{QW}_{ID-IV} < -20$ ns. The red distribution is the one of all prompt events with $\Delta t^{QW}_{ID-IV} > -10$ ns.

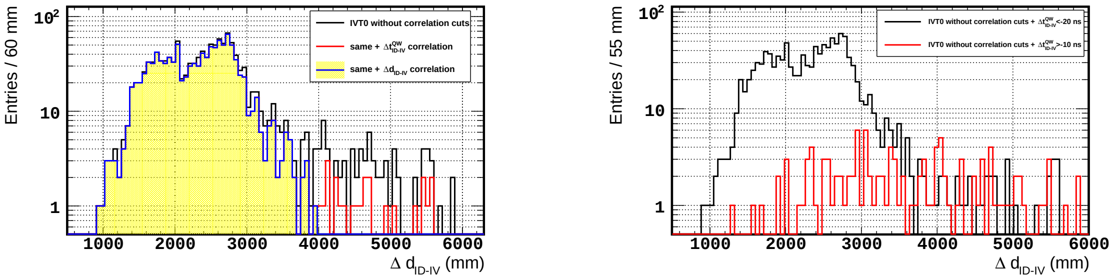


FIGURE 5.25: Distance between ID energy deposition and IV energy deposition for the delayed signals in logarithmic scale. Only delayed signals satisfying the minimal IV tag conditions were selected. The ANN cut was also applied. On the *left* figure, the black distribution is the distribution of all delayed events, in red, the ID-IV time correlation cut is applied and, in blue filled in yellow, both ID-IV space and time correlation cuts are applied. On the *right* figure, the black distribution is the one of all delayed events with $\Delta t^{QW}_{ID-IV} < -20$ ns. The red distribution is the one of all delayed events with $\Delta t^{QW}_{ID-IV} > -10$ ns.

- $\Delta d_{ID-IV} < 4$ m

Inefficiency

The computation of IV veto inefficiency for Hydrogen selection was performed in the same way as for the Gadolinium analysis.

For the accidental correlations method, we build a sample with 34 469 ID signals (all cuts applied except the IV veto[1]) and 824 188 IV signals. As for the Gadolinium analysis, we associated each ID signals to 100 IV signals randomly selected. We found that the IV veto reject 9 924 ± 54 events when applied on the prompt signals and 10 033 ± 54 events when applied on the delayed signals.

FIG. 5.26 and FIG. 5.27 show, respectively, the space and the time correlations of events rejected by the IV veto in the accidental correlations sample, for both prompt and delayed signals.

For the Hydrogen analysis we have $N_{IBD} = 31\,898$ events and we measured $N^{Prompt}_{IBD\,with\,Q_{IV}>0} = 2890$ events (all cuts applied except the IV veto applied on the prompt signals) and $N^{Delayed}_{IBD\,with\,Q_{IV}>0} = 2207$ events (all cuts applied except the IV veto applied on the delayed signals). Therefore, we measured the following IV veto inefficiencies with EQ. 5.11:

$$\bar{\varepsilon}^{AccCorr}_{IVV,\,prompt} = 0.0261 \pm 0.0002\,(stat)\% \tag{5.18}$$

$$\bar{\varepsilon}^{AccCorr}_{IVV,\,delayed} = 0.0201 \pm 0.0002\,(stat)\% \tag{5.19}$$

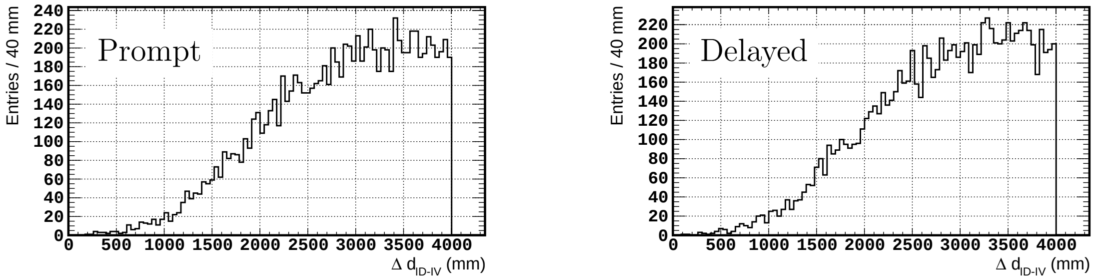


FIGURE 5.26: ID-IV space correlation for events from the accidental correlations sample rejected by the prompt IV veto (*left*) and the delayed IV veto (*right*). The IV veto rejects 9 924 ± 54 events when applied on the prompt signals, and 10 033 ± 54 events when applied on the delayed signals.

As for Gadolinium selection, we used the excess method to compute an inefficiency upper limit. FIG. 5.28 and FIG. 5.29 show the Δt_{ID-IV} distributions used

---

[1] The prompt signal energy window was $E_{vis} \in [0.7, 20]$ MeV, as it was shifted latter to $[1, 20]$ MeV. This should be negligible on this analysis.

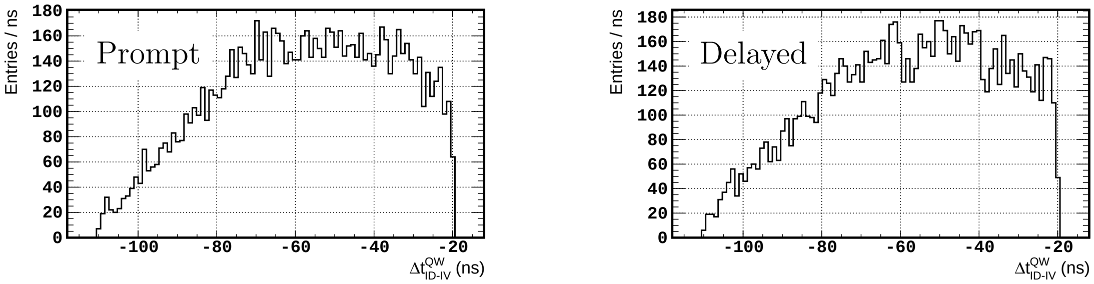


FIGURE 5.27: ID-IV time correlation for events from the accidental correlations sample rejected by the prompt IV veto (*left*) and the delayed IV veto (*right*). The IV veto rejects $9\,924 \pm 54$ events when applied on the prompt signals, and $10\,033 \pm 54$ events when applied on the delayed signals.

to estimate the inefficiencies and the tables show the different number computed in this analysis for, respectively, the IV veto applied on the prompt signals and the IV veto applied on the delayed signals. No clear excesses or holes were observed in both Δt^{QW}_{ID-IV} distributions. The ratio computed for the *IV veto sample* was seen to be far higher than the two other ratios: about 1 event/bin higher for the IV veto applied on the prompt signals and 0.4 event/bin higher for the IV veto applied on the delayed signals. This is interpreted as an indication that the IV veto is missing a part of the background events with ID-IV correlations. We suspect that the IV veto is missing $\gamma - \gamma$ events which deposit a small amount of energy in the IV (i.e. with $Q_{IV} < 400$ a.u.). Using EQ. 5.14, we compute an upper limit for the IV veto:

$$\overline{\varepsilon}^{Excess}_{IVV,prompt} < 0.123\% \tag{5.20}$$

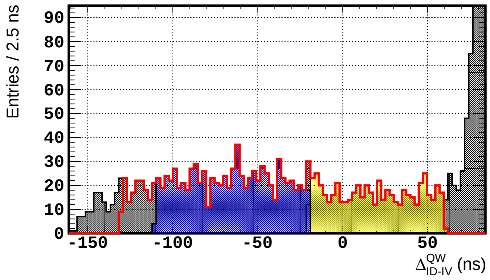
$$\overline{\varepsilon}^{Excess}_{IVV,delayed} < 0.116\% \tag{5.21}$$

Both methods give consistent results for the two IV veto.

With both methods, we were not able to measure the inefficiency of the full IV veto and we separate it in the prompt IV veto (i.e. the IV veto applied on the prompt signals) and the delayed IV veto (i.e. the IV veto applied on the delayed signals). If we assume no correlation between both applications, we can compute a combined inefficiency as:

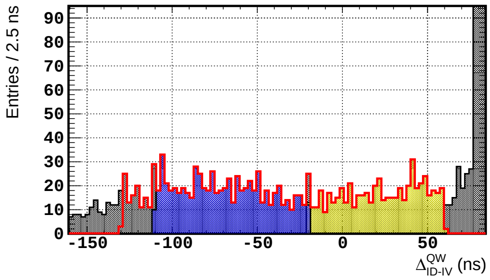
$$\overline{\varepsilon}^{Excess}_{IVV} < 0.169\% \tag{5.22}$$

This last value was used as the inefficiency due to the IV veto in the fit performed to extract θ_{13} (see section 5.3.3.1).



Sample	Entries	Ratio
Uncorrelated $[-20, 60]$ ns	558	6.97 ± 0.30
IV veto $[-110, -20]$ ns	834	9.27 ± 0.32
Total $[-130, 60]$ ns	1547	8.14 ± 0.21

FIGURE 5.28: *Left*: ID-IV time difference distribution for IBD prompt signal candidates after application of all cuts. The different samples (red, blue and yellow) are used to compute the IV veto inefficiency. *Right*: Summary of the different samples from the time distribution.



Sample	Entries	Ratio
Uncorrelated $[-20, 60]$ ns	555	6.94 ± 0.30
IV veto $[-110, -20]$ ns	692	7.69 ± 0.30
Total $[-130, 60]$ ns	1380	7.26 ± 0.20

FIGURE 5.29: *Left*: ID-IV time difference distribution for IBD delayed signal candidates after application of all cuts. The different samples (red, blue and yellow) are used to compute the IV veto inefficiency. *Right*: Summary of the different samples from the time distribution.

IV veto rejection efficiency

As for the Gadolinium analysis, we evaluated the accidental background rejection power of IV veto with the accidental background sample. The efficiency of the correlated background rejection was not estimated for the same reasons than in the Gadolinium analysis (see section 5.1.4.2).

FIG. 5.30 shows the spectrum of accidental events rejected by the IV veto. Before application of the IV veto, we measured 520 984 accidental events. The IV veto, applied on both prompt and delayed signals, allows to reject 140 837 accidental events. This leads to the following accidental background rejection efficiency:

$$\varepsilon_{IVV}^{Acc} = 27.03 \pm 0.04(stat)\% \tag{5.23}$$

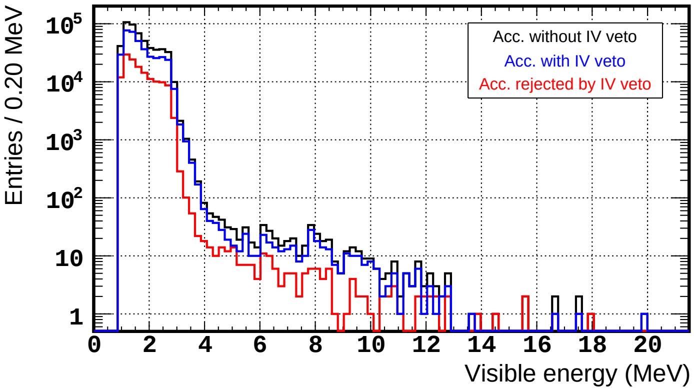


FIGURE 5.30: Accidental prompt energy spectrum before application of IV veto (black), after application of IV veto (blue) and for events rejected by IV veto (red). All Hydrogen IBD selection cuts were applied on the accidental sample.

The ANN cut in Hydrogen selection allows to reject 87.5% of accidental background. However, the ANN looks only at the delayed signal energy and at prompt-delayed time and space correlations and do not care about the ID-IV correlations. There should be no correlations between the ANN cut rejection method and the IV veto rejection method. Therefore, it is expected the accidental background rejection efficiency of the IV veto should be the same if we do not apply the ANN cut. In order to check this, we performed the estimation of the IV veto accidental background rejection efficiency after application of all other cuts of the Hydrogen selection, except the ANN cut. The accidental spectrum without application of the ANN cut is shown in FIG. 5.31. In this case, we measured a 14 480 490 accidental events before the application of the IV veto. The IV veto rejects here 3 727 087 accidental events. This leads to the following accidental background rejection efficiency:

$$\varepsilon_{IVV}^{no\,ANN} = 25.74 \pm 0.01(stat)\% \tag{5.24}$$

These two accidental background rejection efficiencies are consistent within 1.5%. We considered it as a good result to demonstrate the stability of ANN and IV veto rejection.

FIG. 5.32 shows the correlation between the energy deposited in the ID and the in IV for events rejected by IV veto. A triangular shape at low energy indicates an energy conservation between both volumes. This could be the signature of an

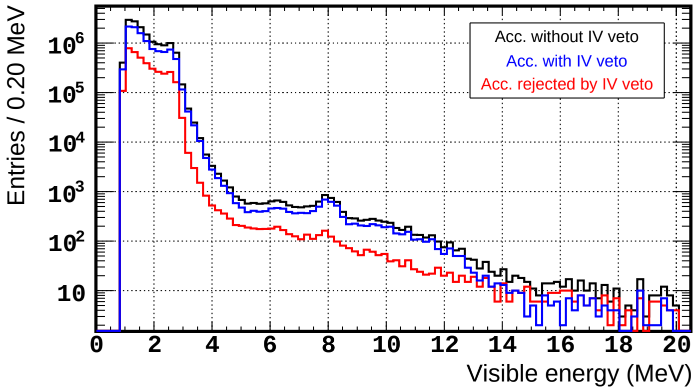


FIGURE 5.31: Accidental prompt energy spectrum before application of IV veto (black), after application of IV veto (blue) and for events rejected by IV veto (red). All Hydrogen IBD selection cuts, except ANN, were applied on the accidental sample.

interaction in the IV and in the ID by the same particle. We interpreted this feature as due to $\gamma - \gamma$ events. Fast neutrons and stopping muons events are not expected to show energy conversation between the IV and the ID as they have a wide energy spectrum and not a fixed energy like γ from radioactivity. This figure could be a demonstration that we are rejecting $\gamma - \gamma$ events with the IV veto.

Summary

We defined for the Hydrogen selection a new version of the IV veto, which is applied on both prompt and delayed candidates. This cut allowed to reject $\sim 27\%$ of the accidental background, the main background in the Hydrogen selection, with a negligible inefficiency ($< 0.2\%$). The IV veto is one of the main development of the new Hydrogen selection, contributing to the large accidental reduction and therefore, to the reduction of the θ_{13} background systematics for this particular analysis.

5.2 Remaining background estimations

While the selection cuts tried to reject most of the background, in order to keep a high signal efficiency a background-free analysis cannot be provided. The remaining background after selection cuts were estimated and used in order to extract θ_{13} in the final rate+shape and RRM fit (see section 5.5).

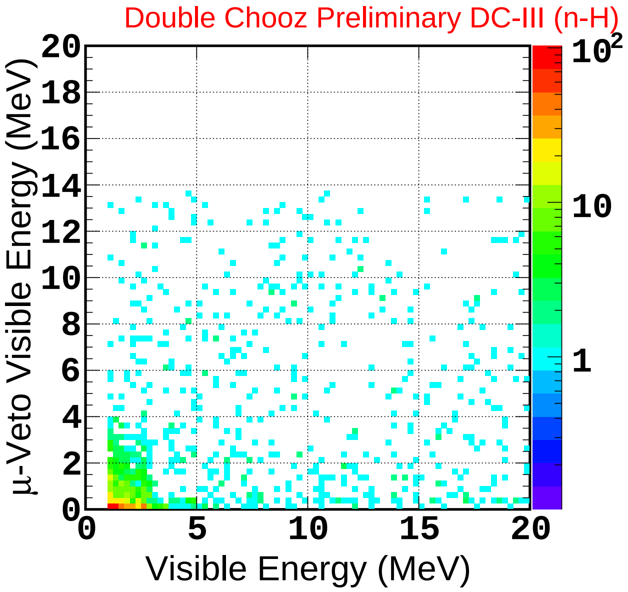


FIGURE 5.32: Energy correlation between ID energy deposition and IV energy deposition for events rejected by IV veto applied on the prompt and the delayed signals. The triangular shape at low energy indicates an energy conservation between both volume, which could be the signature of an interaction in the IV and in the ID by the same particle. It is interpreted as due to $\gamma - \gamma$ events. Events without energy conversation between IV and ID are expected to be fast neutrons and stopping muons.

5.2.1 Accidental background

The accidental background was estimated using an "Off-time" window method. It consists in the selection of a prompt and a delayed signals with random coincidences. In this method, all standard IBD selection cuts are applied but the time coincidence cut between prompt and delayed candidates. The coincidence time window is opened more than 1 s after the prompt signal, allowing to select random coincidence. For the Gadolinium analysis 2 000 time windows were opened in order to reach a significant statistic and improve the computation precision. For the Hydrogen analysis 200 time windows were opened. FIG. 5.33 is an illustration of the "Off-time" window method.

The resulted rate was then corrected for the dead times or the other similar time biases coming from IBD selection (Muon veto dead time, multiplicity cut, ^9Li and finite run length).

The remaining accidental rate was measured and the result is in EQ. 5.25 for Gadolinium analysis and in EQ. 5.26 for Hydrogen analysis. The prompt energy

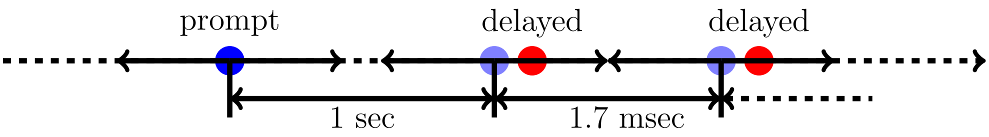


FIGURE 5.33: Illustration of the “Off-time” method to measure the accidental background. The blue point represents a prompt-like event. The red points represent delayed-like events. The multiplicity cut is applied around the virtual time position of the prompt-like event. This leads to an isolation time windows of 1.7 ms in the Hydrogen analysis, and of 0.8 ms in the Gadolinium analysis.

spectrum of the accidental background is shown in FIG. 5.34.

$$R^{Gd}_{Acc} = 0.0701 \pm 0.0003\,(stat) \pm 0.0054\,(syst)\,\text{events per day} \tag{5.25}$$

$$R^{H}_{Acc} = 4.334 \pm 0.011\,\text{events per day} \tag{5.26}$$

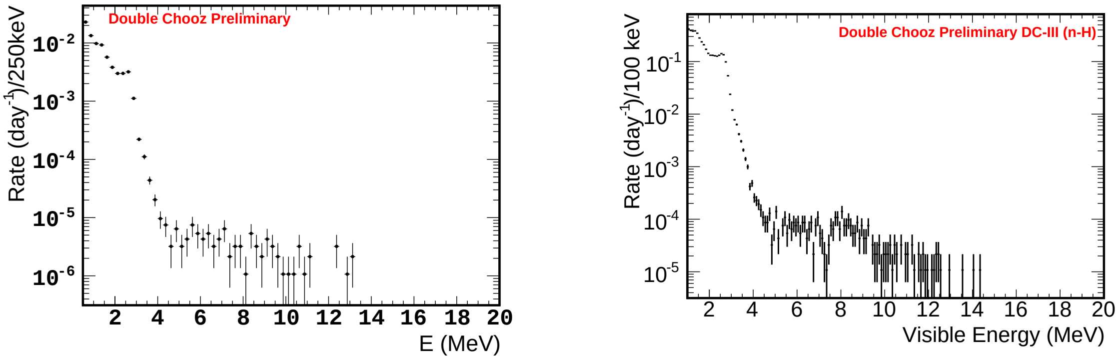


FIGURE 5.34: Prompt energy spectrum of the accidental sample for the Gadolinium analysis (*left*) and for the Hydrogen analysis (*right*).

5.2.2 Cosmogenic radio-isotope background

The remaining contamination from the cosmogenic isotopes, mainly ^{9}Li, was measured from fits of the ΔT_{μ} distribution, i.e. the time interval between IBD candidates and the previous muon. The cosmogenic radio-isotope selection also applies cuts on the distance between the track of the previous muon and the prompt signal, as well as a condition on the muon energy [178].

The remaining cosmogenic rate was measured from the selected sample and the result is in Eq. 5.27 for Gadolinium analysis and in Eq. 5.28 for Hydrogen analysis. The background shape is the same for both analysis and is shown in Fig. 5.35.

The cosmogenic radio-isotope background is the most dangerous background in both analysis and dominates both background systematic budgets by more than 50% (see section 5.2.4). ^{12}B remaining background was also estimated but found to be negligible.

$$^{9}Li\ Rate_{Gd} = 0.97^{+0.41}_{-0.16}\,\text{events per day} \tag{5.27}$$

$$^{9}Li\ Rate_{H} = 0.95^{+0.57}_{-0.33}\,\text{events per day} \tag{5.28}$$

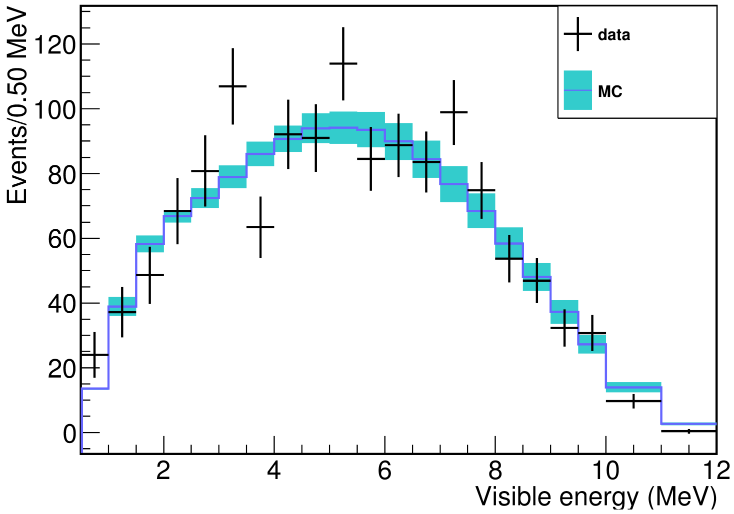


Figure 5.35: Cosmogenic isotope background energy spectrum.

5.2.3 Correlated background

Remaining fast neutron and stopping muon was also measured for both selection. Details can be found in the next chapter, chapter 6.

5.2.4 Summary

The different background rates and systematics due to the background are detailed in Tab. 5.2 for the Gadolinium selection and in Tab. 5.3 for the Hydrogen selection.

BG	rate (event/day)	BG/S (%)	δ(BG) (%)	suppression (w.r.t. Gd-II)
Accidental	0.070 ± 0.003	0.19	0.01	×3.7
Fast neutron stopped-μ	0.604 ± 0.051	1.62	0.1	×1.9
Cosmogenic radio-isotope	$0.97^{+0.41}_{-0.16}$	2.61	0.78	×1.3

Table 5.2: Summary of the remaining backgrounds for the Gadolinium analysis.

BG	rate (event/day)	BG/S (%)	δ(BG) (%)	suppression (w.r.t. H-II)
Accidental	4.334 ± 0.011	6.3	0.02	×16.9
Fast neutron stopped-μ	1.55 ± 0.15	2.2	0.23	×2.0
Cosmogenic radio-isotope	$0.95^{+0.57}_{-0.33}$	1.4	$^{+0.86}_{-0.50}$	×2.9

Table 5.3: Summary of the remaining backgrounds for the Hydrogen analysis.

These results highlight the major improvements of the present Gadolinium and Hydrogen analyses with respect to the previous Gadolinium and Hydrogen analyses. In the Gadolinium analysis, the accidental and the correlated background were highly reduced. In the Hydrogen analysis, all background rates were reduced by a factor > 2, and with the accidental background which was reduced by factor 16.9.

It is interesting to notice that the cosmogenic radio-isotope background rate is similar in both Gadolinium and Hydrogen analyses. The accidental background, despite to be the main background in rate in the Hydrogen analysis, is now negligible on the resolution of the analysis for both Gadolinium and Hydrogen analyses (with a $\delta(\mathrm{BG}) \leq 0.02$).

5.3 Selection efficiency and systematic uncertainties

The Gadolinium and the Hydrogen selections were optimized to provide a high IBD efficiency. However, the inefficiency induced by the cuts cannot be 0% and was then evaluated for each selection. In addition, other parameters, like the good understanding of the neutron detection, induced a systematic uncertainty on the θ_{13} measurement. Most of these uncertainties are different between the Gd and H analyses, due to the differences between both selections.

5.3.1 Reactor uncertainties

Source	**Uncertainty (%)**
Bugey-4 measurement	1.4
Fractional fission rate of each isotope	0.8
Thermal Power	0.5
IBD cross-section	0.2
Mean energy per fission	0.2
Distance to reactor core	< 0.1
Total	1.7

TABLE 5.4: Summary of reactor $\overline{\nu}_e$ rate uncertainties.

The reactor flux uncertainty is the only common systematic uncertainty between both analyses, as both are looking at the same $\overline{\nu}_e$ flux. It was computed to be 1.7%. TAB. 5.4 summarizes the different reactor flux uncertainties. Double Chooz uses the Bugey-4 measurement [Y. Declais et al., Phys. Lett. B338, 383 (1994).] to normalize its prediction and reduce the uncertainties (see section 2.5.5). This measurement was adapted to take into account the different fuel composition inside Bugey-4 and Chooz-B reactor core. Without Bugey-4 measurement, systematic would have been 2.8% [79].

5.3.2 Gadolinium analysis

5.3.2.1 IBD inefficiency

The IBD inefficiency due to the selection cuts was estimated using different methods. Results can be find in the Tab. 5.5. The computation of IV veto inefficiency is described in section 5.1.4. These inefficiencies were applied as corrections to the MC simulation.

Source	**Inefficiency** (%)
Multiplicity cut	1.06
IV veto	0 (< 0.1%)
FV veto	0.06 ± 0.11
^9Li reduction	0.536 ± 0.019
Muon veto dead time	0.14

Table 5.5: Summary of Gd-n IBD inefficiency

5.3.2.2 Neutron detection efficiency

Biases coming from the neutron detection efficiency were evaluated for Gadolinium selection using data from ^{252}Cf calibration runs and IBD events (see below). This study allowed to compute a correction factor from data-MC comparison. This correction was applied to the MC simulation for the final fit and consists in three independent contributions:

$$c_{Gd} = c_{Gd}^{GdF} \times c_{Gd}^{IBD} \times c_{Gd}^{Sio} \tag{5.29}$$

where c_{Gd}^{GdF} is the correction from Gd neutron capture fraction, c_{Gd}^{IBD} from the selection efficiency and c_{Gd}^{Sio} from spill-in/out effects. The correction value was used to normalize the MC spectrum, whereas its uncertainty was used in the final fit (see section 5.5).

Gd fraction

The fraction of Gadolinium neutron capture is related to the Gd concentration within liquid scintillator. In the ν-Target it is expected to have about 85% of neutron capture on Gadolinium. The remaining neutron captures are mainly on

Hydrogen, while neutron capture on Carbon are negligible within the ν-Target. A precise Gd fraction measurement was performed using calibration runs with ^{252}Cf source. A similar selection than the one detailed in section 3.2.1.1 was applied. Background subtraction was ensure with a tighter prompt energy window starting at 4 MeV. The Gd fraction was computed with the following formula:

$$f_{Gd} = \frac{N(Gd-n)}{N(Gd-n\,+\,H-n)} = \frac{N(E_{vis} \in [3.5,\,10]\,MeV)}{N(E_{vis} \in [0.5,\,10]\,MeV)} \tag{5.30}$$

where $N(Gd-n)$ is the number of Gadolinium neutron captures, selected with an energy window of $[3.5, 10]$ MeV, and $N(Gd-n\,+\,H-n)$ is the number of Gadolinium and Hydrogen neutron captures, selected with an energy window of $[0.5, 10]$ MeV. $N(Gd-n)$ can suffer for a negligible contamination of Carbon neutron capture contamination. Gd fraction was measured to be:

$$f_{Gd}^{Data} = 85.30 \pm 0.08\% \tag{5.31}$$

$$f_{Gd}^{MC} = 87.49 \pm 0.04\% \tag{5.32}$$

The Gd fraction correction faction is then:

$$c_{Gd}^{GdF} = \frac{f_{Gd}^{Data}}{f_{Gd}^{MC}} = 0.9750 \pm 0.0011\,(stat) \pm 0.0041\,(syst) \tag{5.33}$$

IBD selection efficiency

The correction for IBD selection efficiency is a normalization factor between data and MC in order to match MC and data selection efficiencies. Selection cuts were evaluated simultaneously in order to take into account the possible correlations between them. This correction factor was computed with two different methods, both giving consistent results.

IBD efficiency in ^{252}Cf calibration runs: The first method used calibration runs using a ^{252}Cf source deployed at different position along the ZAxis. This allowed to evaluate the full ν-target volume efficiency with a dependence over z. The dependence over ρ could not be evaluated with ^{252}Cf calibration runs as the source was not deployed off the ZAxis in the ν-target. The dependence over ρ was assumed to be the same as over z. This was validated using the MC simulation. The efficiency of the neutron capture signal was defined as the following:

$$\varepsilon = \frac{N(0.5 < \Delta T < 150\,\mu s \cap 4 < E_{vis} < 10\,MeV \cap \Delta R < 1\,m)}{N(0.25 < \Delta T < 200\,\mu s \cap 3.5 < E_{vis} < 10\,MeV)} \tag{5.34}$$

where N is the number of events selected by the cuts shown in parentheses. The numerator selection is similar to the Gadolinium selection, whereas the denominator selection is a relaxed one, which should allow to select all the Gadolinium neutron captures. This efficiency is defined for the full ν-target as:

$$\varepsilon(z,\rho) = \varepsilon_0 \times f_1(z) \times f_2(\rho) \tag{5.35}$$

where ε_0 is the efficiency at the ν-target center and $f_1(z)$ $f_2(\rho)$ describe the z and the ρ dependences, respectively. Efficiency was computed to be $98.29 \pm 0.06\%$ for data and $98.26 \pm 0.3\%$ for MC simulation. The correction factor was then computed as:

$$c_{Gd}^{IBD,Cf} = \frac{\varepsilon^{Data}}{\varepsilon^{MC}} = 1.0003 \pm 0.0032 \tag{5.36}$$

IBD efficiency in physic runs: The second method used IBD candidates to measure the IBD selection efficiency. This efficiency included the ρ dependence as IBD candidates are uniformly distributed in the detector volume. The efficiency of the neutron capture signal was measured using the following formula:

$$\varepsilon = \frac{N(0.5 < \Delta T < 150\,\mu s \cap 4 < E_{vis} < 10\,MeV \cap \Delta R < 1\,m)}{N(0.25 < \Delta T < 200\,\mu s \cap 3.5 < E_{vis} < 10\,MeV \cap \Delta R < 1.7\,m)} \tag{5.37}$$

where N is the number of IBD candidates. These IBDs candidates were selected with the cut shown in parentheses. As for the selections for the ^{252}Cf runs, the numerator selection is similar the Gadolinium selection, whereas the denominator selection is a relaxed one which should allow to select all Gadolinium neutron captures. In addition to these cuts, the other cuts from the standard selection was applied, except the following cuts, which were applied to the prompt signal in order to reduce the background contamination:

- $0.5 < E_{vis} < 8$ MeV
- $FV < 5.8$

The accidental background contamination was subtracted from IBD candidates using an off-time window measurement (cf. section 5.2.1). After data-MC comparison, the correction faction obtained was:

$$c_{Gd}^{IBD,IBD} = \frac{\varepsilon^{Data}}{\varepsilon^{MC}} = 0.9996 \pm 0.0021 \tag{5.38}$$

Combination: Both correction factors (EQ. 5.36 and EQ. 5.38) were combined in order to extract the IBD efficiency factor. The weighted mean was computed to be 0.99982 ± 0.00176. However, as no evidence for a systematic bias was found in MC simulation, the IBD efficiency correction factor was set to be unity, as the following:

$$c_{Gd}^{IBD} = 1.00000 \pm 0.00176 \tag{5.39}$$

Spill-in/out

The Gadolinium analysis looks for neutron captures on Gadolinium. However, this does not ensure that all IBD reactions occurred in the ν-target. IBD reactions within γ-catcher can produce a neutron coming in the ν-target (spill-in), while neutrons from IBD reaction within the ν-target can escape from the γ-catcher and be captured on Hydrogen nuclei (spill-out).

Due to the fact that the γ-catcher volume is larger than ν-target volume, there are more spill-in events than spill-out events. The measure of the spill-in and spill-out currents suffers from a systematic uncertainty due to the modeling of neutron thermalization. Differences were shown using Tripoli4 neutron simulation and the default Geant4 neutron physics.

Double Chooz developed a custom neutron simulation code to improve the MC simulation: NeutronTH. It was shown that NeutronTH reproduces the observed neutron capture time better than the default Geant4 simulation.

The systematic uncertainty was computed by comparing the spill currents (i.e. the difference between the spill-in and the spill-out currents) obtained with Tripoli4 and NeutronTH. The NeutronTH spill current had been estimated to be $\Phi_{Spill} = 2.08 \pm 0.27\%$ where the uncertainty is the discrepancy from Tripoli4 model. The correction factor was then computed as:

$$c_{Gd}^{Sio} = 1.0000 \pm 0.0027 \tag{5.40}$$

Summary

After combination of the correction factors for the Gd fraction (Eq. 5.33), the IBD selection efficiency (Eq. 5.39) and the spill-int/out effect (Eq. 5.40), the final MC simulation correction factor was computed as:

$$c_{Gd} = 0.9750 \pm 0.001\,(stat) \pm 0.0052\,(syst) = 0.9750 \pm 0.0053 \tag{5.41}$$

As explained above, it is applied as a normalization factor to the MC simulation for the final fit (see section 5.5).

5.3.3 The Hydrogen analysis

The Hydrogen analysis suffers from similar uncertainties than the Gadolinium analysis. However, each of them needed to be recomputed to take into account the discrepancies between Hydrogen and Gadolinium.

5.3.3.1 IBD inefficiency

As for the Gadolinium analysis, the Hydrogen analysis suffers from an IBD inefficiency due to the selection cuts. This inefficiency was estimated and results can be find in the Tab. 5.6. The computation of IV veto inefficiency is described in section 5.1.4. These inefficiencies were applied as corrections to the MC simulation.

Source	**Inefficiency** (%)
Multiplicity cut	2.25
IV veto	$< 0.169\%$
FV veto	0.046 ± 0.015
MPS veto	$< 0.1\%$
^{9}Li reduction	0.508 ± 0.012
Light noise cut	$0.0604 \pm .0012$
Muon veto dead time	0.613

Table 5.6: Summary of H-n IBD inefficiency

5.3.3.2 Neutron detection efficiency

As for the Gadolinium selection, biases from the neutron detection efficiency were evaluated. This evaluation considers the Hydrogen instead of the Gadolinium, and so demonstrates major differences with Gadolinium evaluation. It allowed to compute a MC correction factor from data-MC comparison. This MC correction factor is applied separately to events occuring inside and outside the ν-target.

It consists in four independent contributions c_H^{HF}, the correction from H neutron capture fraction, c_H^{IBD}, from the selection efficiency, c_H^{Spill}, from spill effects, and c_H^{P}, from the proton number.

Hydrogen fraction

Hydrogen fraction was determined with ^{252}Cf calibration data and with IBD candidates. In the ν-target volume, the Hydrogen fraction was derived from the Gadolinium fraction evaluated for the Gadolinium analysis (cf. sec. 5.3.2.2). The number of Carbon neutron captures was neglected as estimated to be $\sim 0.1\%$ of all captures occuring within the ν-target. It was computed with the following formula:

$$c_H^T = \frac{(1 - f_{Gd}^{Data})|_{(x,y,z)=(0,0,0)})}{(1 - f_{Gd}^{MC})|_{(x,y,z)=(0,0,0)})} \tag{5.42}$$

This gave the following results:

$$c_H^{T,Cf} = 1.1667 \pm 0.0481 \tag{5.43}$$

$$c_H^{T,IBD} = 1.1750 \pm 0.0277 \tag{5.44}$$

In the γ-catcher, the Hydrogen fraction was defined as the following for ^{252}Cf source:

$$f_H^{GC,Cf} = \frac{N(0.5 < E_{vis} < 3.5\,MeV)}{N(0.5 < E_{vis} < 10\,MeV)} \tag{5.45}$$

The formula is similar to the one for the Gadolinium fraction, but inverted: the numerator is the number of Hydrogen neutron captures and the denominator is the number of neutron captures on Hydrogen, Gadolinium and Carbon.

For the ^{252}Cf method, in order to avoid neutron capture on Gadolinium, only data with the source deployed at the farthest position from the ν-target were used. For IBD method, the Hydrogen fraction in γ-catcher was computed as:

$$f_H^{GC,IBD} = \frac{N(1.6 < E_{vis}^d < 3.5\,MeV \cap 0.25 < \Delta T < 600\,\mu s \cap \Delta R < 0.8m)}{N(1.6 < E_{vis}^d < 10\,MeV \cap 0.25 < \Delta T < 600\,\mu s \cap \Delta R < 0.8m)} \tag{5.46}$$

where E_{vis}^d is the delayed signal visible energy. Only IBD candidates with prompt an delayed signals in γ-catcher were taken into account. In addition to the cuts shown in parentheses, the other standard cuts were applied, with the exception of the ANN cut, the prompt energy window and the FV veto which were modified. ANN cut was not applied as it was not designed for a delayed energy upper than 3.5 MeV. For the prompt energy window, a cut at $E_{vis} > 4$ MeV was applied to remove accidental contamination. The Gadolinium FV veto was used instead of the Hydrogen FV veto, as this one is not design to work at the energy of the Gadolinium neutron capture. The remaining accidental background was measure by the off-time method and subtracted.

These two methods allowed to evaluate consistent correction factor:

$$c_H^{GC,Cf} = 1.0020 \pm 0.0008 \tag{5.47}$$

$$c_H^{GC,IBD} = 1.0040 \pm 0.0013 \tag{5.48}$$

IBD efficiency

For the Hydrogen analysis, the IBD method is the main method which was used to compute the IBD efficiency. The ^{252}Cf method was applied, the limited deployment of the ^{252}Cf does not allow to cover the full γ-catcher. The neutron detection efficiency for the IBD method was defined as:

$$\varepsilon = \frac{N(ANN > -0.23 \cap 0.5 < \Delta T < 800\,\mu s \cap 1.3 < E_{vis}^{d} < 3\,MeV \cap \Delta R < 1.2\,m)}{N(ANN > -0.40 \cap 0.25 < \Delta T < 1000\,\mu s \cap 1.3 < E_{vis}^{d} < 3.1\,MeV \cap \Delta R < 1.5\,m)} \tag{5.49}$$

Standard IBD cuts were applied with the addition of the following: $0.7 < E_{vis}^{p} <$ 9 MeV, $FV < 5.8$ and the multiplicity cut was extended to 1 ms. The accidental background has been subtracted. This allowed to measure the efficiency for Data and MC to be:

$$\varepsilon_{Data} = 95.12 \pm 0.12\% \tag{5.50}$$

$$\varepsilon_{MC} = 95.16 \pm 0.01\% \tag{5.51}$$

The MC correction factor was then computed with Eq. 5.38 to be:

$$c_H^{IBD} = 0.9995 \pm 0.0017 \tag{5.52}$$

As no evidences for a difference between data and MC were seen, the correction faction was put to unity:

$$c_H^{IBD} = 1.0000 \pm (1.0000 - 0.9995 + 0.0017) = 1.0000 \pm 0.0022 \tag{5.53}$$

A cross-check with ^{252}Cf calibration data was performed and provided consistent results.

Spill uncertainty

In the Hydrogen analysis, neutron captures can occur in both γ-catcher and ν-target. Therefore, several types of spilling can occur. IBD events can be split in three categories:

- IBD candidates with both prompt and delayed signals inside the ν-target volume.
- IBD candidates with both prompt and delayed signals inside the γ-catcher volume.
- Spill events, as any other event configuration.

Spill effect can be measured only by MC simulations and depends mainly on the neutron capture model. As for the Gadolinium analysis, a study was performed with Tripoli4 and Geant4 custom neutron simulation, NeutronTH. The spill uncertainty is evaluated from the discrepancy between Tripoli4 and NeutronTH results. It was computed to be:

$$c_H = 1.0000 \pm 0.0029 \tag{5.54}$$

Proton number

For the Hydrogen analysis the number of protons within the whole volume needs to be measured (it was computed for the Gadolinium analysis as the analysis was focused on the ν-target, see section 2.2.1.1). Studies [163] showed that the γ-catcher proton number was well defined in the MC simulation. However, the proton number within acrylic demonstrated discrepancies between reality and the MC simulation as detailed in table 5.7. A correction factor was computed from the proportion of prompt events in the different volumes, and leads to the following correction:

$$c_H^P = 1.00136 \pm 0.00911 \tag{5.55}$$

Summary

Due to some differences in the ν-target and γ-catcher prompt signal spectra, different corrections were applied for the events occuring inside and outside the ν-target. The ν-target total spectrum was corrected by a factor 1.175 (coming from c_H^T) and the other by a factor 1.002 (coming from c_H^{GC}). No correction was applied to spill events and for the volume-wide neutron detection efficiency, both corrections were set to unity. The combined uncertainty coming from the neutron detection efficiency was estimated to be 1.02%.

Volume	**Correction factor**	**Uncertainty (%)**	**Proportion (%)**
ν-target	1.00	0.30	8.859
γ-catcher	1.00	1.04	90.462
Buffer	1.00	1.00	0.132
ν-target tank	1.39	23.00	0.350
γ-catcher tank	1.00	3.00	0.197

TABLE 5.7: Correction factor and associated uncertainty related to the number of free proton and repartition of prompt events in the different volumes of interest.

5.4 Reactor OFF-OFF period

Thanks to the reactor plant configuration, Double Chooz is the only $\overline{\nu}_e$ reactor experiment able to take data when no reactor cores are running. This period is called reactor “OFF-OFF” or 2-Reactor Off. Selection cuts were applied to runs taken in this period, allowing to test our background model. However, even if both reactor cores are not running, the expected $\overline{\nu}_e$ flux is not zero. β decays from fission products still occurred in the reactor core. These decays produced a residual power which is responsible of $\sim 6\%$ of the reactor core total power after the shutdown. It decreases exponentially due to the fission products decays. However, as IBD interaction has a threshold of 1.806 MeV, the number of fission products with a long enough lifetime and producing such *high* energy $\overline{\nu}_e$ is reduced. ^{106}Ru and ^{144}Ce decay series are the main contributors to the reactor 2-Off flux.

$$^{106}Ru \xrightarrow[371.8\,days]{\beta^-} {}^{106}Rh \xrightarrow[30.1\,s]{\beta^-\,(3.5\,MeV)} {}^{106}Pd \tag{5.56}$$

$$^{144}Ce \xrightarrow[284.9\,days]{\beta^-} {}^{144}Pr \xrightarrow[17.3\,min]{\beta^-\,(3.0\,MeV)} {}^{144}Nd \tag{5.57}$$

The standard $\overline{\nu}_e$ flux prediction cannot be used to predict the flux in this period, as it is computed from fission rate (cf. 2.5.3). Another method was used, using Fispact simulation code [164] to compute the Bateman equations and with Bestiole code [72] which computes the $\overline{\nu}_e$ spectrum.

TAB. 5.8 summarized results for the Gadolinium and Hydrogen selections. The indicated number of IBD candidates in this table is the number of IBD candidates selected by the Gadolinium or the Hydrogen selections. It is a direct measurement of the background. The expected number of IBD candidates is the expectation coming from the background measurements, it included few IBD reactions, due to the non-zero $\overline{\nu}_e$ as explained above.

Analysis	**Gd-n**	**H-n**
Live time	7.238 days	7.15 days
Number of IBD candidates	7	63
Expected number of IBD	1.57 ± 0.47	2.36 ± 0.72
Expected number of IBD candidates	$12.93^{+3.07}_{-1.41}$	$50.41^{+4.30}_{-2.86}$

Table 5.8: Summary of Gadolinium and Hydrogen results for reactor 2-Off data. The expected number of IBD candidates is the combination of the $\overline{\nu}_e$ prediction and of our background model.

5.5 "Final Fit" : θ_{13} determination

Double-Chooz uses two different methods to measure θ_{13}. The main method is the "rate+shape" fit, based on a combination of a fit of the rate and of the spectral shape. The second method is the Reactor Rate Modulation analysis (RRM) which extract θ_{13} from observed and predicted rates at different reactor powers.

5.5.1 Rate and spectral shape analysis

After application of the $\overline{\nu}_e$ selection, results from data and MC simulation were compared in the rate+shape fit. This fit took into account the background model from the background analyses (cf. section 5.2). It was performed with RooFit and Minuit packages [165, 166]. Two fits were performed in order to evaluate θ_{13} for both mass hierarchies.

The fit χ^2 was defined with respect to $\sin^2 2\theta_{13}$ taking into account eight nuisance parameters: Δm^2, the predicted number of residual neutrino in the OFF-OFF period ($\alpha_{residual}$), the cosmogenic nuclei background rate ($\alpha_{^9Li}$), the accidental background rate α_{acc}, the correlated background rate α_{corr}, and the three energy scale parameters (a', b' and c') (cf sec. 3.2.9). It is written as follows:

$$
\begin{aligned}
\chi^2 \; &= \textstyle\sum_i^B \sum_j^B (N_i^{pred} - N_i^{obs}) M_{ij}^{-1} (N_j^{pred} - N_j^{obs}) \\
&+ \tfrac{(\Delta m^2 - \Delta m^2_{ee})^2}{(\sigma_{ee}^{(up,down)})^2} + \tfrac{(\alpha_{residual} - 1)^2}{(\sigma_{residual})^2} \\
&+ \tfrac{(\alpha_{^9Li} - 1)^2}{(\sigma_{^9Li}^{(up,down)})^2} + \tfrac{(\alpha_{acc} - 1)^2}{(\sigma_{acc\,(syst)})^2} + \tfrac{(\alpha_{corr} - 1)^2}{(\sigma_{corr})^2} \\
&+ \left[(a' - a'_{CV}) \quad (b' - b'_{CV}) \quad (c' - c'_{CV})\right] \qquad (5.58) \\
&\times \begin{bmatrix} (\sigma_{a'})^2 & \rho_{a'b'}(\sigma_{a'}\sigma_{b'}) & \rho_{a'c'}(\sigma_{a'}\sigma_{c'}) \\ \rho_{b'a'}(\sigma_{b'}\sigma_{a'}) & (\sigma_{b'})^2 & \rho_{b'c'}(\sigma_{b'}\sigma_{c'}) \\ \rho_{c'a'}(\sigma_{c'}\sigma_{a'}) & \rho_{c'b'}(\sigma_{c'}\sigma_{b'}) & (\sigma_{c'})^2 \end{bmatrix}^{-1} \times \begin{bmatrix} (a' - a'_{CV}) \\ (b' - b'_{CV}) \\ (c' - c'_{CV}) \end{bmatrix} \\
&+2 \left(N_{off}^{obs} \cdot \ln \left[\tfrac{N_{off}^{obs}}{N_{off}^{pred}} \right] + N_{off}^{obs} - N_{off}^{pred} \right)
\end{aligned}
$$

where B is the number of bins in the energy spectrum. The binning used depends of the analysis and is detailed in the following. N_i^{obs} and N_i^{pred} are respectively the number of observed and predicted events within the bin i, with N_i^{pred} computed as follows:

$$
\begin{aligned}
N_i^{pred} = \; & \textstyle\sum_{R=1,2} N_i^{\overline{\nu}_e,R}(\sin^2 2\theta_{13}, \Delta m^2, a', b', c') + N_i^{^9Li}(\alpha_{^9Li}) \\
& + N_i^{corr}(\alpha_{corr}) + N_i^{acc}(\alpha_{acc})
\end{aligned}
\qquad (5.59)
$$

where the sum is done over the expected neutrinos from both reactors (R). The expected number of background events in a given bin is function of the background's nuisance parameter, α_{BG}.

M_{ij} is a sum of several covariance matrix:

$$
M_{ij} = M_{ij}^{stat} + M_{ij}^{reactor} + M_{ij}^{efficiency} + M_{ij}^{^9Li\,shape} + M_{ij}^{acc\,(stat)} \qquad (5.60)
$$

where M_{ij}^{stat} is the diagonal matrix taking into account the statistic uncertainty. $M_{ij}^{reactor}$ accounts for the uncertainty on the $\overline{\nu}_e$ prediction. $M_{ij}^{efficiency}$ for the uncertainties due to neutron detection inefficiency. $M_{ij}^{^9Li}$ is the shape error of the cosmogenic nuclei spectrum and $M_{ij}^{acc\,(stat)}$ is the diagonal matrix taking into account the statistical uncertainty of the accidental background rate.

The last line of equation 5.58 shows the contribution of the OFF-OFF reactor period candidates. A Poisson statistics log-likelihood is used to compared N^{obs}_{off} and N^{pred}_{off} due to the low statistics.

Using the χ^2 formula, the fit was performed in two ways to extract θ_{13}. The first way, called “rate only” consists in the comparison between predicted and observed spectra on the whole range of the prompt energy spectrum, without taking into account the binning. The second way, called “rate+shape”, is the main fit used by Double Chooz. It takes into account the spectral behavior of the prompt energy spectrum.

5.5.1.1 Gadolinium results

For the Gadolinium analysis the rate only analysis allowed to measured the following value for θ_{13} [79]:

$$\sin^2 2\theta_{13} = 0.090^{+0.036}_{-0.037} \tag{5.61}$$

Range	**Bin width**	**Number of bins**
0.5 – 8 MeV	250 keV	30 bins
8 – 10 MeV	500 keV	4 bins
10 – 12 MeV	1 MeV	2 bins
12 – 20 MeV	2 MeV	4 bins
0.5 – 20 MeV		40 bins

Table 5.9: Bins distribution for the rate+shape analysis of the Gadolinium selection.

For the rate+shape analysis, we used the 40 bins detailed in Tab. 5.9. This binning allows to extract information about background rates since there is no possibility to have IBD for a prompt energy above 12 MeV. The rate+shape analysis was performed for inverted and normal hierarchy. The Δm^2_{31} used were taken from MINOS results [167] and are in agreement with T2K results [37, 38]. For normal hierarchy we used $\Delta m^2_{31} = 2.44^{+0.09}_{-0.10} \times 10^{-3}$ eV^2 , and $|\Delta m^2_{31}| = 2.38^{+0.09}_{-0.10} \times 10^{-3}$ eV^2 for inverted hierarchy. The results are the following [79]:

$$\sin^2 2\theta^{NH}_{13} = 0.090^{+0.032}_{-0.029} \; with \; \chi^2_{min}/d.o.f. = 52.2/40 \tag{5.62}$$

$$\sin^2 2\theta_{13}^{IH} = 0.092^{+0.033}_{-0.029} \; with \; \chi^2_{min}/d.o.f. = 52.2/40 \tag{5.63}$$

The fit was also performed without constraints on ^9Li and correlated background rates in order to cross-check the results. It allowed to find $\sin^2 2\theta_{13} = 0.088^{+0.030}_{-0.031}$ with a ^9Li rate of $0.49^{+0.16}_{-0.14}$ events per day and $0.541^{+0.052}_{-0.048}$ events per day for the correlated background rate. These results are consistent with the ones obtained in the standard analysis.

The prompt visible energy spectrum is shown in FIG. 5.36. The best fit prediction is shown on this spectrum, as well as the non-oscillation hypothesis and the estimated background shapes. The ratio of the data to the non-oscillation prediction demonstrates a clear deficit between 1 MeV and 4 MeV, which demonstrates the neutrino oscillation.

A spectrum distortion is observed around 5 MeV, called “bump”. Double Chooz investigated deeply to understand this unexpected behavior before presented it with a significance of 3σ [175]. RENO collaboration confirmed this excess at the Neutrino 2014 conference [176], following by the Daya Bay collaboration at the ICHEP 2014 conference [177] with a significance of 4σ.

The investigations on the cause of the distortion disfavored the energy scale, thanks to the good agreement between data and MC at this energy (cf. section 3.2.7). FIG. 5.37 shows that the distortion depends on the nuclear reactor power which disfavored a new background. The preferred hypothesis is a deviation from the reactor flux prediction (see section 2.5). Study [169] indicates that the distortion could be corrected by using different nuclei database to perform the summation method. This study consolidates the goal of the TAGS experiments, which aim to remeasure the fission products β spectra, in order to correct for previous measurements errors like the pandemonium effect [168].

Hayes and collaborators [148] showed that the conversion method used to predict the reactor spectra could be biased by an underestimation of the errors on the Schreckenbach measurements. They ask for a remeasurement of the different ^{235}U, ^{238}U and ^{239}Pu β spectra. They suggest that short baseline experiments, operated close to reactor using a highly enriched ^{235}U fuel, could be interesting as they restrict the resulting antineutrino flux to the fragments produced by a single actinide.

The Double Chooz Near Detector construction finished in the end of 2014. ÉDF, which operates the Chooz-B nuclear power plant, is planning to run only one reactor for a long period. Taking advantage of this situation, a measurement of the $\overline{\nu}_e$ spectrum, without oscillations and without uncertainties coming from the convolution of several reactor fluxes, will provide a pure $\overline{\nu}_e$ spectrum of a PWR. This could be an interesting result to understand the distortion.

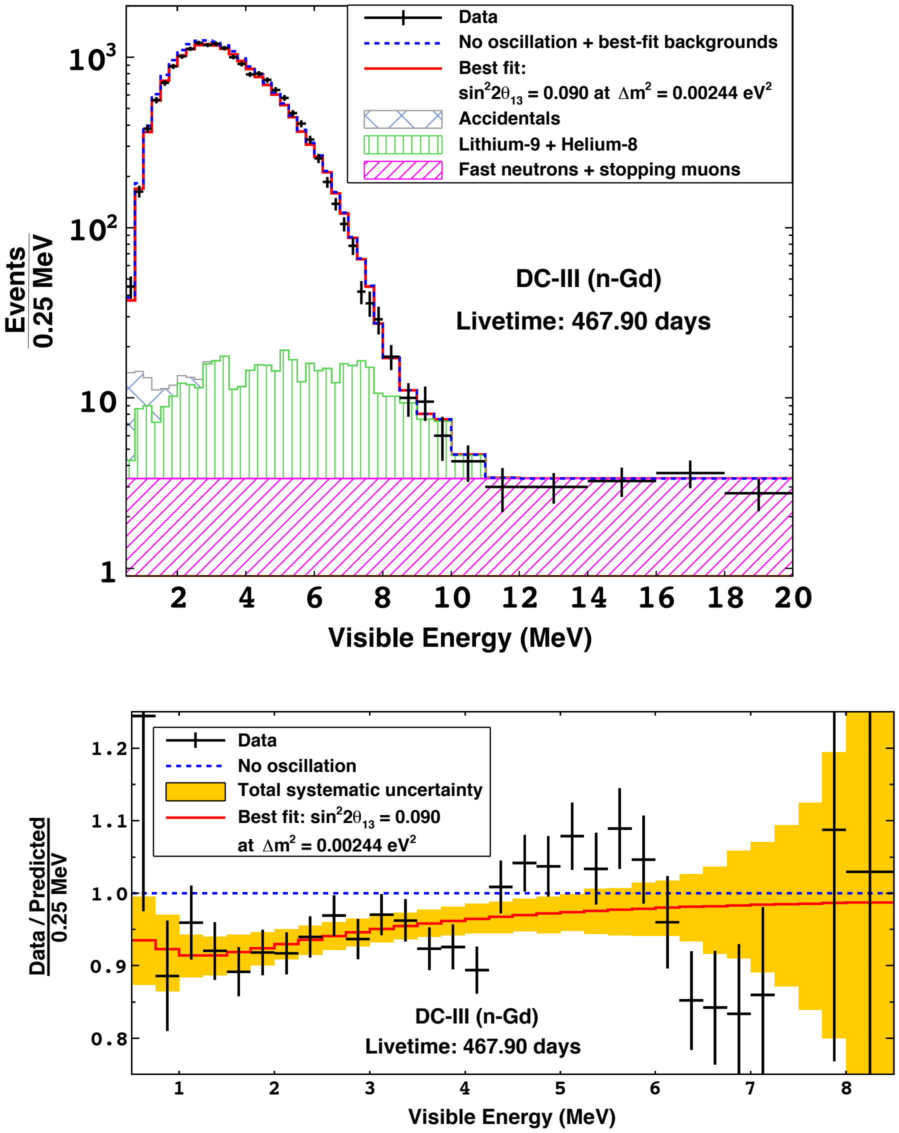


FIGURE 5.36: *Top*: Prompt spectrum in black compared to null oscillation and best-fit predictions in blue and red, respectively, stacked on the backgrounds. *Bottom*: Ratio of the number of IBD candidates over the prediction as a function of energy, with the best θ_{13} fit.

5.5.1.2 Hydrogen results

For the Hydrogen analysis, due to the discovery of $[4, 6]$ MeV spectral distortion, the possibility to have a rate+shape fit biased by this distortion was considered. Therefore, in order to be safe from the $[4, 6]$ MeV spectral distortion, the

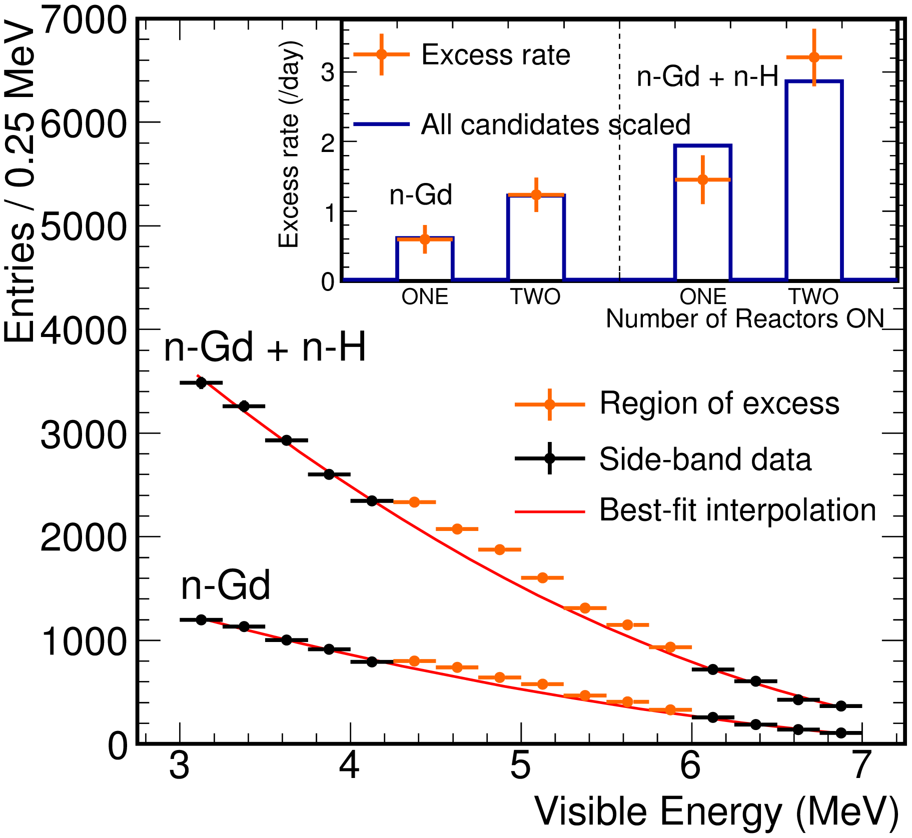


FIGURE 5.37: The energy spectrum of the prompt signal for IBD candidates with neutrons captured on Gd and one including H captures (Gd+H), H captures were selected with the H-II analysis. Points show the data and lines show the second order polynomial functions. *Inset figure*: points show the correlations between the observed rate of the excess and the number of operating reactors, and the histograms show the total IBD candidate rate (area normalized). The H capture sample includes accidental background with a rate comparable to the IBD signals and therefore the total rate of the Gd+H sample has an offset due to this background in addition to IBD signals which is proportional to the reactor power.

RRM analysis was the main analysis performed for the Hydrogen anaysis, and the rate+shape analysis was only performed as a cross-check. Analyses are still ongoing for this rate+shape fit. Preliminary results allowed to provide the following result:

$$\sin^2 2\theta_{13} = 0.124^{+0.030}_{-0.039} \tag{5.64}$$

The prompt visible energy spectrum is shown in FIG. 5.38. The best fit prediction is shown on this spectrum, as well as the non-oscillation hypothesis and the estimated background shapes.

The $[4, 6]$ MeV spectral distortion was confirmed with the Hydrogen analysis. Consistent structures are seen between Hydrogen and Gadolinium spectral distortion, as shown in FIG. 5.39. The significance of the Hydrogen distortion is not yet computed.

5.5.2 Reactor Rate Modulation analysis

The Reactor Rate Modulation analysis (RRM) consists in the comparison between observed and predicted rate at different reactor powers. In addition to provide a θ_{13} measurement, it allows to extract an independent background model, as both are computed simultaneously. Double Chooz is the only experiment to use such analysis with a reactor 2-Off period.

Thanks to the reactor central configuration, with only two reactors, Double Chooz is able to measure the $\overline{\nu}_e$ flux in different reactor configurations: with two reactors on (2-reactor-On), with one reactor off (1-reactor-On) and with two reactors off (reactor-Off). RRM used these measurements to perform a fit. RRM can also be applied without the reactor-Off period, but this reduces the background constraint.

From the comparison between the observed and the predicted rates, this analysis uses the following correlation between $\sin^2 2\theta_{13}$ and the total background rate:

$$R^{obs} = B + R^{pred} = B + (1 - \alpha_{osc} \sin^2 2\theta_{13}) R^{\overline{\nu}_e} \tag{5.65}$$

where R^{obs} is the observed rate, R^{pred}, the predicted rate, $R^{\overline{\nu}_e}$, the expected $\overline{\nu}_e$ rate in the non-oscillation hypothesis and α_{osc} the mean disappearance coefficient $< \sin^2(\Delta m^2 L/4E) >= 0.55$. This last coefficient was computed by means of simulations for each one of the data points.

RRM has three sources of systematics: the detection efficiency ($\sigma_d = 0.63\%$), the residual $\overline{\nu}_e$ prediction ($\sigma_\nu = 30\%$) in the reactor 2-Off period (see section 5.4) and the $\overline{\nu}_e$ prediction in reactor-on data (σ_r). σ_r value depends on the reactor power. The RRM fit is performed with the following χ^2:

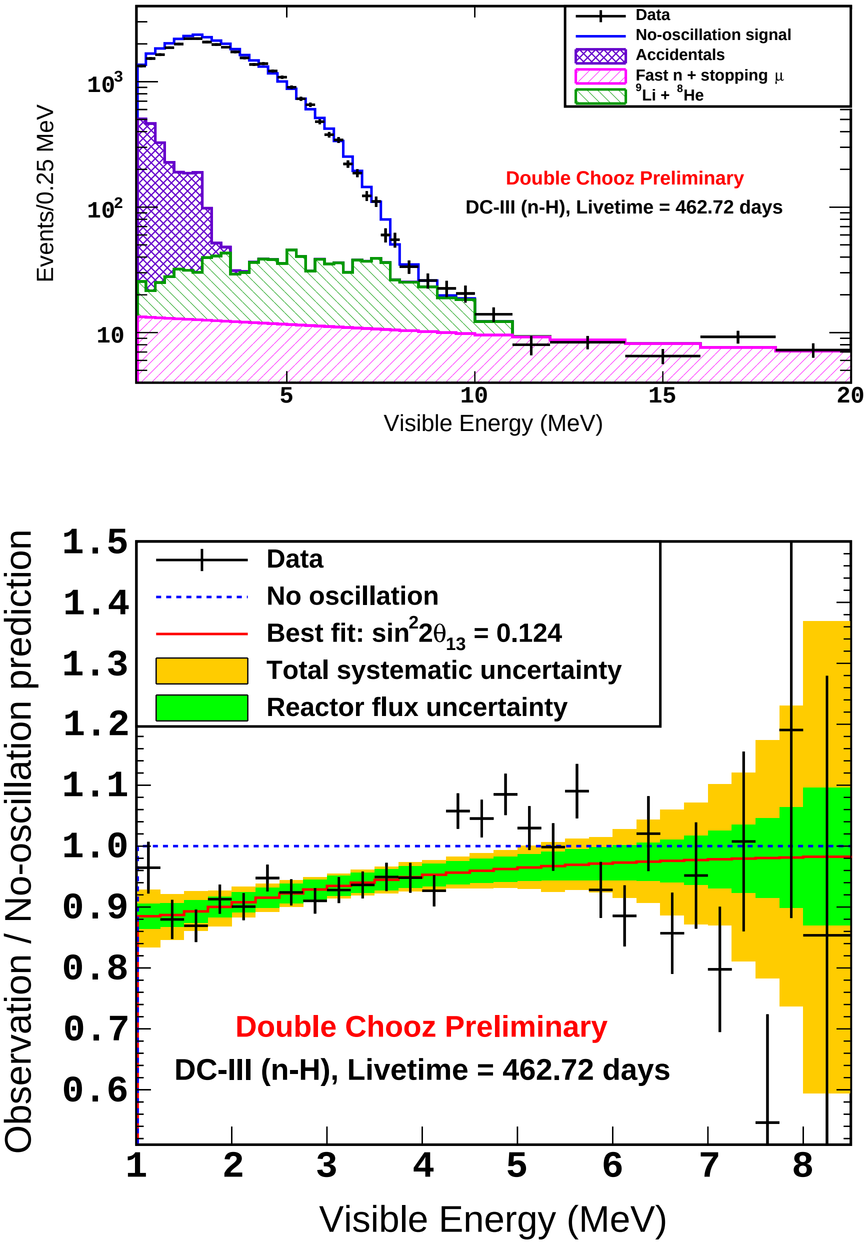


FIGURE 5.38: *Top*: Prompt spectrum in black compared to null oscillation and best-fit predictions in blue and red, respectively, stacked on the backgrounds. *Bottom*: Ratio of the number of IBD candidates over the prediction as a function of energy, with the best θ_{13} fit.

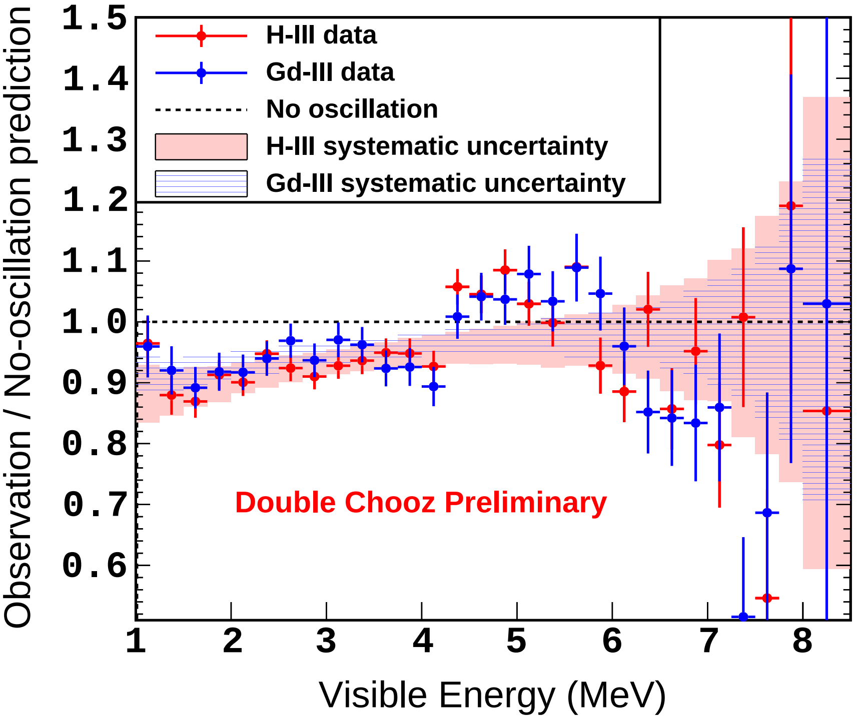


FIGURE 5.39: Ratio of background-subtracted data to no-oscillation prediction in the DC-III (H-n) (red) and DC-III (Gd-n) (blue) analyses. Bands show total systematic error for each analysis. The dashed black line is the no-oscillation prediction.

$$\begin{aligned}\chi^2 = \; & \sum_{i=1}^{6} \frac{(R_i^{obs}-R_i^{pred}-B)^2}{(\sigma_i^{stat})^2} + \frac{\varepsilon_d^2}{\sigma_d^2} + \frac{\varepsilon_\nu^2}{\sigma_\nu^2} + \frac{\varepsilon_r^2}{\sigma_r^2} \\ & +2\left(N_{off}^{obs}\cdot\ln\left[\frac{N_{off}^{obs}}{N_{off}^{pred}}\right] + N_{off}^{obs} - N_{off}^{pred}\right) \\ & +\frac{(B-B^{pred})^2}{(\sigma_B)^2}\end{aligned} \tag{5.66}$$

where R_i^{obs} and R_i^{pred} are, respectively, the observed and the predicted rate for a given bin, and B the total background rate. R_{off}^{obs} and R_{off}^{pred} are, respectively, the observed and the predicted rate for the OFF-OFF period.

It is possible to run the RRM analysis with and without the reactor-Off data, but also with and without the background model. The different configurations were run and allows the computation of consistent $\sin^2\theta_{13}$ and B values.

For the Gadolinium analysis, results are summarized in Tab. 5.10. Fig. 5.40 shows the two-parameter fit and the contour plot when using reactor-on, reactor-off and the background model in RRM.

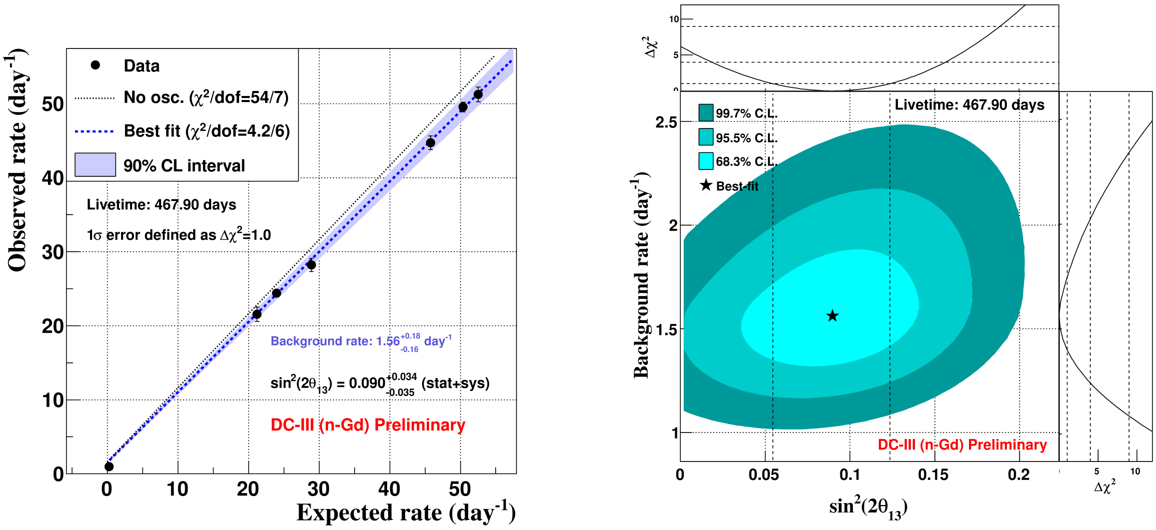


Figure 5.40: Gadolinium RRM fit including the reactor-off data and the constraint provided by the DCIII (Gd-n) background model. *Left*: (θ_{13},B) fit, superimposed to the null oscillation hypothesis assuming the background estimates in DCIII. *Right*: 1, 2 and 3σ (θ_{13},B) contour plot.

Configuration	$\sin^2 2\theta_{13}$	B	$\chi^2/d.o.f.$
Reactor-On + Reactor-Off + Background model	$0.090^{+0.034}_{-0.035}$	$1.56^{+0.18}_{-0.16}$	4.2/6
Reactor-On + Reactor-Off	0.060 ± 0.039	$0.93^{+0.43}_{-0.36}$	1.9/5
Reactor-On	0.089 ± 0.052	1.56 ± 0.86	1.3/4

Table 5.10: Summary of the different configurations and results of the RRM analysis for Gadolinium selection

For the Hydrogen analysis, results are summarized in Tab. 5.11. Fig. 5.41 shows the two-parameter fit and the contour plot when using reactor-on, reactor-off and the background model in RRM.

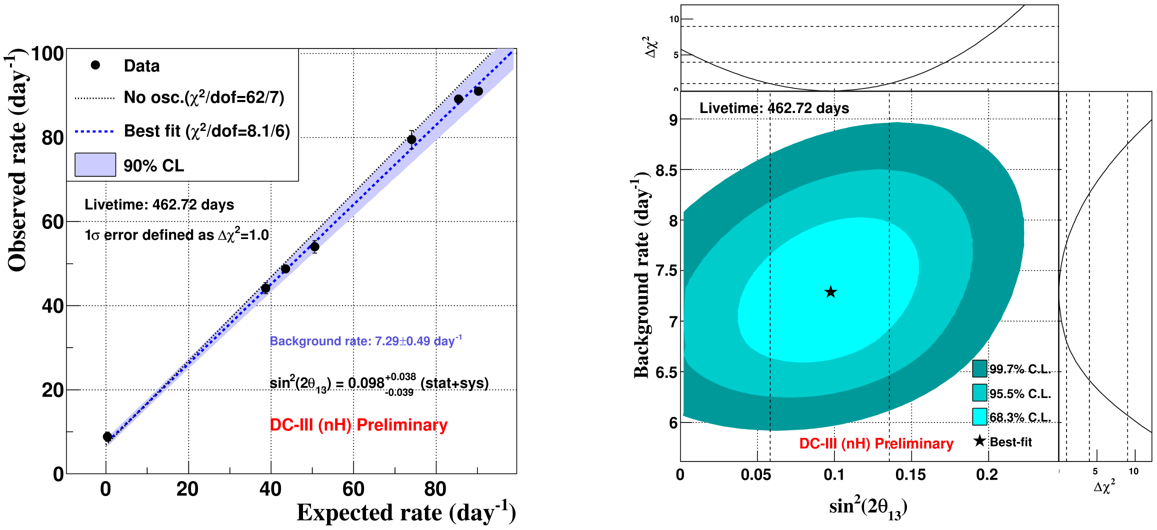


FIGURE 5.41: Hydrogen RRM fit including the reactor-off data and the constraint provided by the DCIII (H-n) background model. *Left*: (θ_{13},B) fit, superimposed to the null oscillation hypothesis assuming the background estimates in DCIII. *Right*: 1, 2 and 3σ (θ_{13},B) contour plot.

Configuration	$\sin^2 2\theta_{13}$	B	$\chi^2/d.o.f.$
Reactor-On + Reactor-Off + Background model	$0.098^{+0.038}_{-0.039}$	7.29 ± 0.49	8.1/6
Reactor-On + Reactor-Off	$0.123^{+0.042}_{-0.043}$	8.28 ± 0.87	6.2/5

TABLE 5.11: Summary of the different configurations and results of the RRM analysis for Hydrogen selection

5.6 Summary and Outlooks

Thanks to the increase of statistic and to the development of new background rejection methods, like the IV veto, the Gadolinium analysis results have been improved with respect to the previous Double Chooz Gadolinium analysis. The improved knowlegde on the different backgrounds allowed to reduce the background systematic by a factor 2. The development on the energy reconstruction (see chapter 3) led to a improvement by a factor ~ 2 of the energy reconstruction uncertainty. Thanks to these huge development of analysis, the total uncertainty has been reduced by a factor 1.25 with respect to the previous Gadolinium analysis. Therefore, the Double Chooz collaboration was able to measure θ_{13} as $\sin^2 2\theta_{13} = 0.090^{+0.032}_{-0.029}$.

Double Chooz was the first reactor $\overline{\nu}_e$ experiment to perform a Hydrogen analysis. Hydrogen analysis presented here is a major improvement of the previous one. Thanks to the innovative signal selection, using an artificial neural network, and to the background rejection method, with IV veto, we were able to reduce the background systematics to a level comparable to the Gadolinium analysis. Other systematics are also at the same level between both analyses, with the notable exception of one due to the proton number inside the γ-catcher. Statistics are also higher with the Hydrogen analysis than with the Gadolinium analysis, due to the large volume of the γ-catcher. Due to the high accidental background, and to the results of the first Hydrogen analysis, it was not expected to reach such level of systematics. The Double Chooz collaboration was able to measure $\sin^2 2\theta_{13} = 0.098^{+0.038}_{-0.039}$.

We were able to combine the Gadolinium and Hydrogen results with a RRM fit, leading to a combine value of θ_{13} of $\sin^2 2\theta_{13} = 0.090 \pm 0.033$. FIG. 5.42 shows the three-parameter fit (θ_{13},B_{Gd},B_H) and the contour plot when using reactor-on, reactor-off and the background model in RRM. A combination of the Gadolinium and Hydrogen results with a rate+shape fit is currently studied by the collaboration.

It is important to notice that Double Chooz currently has only one detector running, whereas the other experiments, RENO and Daya Bay, have several detectors running. Therefore, since they compare a near detector $\overline{\nu}_e$ flux and a far detector $\overline{\nu}_e$ flux, they are able to supress some detection systematics arising from the comparison between a data flux and a MC flux. The current published θ_{13} measurements of Daya Bay and RENO are: $\sin^2 2\theta_{13} = 0.084 \pm 0.005$ (Daya Bay [54]) and $\sin^2 2\theta_{13} = 0.113 \pm 0.013(stat) \pm 0.019(syst)$ (RENO [49]). The second phase of Double Chooz started in the end of 2014, with the start of the Near Detector data taking. FIG. 5.43 shows the sensibility of Double Chooz on the $\sin^2 2\theta_{13}$ measurement. This figure shows that with the 6 month of Near Detector data already taken, the sensibility will be improved by a factor ~ 1.5.

On possible outlook of the Hydrogen result is that we could build a Double Chooz-like reactor $\overline{\nu}_e$ experiment without Gadolinium. Such experiments would have a detector with a large ν-target, encompassing the current Double Chooz ν-target and γ-catcher volumes. Double Chooz demonstrates with this last Hydrogen analysis, that the presence of a γ-catcher is not a necessity for a Hydrogen analysis. The use of only one volume will reduce the systematics linked to the detection of neutron, like spill effects, as well as the ones linked to the energy reconstruction: a single volume will be more uniform and therefore simpler to calibrate. The proton number is an issue for the Double Chooz Hydrogen analysis as it was not computed when the detector was designed and build. This non-computation is due

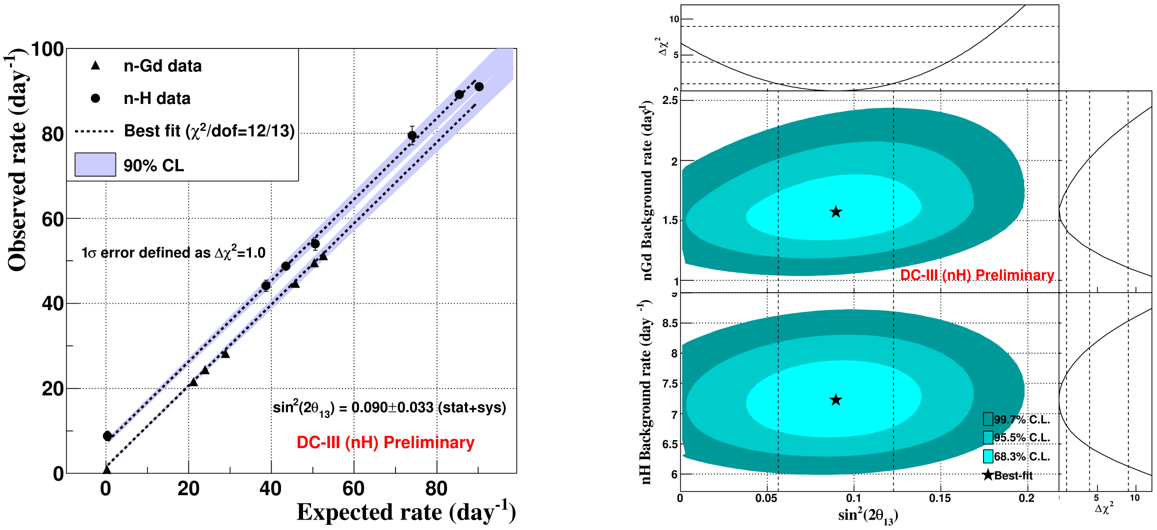


FIGURE 5.42: Gadolinium+Hydrogen combined RRM fit including the reactor-off data and the constraint of the background models. The reactor flux and residual $\overline{\nu}_e$ rate uncertainties are fully correlated in the DCIII (Gd-n) and (H-n) samples, while the detection systematics are assumed to be fully uncorrelated. The errors on the B_{Gd} and B_H rates are assumed to be fully uncorrelated as correlations up to 50% do not affect significantly the fit outcome (the maximum possible correlation has been estimated to be below 50%). *Left*: (θ_{13},B_{Gd},B_H) fit, superimposed to the null oscillation hypothesis assuming the background estimates in DCIII. *Right*: 1, 2 and 3σ (θ_{13},B) contour plots.

to the fact that it was not planned to use γ-catcher for $\overline{\nu}_e$ detection when Double Chooz was designed. Such experiments would be cheaper, as the Gadolinium is expensive. They would also be simpler to build, with only one active volume and with “simpler” liquids, as special measures need to be taken to minimize the risks of a deterioration of the liquid stability by the precense of Gadolinium nuclei.

In both analyses, I contributed to the development of the IV veto rejection method. We succeed in developing a method which allows to reject both accidental and correlated background. Double Chooz is the only current reactor $\overline{\nu}_e$ experiment able to use such method. Daya Bay and RENO have a water inner veto, using Cerenkov detection, which is not efficient to detect fast neutron and γ. Consequently, by design, they cannot reuse this kind of method to have a better control on the correlated and accidental backgrounds.

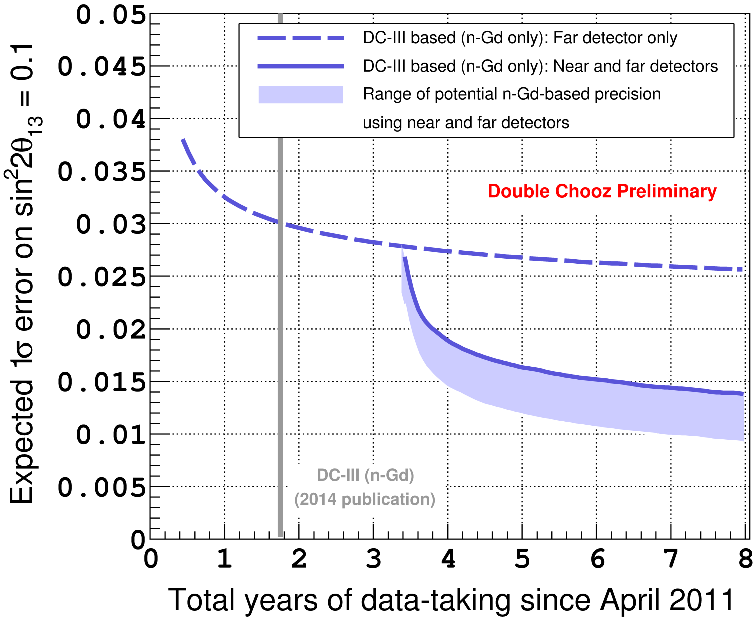


FIGURE 5.43: Expected sensibility on the $\sin^2 2\theta_{13}$ measurement the Gadolinium analysis, with and without the Near Detector.

Chapter 6

Correlated Background

Correlated background is the term used in Double Chooz to indicate a background where prompt and delayed signals come from the same physical process. This background is produced by muon interactions within the detector or its surrounding. More precisely it designates fast neutron and stopping muon backgrounds. Cosmogenic radioisotopes could also be considered as correlated background following this definition, however, Double Chooz considers them separately from fast neutrons and stopping muons.

In this chapter, the Double Chooz correlated background analysis is detailed. During this thesis, I was in charge of the correlated background rate and shape measurement of the Hydrogen analysis. Cross-checks of the official correlated background rate and shape measurement of Gadolinium analysis were also performed and are described here.

6.1 Correlated Background

6.1.1 Fast Neutrons

Fast Neutron background is generated by muon spallation in the rocks surrounding the detector. Due to their low interaction cross-section, fast neutron can cross the detector and be captured within the ID. Their interactions inside the liquid scintillator can be a proton recoil (mimicking a prompt signal) and their captures on a nuclei like Gadolinium or Hydrogen (mimicking a delayed signal). Proton recoils have a large energy spectrum going from few keV to more than a GeV. The time and space correlations between proton recoil and the neutron capture depend on the neutron thermalization. Therefore, such events cannot be separated from $\overline{\nu}_e$ events. However, it is possible to discriminate them using a tagging method, as the IV veto method presented in section 5.1.4. It is also possible to use their large proton recoil spectrum to measure their shape and their rate. Their spatial distribution within the detector volume is expected to be uniform.

6.1.2 Stopping Muons

The Stopping Muon background is due to low energy muons entering in the detector through the chimney hole and decaying. The muon energy loss by ionization mimics the prompt signal, whereas the emission of the Michel electron during the decay mimics the delayed signal.

Due to the muon time-life $\tau_\mu \simeq 2.2\mu$s and since the Michel electron deposit its energy quickly, the time and the space correlations between prompt signal and the delayed signal are expected to be short (few μs and few cm). In addition, since the muon was not tagged by the muon cuts based on the visible energy, stopping muons in the correlated background are expected to deposit a small amount of energy in the detector. Therefore, these muons are expected to have a short track in the detector ($\lesssim$ 10 cm). For these reasons, stopping muons are expected to occur in a small region of the detector near the chimney.

Due to the reflections occuring within the chimney, some characteristics are expected to arise from their light pattern (i.e. hit PMTs, distribution of the charge, etc.). In the present analyses, these characteristics are propagated to the vertex reconstruction algorithm and the FVV method used it in order to reject the stopping muon background (cf. chapter 5). This allowed to have a negligible stopping muon background.

6.2 Measuring Correlated Background Shape and Rate

In order to increase the precision of the θ_{13} measurement, we measured the remaining background shapes and rates, cf. sec. 5.2. For the correlated background, we rely on the background shape estimation to compute the remaining rate. Since both correlated background components show a prompt energy spectrum extending at high energy, we can select a pure sample of the correlated background by requiring a prompt energy $E > 12$ MeV, where the $\overline{\nu}_e$ and other background rates are negligible.

This high energy correlated background sample allows to measure the high energy shape. For the low energy shape ($E < 12$ MeV), a common approach it to assume the low energy shape to be flat and then to extrapolate the rate from the high energy shape measurement. This is the approach used by the Daya Bay and the RENO collaboration [46, 49]. However, the validity of this approach depends on the validity of the flat shape assumption. In Double Chooz, we did not infer directly the low energy shape from the high energy shape. Instead, we measured it by selecting a pure sample of the correlated background using different techniques.

MC techniques could have been used to measure the correlated background shape at low energy. For the stopping muons simulation this requires an accurate modeling of the whole detector volumes in order to have a precise simulation of the acceptance effect for incoming muons. The modeling for the chimney is very complex on the geometrical and mechanical point of view. For the fast neutrons, the simulation is also difficult, as the fast neutron production is not well known. In addition, the neutron transport is not properly modeled by the current simulation.

Due to these complications, we did not use MC simulations to measure the correlated background shape. Therefore, we used tagging methods to select a pure sample of the correlated background. These samples were then used to measure the shape of the correlated background in order to extract the remaining background rate. Several tagging methods are possible. Here, we present these different methods, as well as a general presentation of the use of such tagging method.

Tagging methods

The use of a tagging method implies to select to subsample of the correlated background. In order to have a proper computation of the rate, this subsample needs to be normalized to the full sample of the correlated background. As explain above, it is currently not possible to select the full sample of the correlated background over the full energy range. However, at high energy, the events selected by the IBD analyses can be assumed to be pure correlated background sample. Therefore, it is possible to compute a normalization factor, c_{Tag}, as follows:

$$c_{Tag}[\%] = 100 \times \frac{\int_{E_{min}}^{E_{max}} Tagged\,events}{\int_{E_{min}}^{E_{max}} IBD\,selection} \tag{6.1}$$

where $[E_{min},\, E_{max}]$ is the energy window in which the sample of IBD candidates (i.e. the sample of events selected by the IBD analyses) is expected to be a pure sample of the correlated background. We can also compute a tagging efficiency, ε_{Tag}, defined as the ratio between the number of tagged events and the number of expected correlated background events. If the tagged sample is included within the IBD candidates sample, then c_{Tag} is also the tagging method efficiency. However, if the tagged sample is rejected from the IBD sample, as it is the case in the analyses presented below, the tagging efficiency, ε_{Tag}, is computed as follows:

$$\varepsilon_{Tag}[\%] = 100 \times \frac{\int_{E_{min}}^{E_{max}} Tagged\,events}{\int_{E_{min}}^{E_{min}} IBD\,selection + \int_{E_{min}}^{E_{max}} Tagged\,events} \tag{6.2}$$

The relation between ε_{Tag} and c_{Tag} is:

$$c_{Tag}\,[\%] = \frac{1}{\frac{100}{\varepsilon_{Tag}\,[\%]} - 1} \tag{6.3}$$

After the correlated background shape determination, we can compute the correlated background rate with the following formula:

$$Rate\,[events/day] = \frac{\int_{E_{prompt,min}}^{E_{prompt,max}} BG\,Shape}{\tau_{run}\,[day] \times c_{Tag}} \times \rho_{Tag} \tag{6.4}$$

where $\int_{E_{prompt,min}}^{E_{prompt,max}} BG\,Shape$ is the integral over the IBD selection prompt energy window of the measured correlated background shape, τ_{run} is the live time of the full data sample, consisting in the length of the data sample after subtraction of the dead time due to muon veto (cf. sec. 5.1.2 and 5.1.3), and ρ_{Tag} is the purity of the tagging selection method.

Several analysis were performed and showed that the stopping muon background was negligible, in the Gadolinium analysis [79] and in the Hydrogen analysis [182, 183], after application of the FV veto (see section 5.1.2.7). Therefore, we did not select separately the stopping muons and the fast neutrons, and then did not compute the stopping muon and the fast neutron rates and shapes separately.

Pulse Shape tagging

Pulse Shape discrimination has also been investigated for the correlated background selection. Such methods are expected to tag fast neutrons thanks to the differences between the pulse shapes of the fast neutron proton recoils and of the electrons and positrons. For the Hydrogen analysis, the MPS cut (cf. section 5.1.3) could have been used to tag fast neutron background. However, this cut is expected to have an energy dependence, due to the cut applied at 1.2 MeV. Therefore, it was not used for the correlated background measurement.

Inner Veto tagging

Inner Veto tagging is the main correlated background tagging method of Double Chooz since the first publication [47]. It relies on the same arguments as the Inner Veto Veto rejection method (cf. section 5.1.4) to select the correlated background. FIG. 6.1 illustrates the expected coincidence between the ID and the IV for the correlated background.

In the case of the fast neutrons, the neutron interacting in the IV and the one interacting in the ID are expected not to be the same. Therefore, the energy deposited in the IV and the one deposited in the ID are not correlated. Consequently,

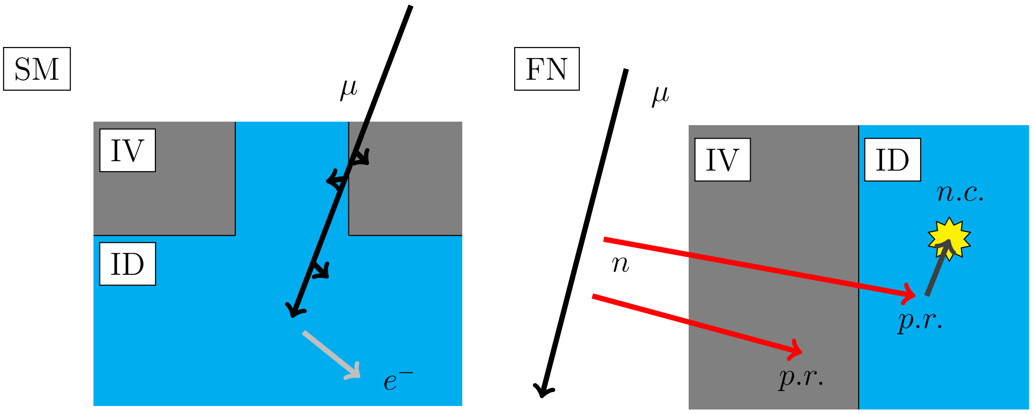


Figure 6.1: Illustration of the Inner Veto tagging method. On the left, the stopping muon deposits a low amount of energy inside the IV, allowing to activate it in correlation with the the background event within the ID. On the right, due to the high neutron yield of muon spallations, we can expect to have a fast neutron interacting within the IV in correlation with the background event within the detector. *p.r.* is for proton recoil and *n.c.* for neutron capture.

the Inner Detector energy spectrum of the part of the fast neutron background with ID-IV coincidence can be assumed to be the same as the one of the fast neutron background without ID-IV coincidence. However, we cannot totally exclude the possibility of having the same neutron interacting in the IV (via proton recoil) and in the ID. Even in this latter case, thanks to the large energy spectrum of proton recoils, we expect no correlation between the amount of energy deposited inside the IV and the one deposited inside the ID.

Since stopping muons interacting inside ID enter through the chimney, the angular acceptance of the Inner Veto for stopping muons is small. Therefore, the efficiency of Inner Veto tagging for stopping muons is expected to be small.

Inner Veto tagging purity suffers from two other types of events which can be tagged by IV in coincidence with an event within ID: $\gamma - \gamma$ and $\overline{\nu}_e$ in accidental coincidence with a γ interacting in the Inner Veto (cf section 5.1.4). The latter can be avoided using the ID-IV space and time correlations, thanks to the fact that IV and ID events are not correlated. However, $\gamma - \gamma$ events have correlation between IV and ID. Therefore, the rejection from an IV tagged sample is more complex, this is detailed in the following section.

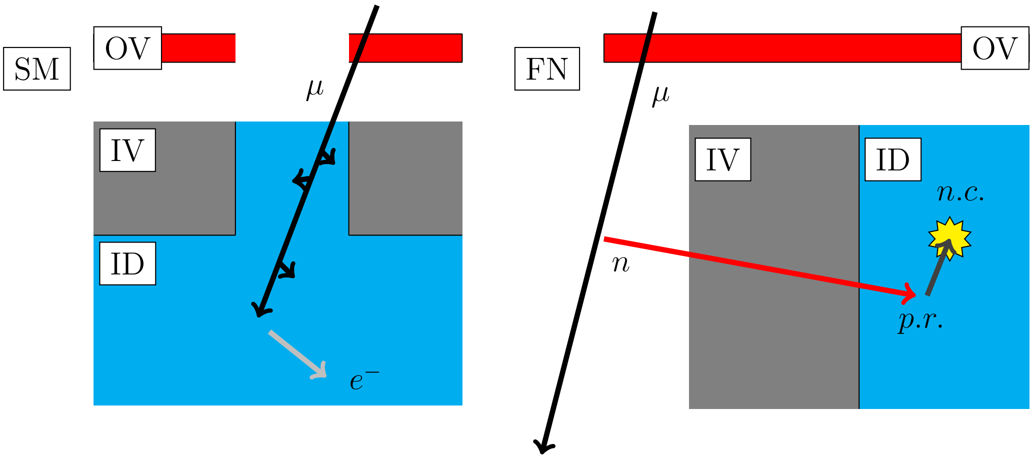


FIGURE 6.2: Illustration of the Outer Veto tagging method. On the left, a stopping muon which activates the OV before entering within the ID. On the right, a fast neutron tagged via its progenitor muon OV activation. *p.r.* is for proton recoil and *n.c.* for neutron capture.

Outer Veto tagging

Outer Veto can be use to tag the correlated background through the tagging of their progenitor muon, as illustrated in FIG. 6.2. This tagging method is expected to have a higher efficiency for stopping muons tagging, since they enter within the ID through the chimney, which is surrounded by OV. However, it is expected to have a smaller efficiency related to fast neutrons tagging, since their progenitor muon can interact far away from the detector and then miss OV.

In addition, Outer Veto was not available for the full data taking period, therefore a normalisation factor had to be applied to scale the lifetime period.

6.2.1 Gadolinium selection

The official correlated background analysis for the Gadolinium selection was done using the Inner Veto tagging method. It consists in the selection of events passing all the selection cuts but the IV veto. Due to the low accidental rate in the Gadolinium analysis, the contamination of the $\gamma - \gamma$ events (cf. section 5.1.4) was expected to be negligible. The sample of events rejected only by IV veto was then expected to be a pure sample of the correlated background.

This analysis was performed with a new IBD candidates selection, using all standard cuts described in section 5.1.2, except the prompt-signal energy window and the muon ID energy definition, which were extended to 30 MeV. The effi-

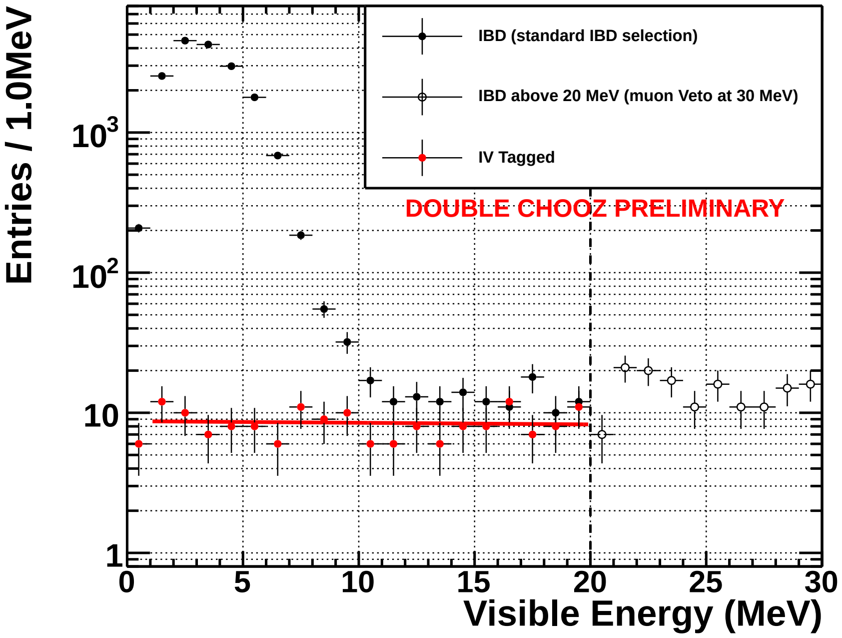


FIGURE 6.3: Prompt energy distribution with IBD selection. The black points show the IBD selection, the selection in [0.5, 20.0] MeV is the official Gd selection, in [20.0,30.0] MeV the Muon tagging is relaxed (energy deposit in ID $>$ 30 MeV) from the official one (energy deposit in ID $>$ 20 MeV). The red points show the correlated background selection with the IV-tag method. The red line is the best fit of a linear function on the IV-tagged sample with a slope of 0.02 ± 0.11 events/MeV^2. IV-tagged events below 1 MeV are not used in the fit to avoid contamination from Compton scattering of γ's in the IV and ID.

ciency was computed in the [20, 30] MeV energy window, where a pure sample of correlated background is expected. This analysis determined that the correlated background shape was consistent with a flat shape. FIG. 6.3 shows the IV tagged sample normalized to the IBD candidates sample with the correlated background sample selected with $E_{vis} > 20$ MeV. This analysis allowed to compute a correlated background rate of:

$$R_{CorrBG} = 0.604 \pm 0.051 \; events \, per \, day \qquad (6.5)$$

The OV configuration was different during the OFF-OFF period (see section 2.2.1.3), the rate of correlated background is slightly different and was estimated to be:

$$R_{CorrBG,OFF-OFF} = 0.529 \pm 0.089 \; events\, per\, day \tag{6.6}$$

As it is detailed in the section 5.1.4.2, the IV veto rejects $\sim 13\%$ of the accidental background. This was not expected when the analysis presented above was performed. Even if the rate of accidental is very low in the Gadolinium analysis, it could have an impact on the measurement of the correlated background shape and rate. Therefore, a cross-check of the official analysis was performed. This analysis is one of the contributions of my thesis on the correlated background measurement.

In order to carry out this analysis, I developed a slightly different Inner Veto tagging method. This method aimed to select the correlated background with a lower accidental contamination. The analysis was performed with a sample of IBD candidates selected with all standard cuts of the Gadolinium selection, except the prompt-signal energy window and the muon ID energy definition which were extended to 45 MeV.

6.2.1.1 Cross-check with a new Inner Veto Tagging

The events rejected by IV veto are mainly correlated background, as the accidental background rate, and therefore the $\gamma-\gamma$ rate, is low in the Gadolinium analysis (cf. section 5.1.2). The number of accidental events rejected by the IV veto is shown in section 5.1.4.2 to be: $10\,515 \pm 37$ events with 2 000 time windows. Therefore, we expect the presence of 5.26 ± 0.02 accidental events in the sample of events rejecting by the IV veto. These accidental events are likely to be $\gamma-\gamma$ events. Thanks to the low inefficiency of IV veto ($\sim 0.2\%$), a negligible amount of events without ID-IV correlations is expected.

It should be possible to increase the purity of this sample using the expected difference of ID-IV correlations between the $\gamma-\gamma$ events and the correlated background events. $\overline{\nu}_e$ events are expected to show no correlation between ID and IV.

Fig. 6.4 shows the distance between ID and IV energy deposition for a pure sample of correlated background rejected by the IV veto (selected with $E > 12$ MeV) and for the sample of accidental background rejected by IV veto. The mean ID-IV distance is clearly different between these two samples. The mean distance is ~ 1.8 m for the accidental background sample and about 2.5 m for the sample of correlated background events rejected by the IV veto. This 0.7 m difference is consistent with the distance between the γ-catcher/Buffer border and the ν-target/γ-catcher border of ~ 0.65 m.

As expected, the accidental background sample has a mean distance at lower than the one of the correlated background sample. Therefore, it should possible to use the ID-IV space correlation in order to improve the purity of the sample.

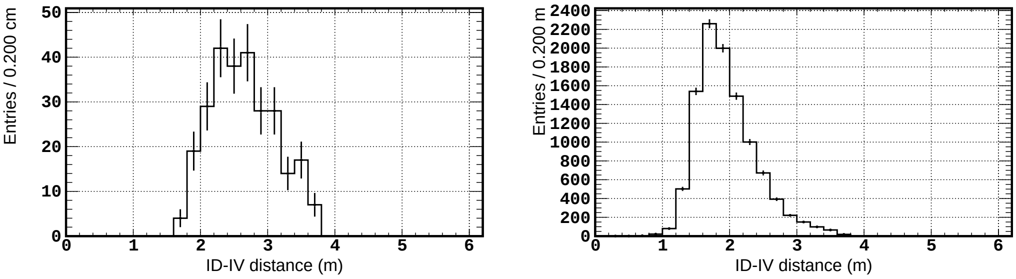


FIGURE 6.4: Distance between ID and IV energy deposition for events rejected by IV veto. *Left*: Pure correlated background sample, selected within the extended IBD sample with a cut on prompt energy: $E \in [12, 45]$ MeV. *Right*: Accidental background sample selected with the off-time window method.

Δd^{ID-IV} cut	IVT Correlated BG	IVT Accidentals BG	ε_{IVT}	ρ_{IVT}
default: < 3.7 m	269	5.26 ± 0.02	$38.99 \pm 1.86\%$	$98.56 \pm 0.01\%$
[1.5, 3.7] m	269	4.64 ± 0.02	$38.99 \pm 1.86\%$	$98.72 \pm 0.01\%$
[1.6, 3.7] m	268	4.18 ± 0.02	$38.90 \pm 1.86\%$	$98.84 \pm 0.01\%$
[1.7, 3.7] m	267	3.62 ± 0.02	$38.81 \pm 1.86\%$	$98.99 \pm 0.01\%$
[1.8, 3.7] m	263	3.05 ± 0.02	$38.45 \pm 1.86\%$	$99.13 \pm 0.01\%$
[1.9, 3.7] m	255	2.52 ± 0.02	$37.72 \pm 1.86\%$	$99.26 \pm 0.01\%$
[2.0, 3.7] m	244	2.05 ± 0.02	$36.69 \pm 1.86\%$	$99.37 \pm 0.01\%$

TABLE 6.1: Summary of the ID-IV distance cut tunning in order to maximize the number of IV tagged correlated background events and to minimize the number of IV tagged accidental ones. Correlated background events come from a subsample of IV tagged events, selected with $E > 12$ MeV. The efficiency was estimated between $[12, 45]$ MeV.

TAB. 6.1 summarizes the different cut values and their associated efficiency and purity. Looking at the best purity improvement over efficiency loss ratio, the analysis allowed to determine the best distance cut as [1.7, 3.7] m, with $\varepsilon_{IVT} = 38.81 \pm 1.86\%$ and $\rho_{IVT} = 98.99 \pm 0.01\%$.

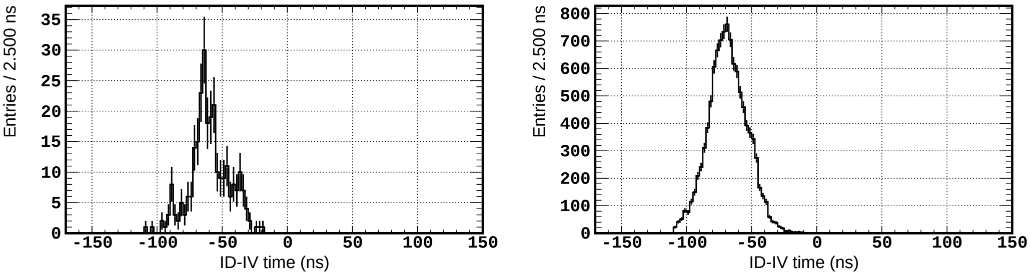


FIGURE 6.5: Time difference between ID and IV energy deposition for events rejected by IV veto. *Left*: Pure correlated background sample, selected within the extended IBD sample with a cut on prompt energy: $E \in [12, 45]$ MeV. *Right*: Accidental background sample selected with the off-time window method.

FIG. 6.5 shows the time difference between ID and IV energy deposition for the two samples. However, it demonstrates no visible difference between correlated background and $\gamma - \gamma$ events. From the difference between the mean ID-IV distances of the correlated background events and of the accidental background events, a difference of $\sim$ 2 ns is expected. The time resolution of Double Choz is 2 ns, therefore it is not expected to see any differences. Therefore, we kept the cut defined for the IV veto.

For this analysis, we selected the IV tagged sample with the following cuts:

- Pass all Gadolinium selection cuts but the IV veto
- Prompt signal in coincidence with an activation of the IV with:
  - Minimal IV tagging condition (cf. section 5.1.4)
  - $\Delta d^{ID-IV} \in [1.7, 3.7]$ m
  - $\Delta t_{QW}^{ID-IV} \in [-10, -110]$ ns

These cuts lead to an efficiency of $\varepsilon_{IVT} = 38.81 \pm 1.86\%$ and a purity of $\rho_{IVT} = 98.99 \pm 0.01\%$. The normalization factor was computed with EQ. 6.1 to be $c_{IVT} = 63.42 \pm 2.35\%$.

Using the IV tagged sample, two correlated background shapes were investigated: a flat line and a linear slope. FIG. 6.6 shows the fit result for the IV tagged sample.

The IVT fit showed a small preference for the linear slope compared. However, the uncertainties on the fit do not allow to give a clear conclusion on the correlated background shape. This conclusion was also the one reached during the official

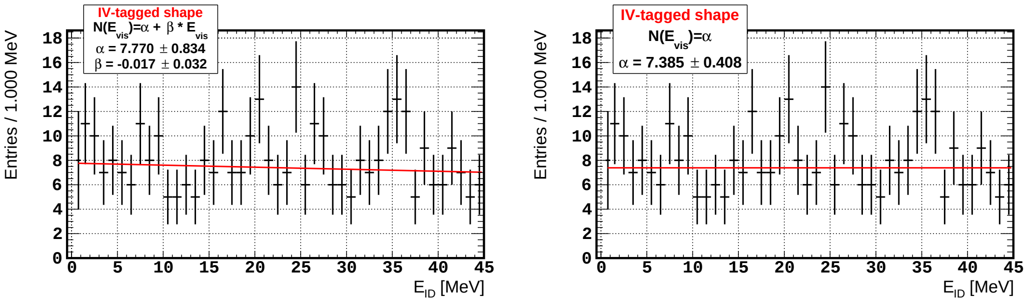


FIGURE 6.6: IV tagged sample spectrum fit with a linear slope (*left*, $\chi^2/ndf = 30.08/43$) and with a flat line (*right*, $\chi^2/ndf = 30.36/44$). The linear slope is compatible with a flat line.

correlated background analysis for the Gadolinium selection. Then, a flat shape was used to compute the correlated background rate using the EQ. 6.4:

$$\begin{aligned} R_{CorrBG} &= 0.494 \pm 0.036(fit) \pm 0.019(eff) \pm 0.001(purity) \\ &= 0.494 \pm 0.041 \text{ events per day} \end{aligned} \tag{6.7}$$

This rate is consistent with the official one, within 2.2σ, and with the one fitted by the rate+shapefit, within 2.0σ (cf. section 5.5.1.1). This is then a demonstration that the impact of the $\gamma - \gamma$ events on the correlated background selection with the IV tagging method in negligible.

We also performed the analysis to compute the correlated background rate during the OFF-OFF period. This analysis was done using only runs where the OV was in the same configuration as during the OFF-OFF period. It led to the following rate, fully consistent with the one from the official analysis:

$$R_{CorrBG}^{OFF-OFF} = 0.484 \pm 0.062 \text{ events per day} \tag{6.8}$$

6.2.2 Hydrogen selection

The contributions to the correlated background in the Hydrogen selection are the same as in the Gadolinium selection: Fast Neutrons and Stopping Muons. However, references [114] showed that the number of Michel electrons around 3 MeV is lower than the one around 8 MeV. Therefore, thanks to the delayed signal energy window, we expect a smaller contamination of stopping muons. Additionally, as in the Gadolinium selection, the FV veto strongly suppresses the stopping muons. Analysis were done and showed a contamination of ~ 0.02 events per day [182, 183].

I was responsible for this correlated background shape and rate measurements. In this analysis we used an IBD sample with a prompt signal energy window extended to 60 MeV. The original goal was to extend the energy window to 100 MeV. However, due an internal cut inside the offline reconstruction algorithms, the reconstruction is performed only for events with $E_{vis} < 60$ MeV. For the future analysis, this internal cut should be removed, allowing to investigate the correlated background shape at higher energy.

We used two methods to select a pure correlated background events sample: the Inner Veto tagging and the Outer Veto tagging. The efficiencies of these tagging methods were estimated between 20 to 60 MeV. We chose to apply [20, 60] MeV in order to keep consistency with the other correlated background measurements performed at the same time [181]. The shift from 12 MeV to 20 MeV for the efficiency computation has a negligible impact on the final rate: ±0.01 events per day.

6.2.2.1 Inner Veto tagging

The IV tagging method measures the background spectrum shape using events which pass all Hydrogen selection cuts except IV veto (Prompt) (see section 5.1.4.3), and which are selected by a set of specific cuts. Since the current IV veto (Prompt) rejects many accidental background events, and since the Hydrogen analysis is very different than the Gadolinium analysis, we had to tune the Inner Veto tagging cuts.

The Inner Veto tagging cuts were defined in order to select a pure sample of correlated background with a negligible accidental background contamination. The IV veto (Delayed) cuts, used for the Hydrogen analysis, reject mainly accidental background, and therefore cannot be used to select the sample of correlated background events. In order to reject the accidental background events in the IV tagging sample, we defined a subsample of events rejected by the IV veto (Prompt), using tighter cuts.

Thanks to the low inefficiency of the IV veto (Prompt) ($< 0.2\%$) we expected a negligible amount of events without ID-IV correlations.

In order to reject the high accidental background events from the sample, we applied a cut on the energy deposited in the IV. Fig. 6.7 shows the correlation between the energy deposited in the ID and the in IV for events rejected by the IV veto. A triangular shape at low energy indicates an energy conservation between both volumes. As explained in section 5.1.4.3, we interpreted this feature as due to $\gamma - \gamma$ events. The energy conservation feature remains until $4 - 5$ MeV in the Inner Veto. Therefore in order to reject most of the $\gamma - \gamma$ events from the sample, a E_{IV} cut was defined. The value was arbitrary set to 6 MeV to ensure a good rejection. Since $E_{IV} = Q_{IV}/2200$, we have:

- $Q_{IV} > 6 \times 2200$ a.u.

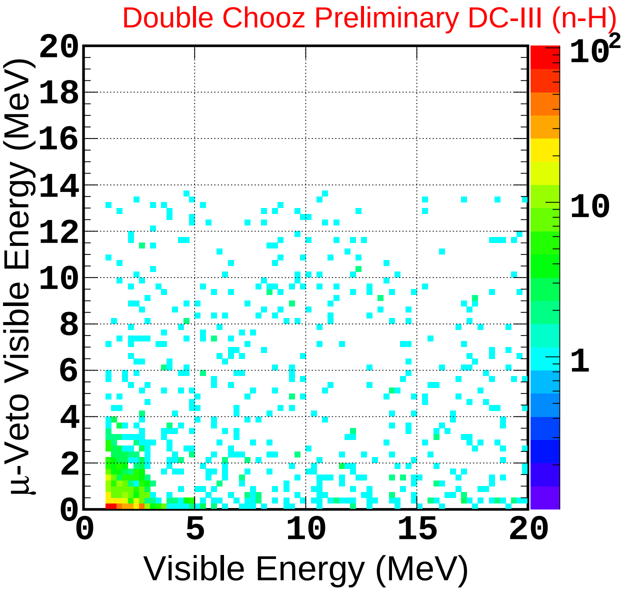


FIGURE 6.7: Energy correlation between ID energy deposition and IV energy deposition for events rejected by IV veto applied on the prompt and the delayed signals. The triangular shape at low energy indicates an energy conservation between both volume, which could be the signature of an interaction in the IV and in the ID by the same particle. It is interpreted as due to $\gamma - \gamma$ events. Events without energy conversation between IV and ID are expected to be fast neutrons and stopping muons.

Therefore, we measured a low accidental background contamination in the sample of events selected by IV veto (Prompt) with this additional cut. We selected the accidental events passing all the selection cuts but the IV veto (Prompt), and selected by the cuts of the IV veto (Prompt) and this new one. Taking into account the number of time windows (200), we measured the accidental background contamination to be 3.27 ± 0.01 events.

The ID-IV correlation definitions have also been modified for the correlated background selection, as the one used in IV veto (Prompt) was defined to include accidental background events. As for the Gadolinium analysis (see above), it was expected that the ID-IV correlations for accidental background events are different than the ones for correlated background. However, since the Hydrogen and the Gadolinium analyses are very different, the differences between the correlations of the accidental background and of the correlated background are not expected to be the same than in the Gadolinium selection.

The ID-IV correlation cuts were defined after the application of the accidental

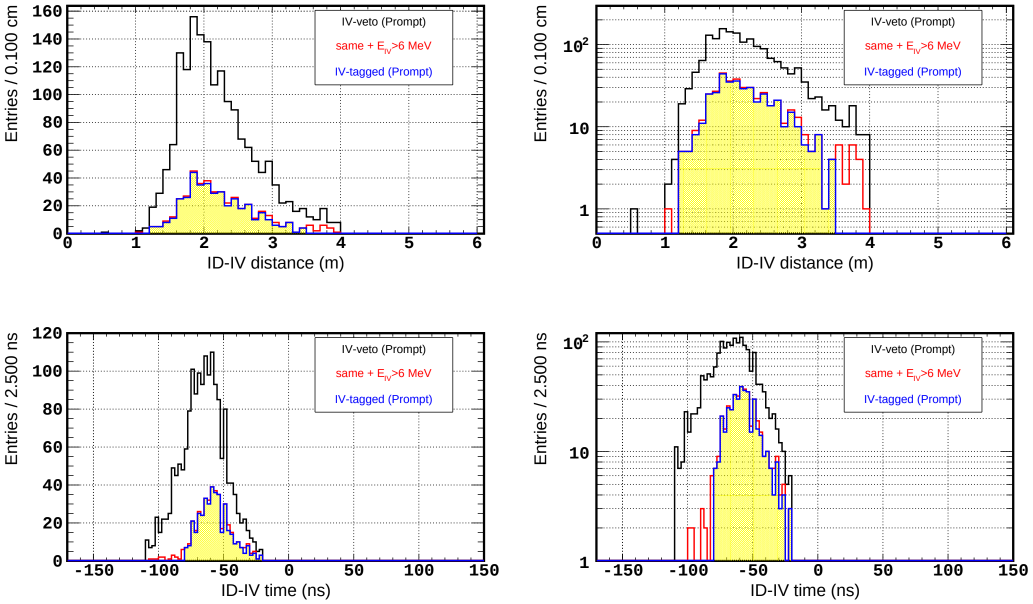


FIGURE 6.8: ID-IV space (*Top*) and time (*Bottom*) correlations for IV tagged events. Plots on the left are in linear scale, plots on the right are in logarithmic scale. Black line shows events rejected by IV veto (IV veto sample), red line shows the remaining subsample of IV veto events after application of the $E_{IV} > 6$ MeV cut. Blue line shows the events selected by IV tagged cuts.

background rejection cut $E_{IV} > 6$ MeV. FIG. 6.8 shows the ID-IV space and time correlations. The application of the accidental background rejection cut reject a high amount of events. Therefore, in order to keep a high statistical power, we defined relaxed ID-IV correlation cuts.

We applied a similar method than the one used to define the ID-IV correlation cuts of the IV veto method: We looked a the intersection between a flat component, due to events without genuine correlations between the IV and the ID signals, and the peak, due to events with genuine correlations between the IV and the ID signals. This method was applied on the samples shown in FIG. 6.8.

After application of the accidental background rejection cut, the ID-IV time correlation figure shows a tighter distribution than before. Consequently, we defined the cuts to select only the peak in shown in the figures. This led us to the following ID-IV correlation cuts:

- $\Delta d^{ID-IV} \in [1.1, 3.5]$ m
- $\Delta t_{QW}^{ID-IV} \in [-20, -80]$ ns

With these cuts the purity of the tagged sample was estimated to be $\rho_{IVT} = 98.92 \pm 0.30\%$. This measurement was done by looking at the number of events selected by the Inner Veto tagged within the accidental background sample of the Hydrogen analysis.

As explained above, the efficiency was estimated between $[20, 60]$ MeV with EQ. 6.2 and give $\varepsilon_{IVT} = 19.116 \pm 1.219\%$. Therefore, using EQ. 6.1, the normalization faction was computed to be $c_{IVT} = 23.634 \pm 1.465\%$.

The $\varepsilon_{IVT} = 19.116 \pm 1.219\%$ efficiency is low compared to the Gadolinium IV tagging one. This low efficiency is due to the $E_{IV} > 6$ MeV. A rought estimatation of the fraction of correlated background events can be performed by assuming that the IV correlated background energy spectrum is flat between $[0, 14]$ MeV, and that the events with $E_{IV} > 6$ MeV are a pure sample of correlated background. Therefore, the ratio of correlated background events rejected can be estimated to be:

$$ratio = \frac{6\,MeV}{14\,MeV} \backsimeq 43\% \tag{6.9}$$

Please note that both assumptions are likely to be wrong but can be use for a rought estimatation.

The sample of events selected by the Inner Veto tagging method was used to measure the shape and the rate of the correlated background, this is detailed in section 6.2.2.3.

6.2.2.2 Outer Veto Tagging

Outer Veto veto rejects almost only correlated background, with a negligible number of rejected accidental background events. We measure with the accidental background sample that the Outer Veto veto was rejecting < 0.0001 event. Therefore, it was not necessary to tune OV tagging cuts to improve its purity. We selected the OV tagged sample with the following cuts:

- Pass all Hydrogen selection cuts but OV veto
- Prompt signal in coincidence with an activation of the OV

The OV tagging purity was assumed to be $\rho_{OVT} = 100.00 \pm 0.01\%$. Using EQ. 6.2, its efficiency was computed between $[20, 60]$ MeV to be $\varepsilon_{OVT} = 32.48 \pm 1.33\%$, leading to the measurement of a normalization factor of $c_{OVT} = 48.10 \pm 1.73\%$.

The sample of events selected by the Outer Veto tagging method was used to measure the shape and the rate of the correlated background, this is detailed in the following section (6.2.2.3).

6.2.2.3 Spectrum shape and rate measurement

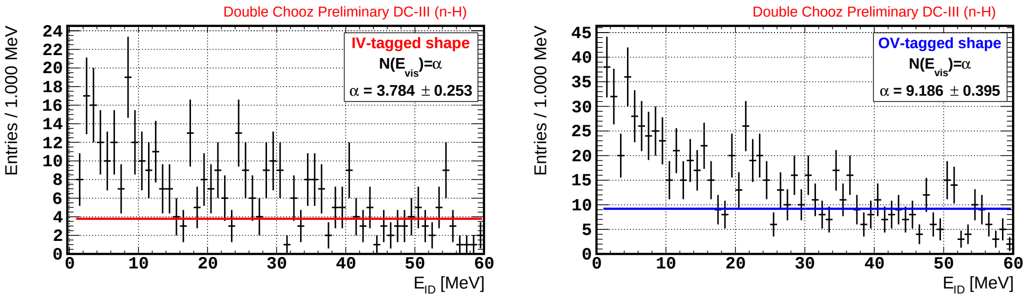


FIGURE 6.9: Spectrum of the events tagged by IV tag (*left*, $\chi^2/ndf = 172.6/59$) and OV tag (*right*, $\chi^2/ndf = 276/59$) with a flat fit.

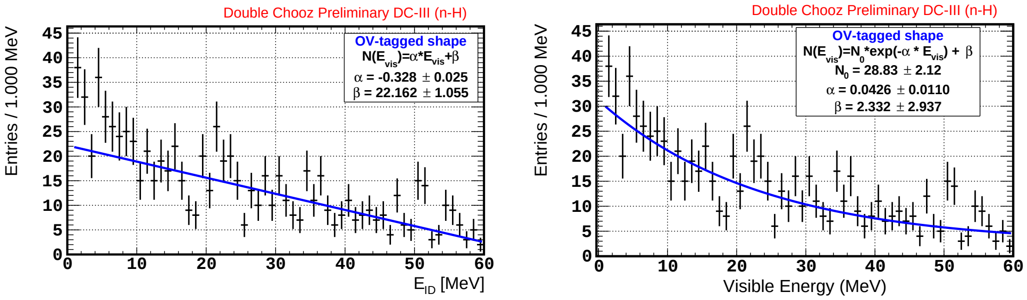


FIGURE 6.10: Spectrum of the events tagged by OV tag, fitted with a linear slope (*left*, $\chi^2/ndf = 99.96/57$) and with an exponential slope (*right*, $\chi^2/ndf = 85.65/57$).

We tested three different shapes of correlated background: a flat line, a linear slope and an exponential slope (cf. EQ. 6.10). FIG. 6.9 shows the results for the flat line, with IV tagged and OV tagged samples. A flat shape is then obviously rejected. FIG. 6.11 shows the fit results for IV tagged sample and FIG. 6.10 for the OV tagged sample. Fits showed a clear preference for the exponential shape. FIG. 6.12 shows the correlated background exponential shape normalized to the IBD spectrum.

$$Shape(x) = N_0 * Exp(-\alpha * x) + \beta \tag{6.10}$$

We computed the correlated background rate using the exponential shape by applying EQ. 6.4, for IV tagged sample and OV tagged sample:

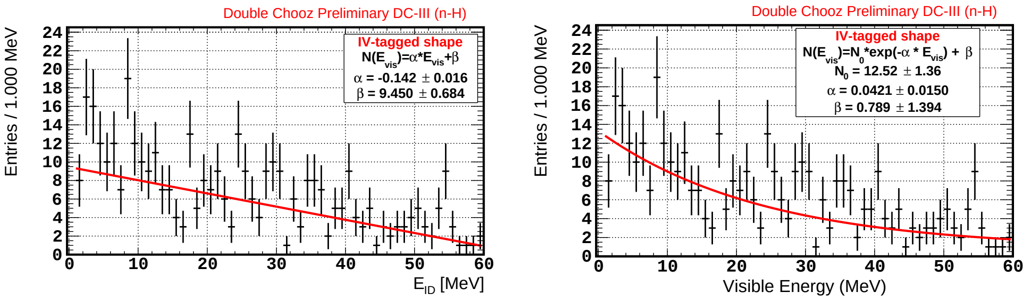


FIGURE 6.11: Spectrum of the events tagged by IV tag, fitted with a linear slope (*left*, $\chi^2/ndf = 92.92/58$) and with an exponential slope (*right*, $\chi^2/ndf = 85.08/57$).

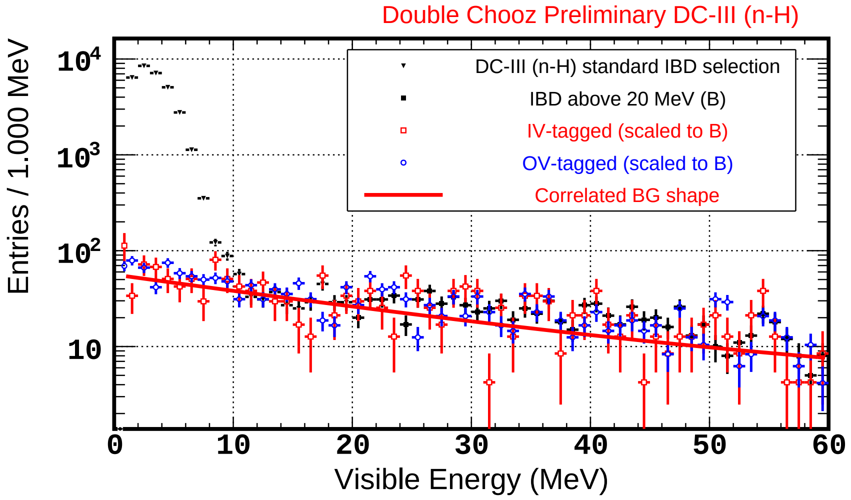


FIGURE 6.12: Prompt energy distribution with IBD selection. The black points showed the IBD selection, the selection in $[1.0, 20.0]$ MeV is the H-III one, in $[20.0, 60.0]$ MeV the Muon tagging is relaxed ($E_{vis} > 60$ MeV) from the Hydrogen official one ($E_{vis} > 20$ MeV). The colored points show the correlated background selection with the IV-tag method (*left*) and with the OV-tag method (*right*), scaled to the IBD selection in $[20.0, 60.0]$ MeV range. The line is the exponential shape computed with the IV-tag method, normalized for $[20.0, 60.0]$ MeV range.

$$\begin{aligned} R_{CorrBG,IVT} &= 1.556 \pm 0.112(fit) \pm 0.097(eff) \pm 0.005(purity) \\ &= 1.556 \pm 0.148 \text{ events per day} \end{aligned} \tag{6.11}$$

$$\begin{aligned} R_{CorrBG,OVT} &= 1.810 \pm 0.089(fit) \pm 0.065(eff) \\ &= 1.810 \pm 0.110 \text{ events per day} \end{aligned} \tag{6.12}$$

The error is computed assuming that fit error and efficiency error are uncorrelated. IVT and OVT rates are in agreement within 2σ. This large discrepancy was interpreted as a bias due to the expected high stopping muon population in the OV tagged sample. This supposition was validated by applying an additional cut to the OV tagging method, in order to avoid stopping muon. Using the fact that stopping muons have a very short time-life ($\sim 2.2\mu$s) with respect to the fast neutron thermalization time, we applied a cut on the time correlation between prompt and delayed signals: $\Delta t_{Prompt-Delayed} > 10\mu$s. From reference [180], this cut allows a stopping muon rejection of $99^{+1}_{-17}\%$. The measured rate with this modified OV tagging method was $R_{CorrBG,OVT\,10\mu s} = 1.647 \pm 0.106$ events per day, consistent within 1σ with the IV tag rate. The IV tag rate was taken as default value, as the Inner Veto allows a 4π coverage. Therefore it allows to better take into account the main correlated background component, the fast neutrons, than the Outer Veto. in comparison to the Outer Veto.

As for Gadolinium, we measured the rate for the OFF-OFF period with the IV tagging method:

$$\begin{aligned} Rate^{OFF-OFF}_{CorrBG} &= 1.453 \pm 0.148(fit) \pm 0.130(efficiency) \pm 0.005(purity) \\ &= 1.453 \pm 0.197 \text{ events per day} \end{aligned} \tag{6.13}$$

6.3 Summary

In this chapter, we presented the methods developed to measure the shape and rate of the correlated background in the Double Chooz experiment: the Inner Veto tagging method and the Outer Veto tagging method. The IV tagging method is the main method for this analysis since the first θ_{13} measurement publication of Double Chooz. This method was improved by the definition of the new cuts in order to reduce the accidental background contamination in the correlated background sample.

For the Gadolinium analysis, with the new cuts proposed, we found a correlated background rate of 0.494 ± 0.041 events per day, smaller than the one found by the official analysis, but consistent within 2σ. This cross-check allows to show that the contamination of the $\gamma - \gamma$ events within the correlated background sample was not biasing the result of the measurement. Both the official analysis and the cross-check performed found a small preference for a linear slope shape, but were not able to give a clear conclusion.

For the Hydrogen analysis, the result of the presented analysis was the official measurement of the correlated background. A correlated background rate of 1.556 ± 0.148 events per day has been measured. Here, with both the IV tagging method and the OV tagging method, an exponential shape was found for the correlated background.

Conclusions

The contributions of this thesis to the Double Chooz analyses have been presented. The Double Chooz experiment, with only one detector running and thanks to an excellent understanding of this detector, has been able to improve its analyses over the years.

The development on the energy reconstruction, which has been one of the contributions of this thesis, allowed to reduce the associated uncertainty by a factor ~ 2. I worked on the Uniformity correction and developed the charge non-linearity correction. In the Gadolinium analysis, the uncertainty associated with the Uniformity correction was reduced by $\sim 16\%$ with respect to the previous analysis. The addition of the charge non-linearity correction allowed to reduce the asymmetry between data and MC from $\sim 1\%$ to $\sim 0\%$ on the energy of the Gadolinium neutron capture for fast neutron. For the Hydrogen analysis, we had to adapt the different methods and were able to improve the energy reconstruction with respect to the previous Hydrogen analysis, but also with respect to the Gadolinium analysis. The Hydrogen analysis uncertainty associated with the Uniformity correction has been reduced by $\sim 81\%$ with respect to the previous Hydrogen analysis and by $\sim 30\%$ with respect to the Gadolinium analysis. It was also shown that with the method developed for the Hydrogen analysis, the Uniformity correction uncertainty would also have been reduced for the Gadolinium analysis.

Innovative background rejection methods have been developed since the previous analyses, leading to a reduction of the background systematics due to the accidental background, to the correlated background or to the cosmogenic nuclei background. In the Gadolinium analysis the total background systematic budget has been reduced by a factor 2. In the Hydrogen analysis, however, the collaboration has been able to reduce the accidental background by a factor ~ 17, while the other background systematics were reduced by a factor ~ 2. My thesis contributions to these improvements are the development of the Inner Veto veto method and the measurement of the remaining correlated background. The IV veto allowed to reject both correlated and accidental backgrounds. In the Hydrogen analysis, this cut is one of the main accidental background rejectors, together with the ANN cut. IV veto alone rejects $\sim 27\%$ of the accidental background. The measurement of the remaining correlated background with the IV tagging method, allowed to have a measurement of the shape of the correlated background. This analysis led to the measurement of a flat correlated background shape for the Gadolinium analysis and of an exponential shape for the Hydrogen analysis. The difference between these both shapes has been discussed in this thesis, and a geometric origin has been suggested. Both these methods, IV veto and IV tagging, are an unique development of the Double Chooz experiment, since the other reactor experiments, with Cerenkov water veto, are not able to detect fast neutrons with enough efficiency.

By the addition of the improvements on the detection systematics, Double Chooz has been able to provide new θ_{13} measurements. In the Gadolinium analysis, the total systematics were reduced by a factor 1.25. Therefore, using a rate+shape fit, we were able to measure $\sin^2 2\theta_{13} = 0.090^{+0.032}_{-0.029}$. In the Hydrogen analysis, the systematics were reduced to a level equivalent to the Gadolinium systematics. Therefore, using the RRM fit, we were able to measure $\sin^2 2\theta_{13} = 0.098^{+0.038}_{-0.039}$.

Finally, the Double Chooz collaboration combined both Gadolinium and Hydrogen analyses results, taking advantage of the high statistical power of the Hydrogen sample. This led to a measurement of $\sin^2 2\theta_{13} = 0.090 \pm 0.033$.

In this thesis, several perspectives for the future analyses of Double Chooz are suggested. Regarding the energy reconstruction, my studies highlighted the presence of a bias in the sample of fast neutrons selected with a $\Delta t_\mu > 50$ μs. We showed that the increase of the Δt_μ cut to 100 μs allows to avoid this bias. It was interpreted as a remaing electronics overload, 50 μs after the muon passage. For future analyses, it would be necessary to increase the Δt_μ cut used to select the fast neutron sample. 100 μs removed 80% of the statistics, a lower value should allow to avoid this bias and to keep a high statistical power.

For the measurement of the correlated background with the IV tagging method, we showed that the fast neutron penetration in the detector has a dependence on the fast neutron energy. Therefore, if the prompt energy window used to select the correlated background is further extended, one should be careful with the ID-IV correlation distance cut. In addition, the hypothesis made to explain the correlated background shape should be validated. This could be done using Monte Carlo simulations able to correctly simulate fast neutron interactions.

A new pulse reconstruction algorithm, *RecoZoR*, is still in development. It aims to improve the linearity of the energy reconstruction, together with an improvement of the resolution on the reconstructed number of photoelectrons hitting the photomulpliers and on their arrival times. Several issues need to be corrected before using it. However, its preliminary results are promising, with an improvement by a factor ~ 2 on both number of PE and time resolution with respect to the current pulse reconstruction method. This allows to consider its use in the near future.. Thanks to the improvements on the time resolution and on the resolution of the number of PE, an overall improvement of the Double Chooz analysis can be expected. Investigation are ongoing in order to use this improvement to directly correct for the non-Uniformity of the detector response, or to use it in order to discriminate the positron and the electron signals in the detector if the time resolution is further increased.

The development of RecoZoR and the good precision of the last Hydrogen analysis of Double Chooz could be used by future experiments. The JUNO detector will consist in a large volume of liquid scintillator not-loaded with Gadolinium, therefore the JUNO analysis to select inverse β decay events is going to be a Hydrogen analysis. The results of Double Chooz with the Hydrogen analysis demonstrate the consistancy of this strategy. In addition, the JUNO collaboration is planning to add several “small” PMTs in its detector. This could lead to the use of a shape dependent method like RecoZoR in order to perform a “photon counting” analysis, i.e. to count each photoelectron seperatly.

The near detector of Double Chooz is taking data since December 2014. New IBD analyses are going to be performed in this second phase of the experiment. The use of both detector will allow to cancel much of the detection systematics in the analyses. In addition, it will be possible to perform the rate+shape comparison between the far detector and the near detector in order to measure the $\overline{\nu}_e$ oscillations without using a MC simulation. Therefore, the precision of the θ_{13} measurement will be improved. Double Chooz will also beneficiate from a long period with only one reactor on, this situation, with the near detector running, could be take as an opportunity to better understand the $[4-6]$ MeV spectrum distortion. The studies presented in this thesis could be used, with some tunings, during these upcoming analyses.

Annex: Contributions of this thesis

Over the three years of this thesis, several contributions have been provided to the Double Chooz experiment. Some other contributions, which are not detailled in this document, were also performed for the SoLi∂ experiment. This annex summarizes and presents the whole contributions performed during this thesis.

Double Chooz:

- Development of a vertex reconstruction algorithm for the Inner Veto. This algorithm was developped to use several artifical neural networks in order to deal with the complex geometry of the Inner Veto. Before its development, there was no possibility to reconstruct the vertex position inside the Inner Veto. This algorithm, called Neural Network Inner Veto reconstruction (NNIVReco), allows to have a resolution of $\sim$ 36 cm on the reconstructed position in the Inner Veto. It is detailled in section 2.7.2.

- Maintenance of the online system performing the onsite data files handling and transfer. This online system was also adapted to be use for the Near Detector online system and to take into account the modifications of the Far Detector online system. It is detailled in section 2.4.2.

- Development of the Inner Veto veto, a rejection method originally aiming to reject the fast neutrons and the stopping muons background of the Double Chooz analysis. This method was developped for both Gadolinium and Hydrogen analyses. It was shown that the IV veto allows to reject 13% of the accidental background in the Gadolinium analysis and 27% in the Hydrogen analysis. It is detailled in section 5.1.4.

- Various studies were performed with the radioactive calibration sources runs during this thesis:

  - Contributions to the determination of the light yield of the liquids and of the Birk's constant (representing the quenching effect), used in section 3.2.6.2.
  - Cross-checks of the PE to MeV conversion factor presented in section 3.2.3.
  - Determination of the correction of the charge non-linearity of the detector response, using the Hydrogen and the Gadolinium neutron capture in ^{252}Cf runs. This study is detailled in section 3.2.6.1.
  - Determination of the light non-linearity correction for the Hydrogen analysis. This study is detailled in section 3.2.8.2.

- Determination of the Uniformity correction of the detector response using a selection of fast neutrons captures on Hydrogen and Gadolinium nuclei. The Uniformity correction was determinated for both Gadolinium and Hydrogen analyses. It is detailled in section 3.2.4 for the Gadolinium analysis and in section 3.2.8.1 for the Hydrogen analysis. For the Hydrogen analysis, a new method was developped.

- Investigation on a bias due to the electronics overload following the high energy depositions in the detector due to muons. Investigations with the so-called *After Muon analysis*, detailled in section 3.2.4.2, led to the conclusion of the absence of bias. Whereas the analysis performed to determine the Uniformity correction in the Hydrogen analysis highlighted the presence of the bias. This discrepancy between the conclusions has been discussed in section 3.2.8.1.

- Determination of the energy reconstruction resolution and validity, using both callibration source runs and the fast neutron selection. This study is detailled in section 3.2.7.

- Fast neutrons analysis for the measurement of the remaining correlated background in the IBD selection. This analysis was performed for both Gadolinium and Hydrogen analyses and it detailled in chapter 6.

- Development of a new calorimetry method using the shape of the pulses in order to improve the resolution on the reconstructed number of photoelectrons hitting a photomultiplier and on their reconstructed arrival time. This new method is presented in chapter 4.

SoLi∂:

- Contributions to the commissionning of the detector in January 2015. A study on the rate per channel was performed in order to tune the detection thresholds used by the experiment and to tune the gain equalization of the detector electronics.
- Development of an analysis framework prototype for the experiment, called *NESPRESSO*. It has been used during the commissionning and a proposal was made.
- Quality checks of the channel response with *NESPRESSO*.

Bibliography

[1] W. Pauli, *"Offener Brief an die Gruppe der Radioaktiven bei der Gauvereins-Tagung zu Tubingen"* (1930)

[2] F. L. Wilson, *"Fermi's Theory of Beta Decay"*, Am.J.Phys.36, 1150 (1968), translation of E. Fermi, *"Versuch einer Theorie der beta-Strahlen. I"*, Zeitschrift fr Physik, vol. 88 (1934)

[3] F. Reines and C. L. Cowan, Jr., *"Detection of the Free Neutrino"*, Phys.Rev.92, 830 (1953)

[4] F. Reines and C. L. Cowan, Jr., *"Detection of the Free Neutrino: A Confirmation"*, Science, Vol. 124 no. 3212 pp. 103-104 (1956)

[5] B. Pontecorvo, J. Exptl. Theoret. Phys. (U. S.S.R.) 37, 1751 (1959) (translation: Soviet Phys. JETP 10, 1236 (1960)).

[6] M. Schwartz, *"Feasibility of Using High-Energy Neutrinos to Study the Weak Interactions"*, Phys.Rev.Lett.4, 306 (1960)

[7] G. Danby et al., *"Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos"*, Phys.Rev.Lett.9, 36 (1962)

[8] E. J. Konopinski and H. M. Mahmoud, *"The Universal Fermi Interaction"*, Phys.Rev.92, 1045 (1953)

[9] DONUT Collaboration, *"Observation of Tau Neutrino Interactions"*, Phys.Lett.B504:218-224 (2001)

[10] Martin L. Perl et al., *"Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation"*, Phys.Rev.Lett.35, 1489 (1975)

[11] The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, the OPAL Collaboration, the SLD Collaboration, the LEP Electroweak Working Group, the SLD electroweak, heavy

flavour groups, *"Precision Electroweak Measurements on the Z Resonance"*, Phys.Rept.427:257-454 (2006)

[12] Raymond Davis, Jr., Don S. Harmer, and Kenneth C. Hoffman, *"Search for Neutrinos from the Sun"*, Phys.Rev.Lett.20, 1205 (1968)

[13] W. Hampel et al. (GALLEX Collaboration), *"Final results of the Cr-51 neutrino source experiments in GALLEX"*, Phys.Lett.B420 114-126 (1998)

[14] SAGE Collaboration, *"Measurement of the solar neutrino capture rate by SAGE and implications for neutrino oscillations in vacuum"*, Phys.Rev.Lett.83:4686-4689 (1999)

[15] K.S. Hirata et al., *"Experimental study of the atmospheric neutrino flux"*, Phys.Lett.B205, 416-420 (1988)

[16] Y. Fukuda et al. (Super-Kamiokande Collaboration), *"Evidence for oscillation of atmospheric neutrinos"*, Phys.Rev.Lett.81:1562-1567 (1988)

[17] Y. Ashie et al. (Super-Kamiokande Collaboration), *"Measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I"*, Phys.Rev.D71, 112005 (2005)

[18] B. Pontecorvo, *"Mesonium and anti-mesonium"*, Zh.Eksp.Teor.Fiz.33: 549551 (1957), reproduced and translated in Sov.Phys.JETP6: 429 (1957)

[19] Z. Maki, M. Nakagawa, and S. Sakata, *"Remarks on the Unified Model of Elementary Particles"*, Progress of Theoretical Physics 28 (5): 870 (1962)

[20] Q. R. Ahmad et al., *"Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory"*, Phys. Rev. Lett. 89, 011301 (2002)

[21] S.L. Glashow, Nucl.Phys.22 (1961), p. 579

[22] S. Weinberg, Phys.Rev.Lett.19 (1967), p. 1264

[23] A. Salam, N. Svartholm (Ed.), Proc. 8-th Nobel Symp., Almquist and Wiksell, Stockholm (1968), p.367

[24] Henrik Bohr, H.B. Nielsen, *"Hadron production from a boiling quark soup: A thermodynamical quark model predicting particle ratios in hadronic collisions"*, Nuclear Physics B128, 275293 (1977)

[25] Marco Drewes, *"The Phenomenology of Right Handed Neutrinos"*, Int.J.Mod.Phys.E, 22, 1330019 (2013)

[26] B. Pontecorvo, *"Neutrino Experiments and the Problem of Conservation of Leptonic Charge"*, Zh.Eksp.Teor.Fiz.53: 1717 (1967), reproduced and translated in B. Pontecorvo, *"Neutrino Experiments and the Problem of Conservation of Leptonic Charge"*, Sov.Phys.JETP26:984 , B. (1968)

[27] L. Wolfenstein, *"Neutrino oscillations in matter"*, Phys. Rev. D 17, 2369 (1978)

[28] E. G. Adelberger et al., *"Solar fusion cross sections II: the pp chain and CNO cycles"*, Rev.Mod.Phys.83, 195 (2011)

[29] A. Gando et al. (KamLAND Collaboration), *"Reactor On-Off Antineutrino Measurement with KamLAND"*, Phys.Rev.D88, 033001 (2013)

[30] A. Gando et al. (KamLAND Collaboration), *"Constraints on θ_{13} from A Three-Flavor Oscillation Analysis of Reactor Antineutrinos at KamLAND"*, Phys.Rev.D83, 052002 (2011)

[31] Y. Fukuda et al. (Super-Kamiokande Collaboration), *"Measurement of a small atmospheric ν_μ/ν_e ratio"*, Phys.Lett.B433, 9 (1998)

[32] M.G. Aartsen et al. (IceCube Collaboration), *"Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data"*, Phys.Rev.D91, 072004 (2015)

[33] P. Adamson et al. (The MINOS Collaboration), *"Measurements of atmospheric neutrinos and antineutrinos in the MINOS far detector"*, Phys.Rev.D86, 052007 (2012)

[34] D. Ayres et al. (The NOvA Collaboration), *"NOvA Proposal to Build a 30 Kiloton Off-Axis Detector to Study Neutrino Oscillations in the Fermilab NuMI Beamline"*, Fermilab-Proposal-0929 (2005)

[35] K. Abe et al. (T2K Collaboration), *"The T2K Experiment"*, Nucl. Instrum. Meth. A 659, 106 (2011)

[36] P. Adamson et al. (The MINOS Collaboration), *"Combined Analysis of ν_μ Disappearance and $\nu_{mu} \rightarrow \nu_e$ Appearance in MINOS Using Accelerator and Atmospheric Neutrinos"*, Phys.Rev.Lett.112, 191801 (2014)

[37] K. Abe et al. (T2K Collaboration), *"Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam"*, Phys.Rev.Lett.112, 181801 (2014)

[38] K. Abe et al. (T2K Collaboration), *"Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6E20 protons on target"*, Phys.Rev.D91, 072010 (2015)

[39] A. Kaboth for the T2K Collaboration, *"First Antineutrino Oscillation Results from T2K"*, KEK High Energy Accelerator Research Organization seminar (May 18th 2015)

[40] K. Duffy for the T2K Collaboration, *"Results and Prospects from T2K"*, NuFact 2015 conference (August 2015)

[41] M. Sanchez for the NOνA Collaboration, *"Results and Prospects from NOvA"*, NuFact 2015 conference (August 2015)

[42] J.P. Athayde Marcondes de André for the IceCube and PINGU Collaborations, *"IceCube/PINGU"*, NuFact 2015 conference (August 2015)

[43] A. Himmel (Super-Kamiokande Collaboration), *"Recent Atmospheric Neutrino Results from Super-Kamiokande"*, AIP Conf. Proc. 1604, 345 (2014)

[44] M. Apollonio (CHOOZ Collaboration), *"Search for neutrino oscillations on a long base-line at the CHOOZ nuclear power station"*, Eur.Phys.J.C27:331-374 (2003)

[45] F. Ardellier et al, *"Double Chooz: A Search for the Neutrino Mixing Angle θ_{13}"*, arXiv:hep-ex/0606025 (2006)

[46] Daya Bay Collaboration, *"A Precision Measurement of the Neutrino Mixing Angle θ_{13} using Reactor Antineutrinos at Daya Bay"*, BNL-77369-2006-IR, LBNL-62137, TUHEP-EX-06-003 (Jan 2007)

[47] Y. Abe et al. (Double Chooz Collaboration), *"Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment"*, Phys.Rev.Lett.108, 131801 (2012)

[48] F. P. An et al. (Daya Bay Collaboration), *"Observation of electron-antineutrino disappearance at Daya Bay"*, Phys.Rev.Lett.108, 171803 (2012)

[49] J.K. Ahn et al. (RENO collaboration), *"Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment"*, Phys.Rev.Lett.108, 191802 (2012)

[50] G. Mention Ph.D. thesis, *"Étude des sensibilités et bruits de fond de l'expérience Double Chooz pour la recherche du paramètre de mélange leptonique θ_{13}"* (2005)

[51] F. P. An et al. (Daya Bay Collaboration), *"Improved measurement of electron antineutrino disappearance at Daya Bay"*, ChinesePhys.C37, 011001 (2013)

[52] F. P. An et al. (Daya Bay Collaboration), *"Spectral Measurement of Electron Antineutrino Oscillation Amplitude and Frequency at Daya Bay"*, Phys.Rev.Lett.112 061801 (2014)

[53] W. Wang for the Daya Bay and JUNO Collaborations, *"From Daya Bay to Jiangmen"*, NuFact 2015 conference (August 2015)

[54] F. P. An et al. (Daya Bay Collaboration), *"A new measurement of antineutrino oscillation with the full detector configuration at Daya Bay"*, to be appear in Phys.Rev.Lett, arXiv:1505.03456 (2015)

[55] Y. Abe et al. (Double Chooz Collaboration), *"First Measurement of θ_{13} from Delayed Neutron Capture on Hydrogen in the Double Chooz Experiment"*, Phys.Lett.B723, 66-70 (2013)

[56] F. P. An et al. (Daya Bay Collaboration), *"Independent measurement of the neutrino mixing angle θ_{13} via neutron capture on hydrogen at Daya Bay"*, Phys.Rev.D90, 071101(R) (2014)

[57] K. Abe et al. (T2K Collaboration), *"Observation of Electron Neutrino Appearance in a Muon Neutrino Beam"*, Phys.Rev.Lett.112, 061802 (2014)

[58] R. Patterson for NOνA Collaboration, *"First oscillation results from NOνA"*, Fermilab Joint Experimental-Theoretical Seminar (August 6, 2015)

[59] K. Abe et al. (T2K Collaboration), *"Measurement of Neutrino Oscillation Parameters from Muon Neutrino Disappearance with an Off-axis Beam"*, Phys.Rev.Lett.111, 211803 (2013)

[60] J. Beringer et al. (Particle Data Group), *"Review of Particle Physics"*, Phys.Rev.D86, 010001 (2012)

[61] M. Goodman for the DUNE Collaboration, *"The Deep Underground Neutrino Experiment (DUNE)"* (July 26th 2015)

[62] K. N. Abazajiana et al., *"Light Sterile Neutrinos: A White Paper"*, arXiv:1204.5379 (2012)

[63] C. Athanassopoulos et al. (LSND Collaboration), *"Evidence for neutrino oscillations from muon decay at rest"*, Phys.Rev.C54, 2685-2708 (1996)

[64] A. A. Aguilar-Arevalo (MiniBooNE Collaboration), *"Improved Search for $\overline{\nu}_\mu \to \overline{\nu}_e$ Oscillations in the MiniBooNE Experiment"*, Phys. Rev. Lett. 110, 161801 (2013)

[65] M. Harada et al., *"Status Report for the 20th J-PARC PAC : A Search for Sterile Neutrino at J-PARC MLF (J-PARC E56, JSNS2)"*, arXiv:1507.07076 (2015)

[66] R. Allen et al. (OscSNS Collaboration), *"The OscSNS White Paper"*, arXiv:1307.7097 (2013)

[67] C. Anderson et al., *"MicroBooNE Proposal Addendum"*, (2008)

[68] K. Terao for MicroBooNE Collaboration and SBN program, *"MicroBooNE and Short Baseline Neutrino Program at Fermilab"*, NuFact 2015 conference (August 2015)

[69] F. Kaether et al., *"Reanalysis of the GALLEX solar neutrino flux and source experiments"*, Phys.Lett.B685, 4754 (2010)

[70] G. Bellini et al., *"SOX: Short distance neutrino Oscillations with BoreXino"*, JHEP 1308, 038 (2013)

[71] G. Alimonti et al. (BOREXINO Collaboration), *"Ultra-low background measurements in a large volume underground detector"*, Astropart.Phys.8, 141-157 (1998)

[72] Th. A. Mueller et al. *"Improved Predictions of Reactor Antineutrino Spectra"*, Phys.Rev.C83, 054615 (2011)

[73] K. Schreckenbach, H. R. Faust, F. von Feilitzsch, A. A. Hahn, K. Hawerkamp, and J. L. Vuilleumier, Phys. Lett. 99B, 251 (1981)

[74] F. von Feilitzsch, A. A. Hahn, and K. Schreckenbach, Phys. Lett. 118B, 162 (1982)

[75] K. Schreckenbach, G. Colvin, W. Gelletly, and F. von Feilitzsch, Phys. Lett. 160B 325 (1985)

[76] A. A. Hahn, K. Schreckenbach, W. Gelletly, F. von Feilitzsch, G. Colvin, and B. Krusche, Phys. Lett. B218, 365 (1989)

[77] P. Huber, *"On the determination of anti-neutrino spectra from nuclear reactors"*, Phys.Rev.C84, 024617 (2011)

[78] G. Mention et al., *"The Reactor Antineutrino Anomaly"*, Phys.Rev.D83, 073006 (2011)

[79] Y. Abe et al. (Double Chooz Collaboration), *"Improved measurements of the neutrino mixing angle $_{13}$ with the Double Chooz detector"*, JHEP 10 (2014) 086 [Erratum ibid. 02 (2015) 074]

[80] *"The Status of Reactor Antineutrino Flux Modelling"*, Workshop held at SUBATECH (2015)

[81] D. A. Dwyer, T. J. Langford, *"Spectral Structure of Electron Antineutrinos from Nuclear Reactors"*, Phys.Rev.Lett.114, 012502 (2015)

[82] D. Benchekroun et al. (STEREO Collaboration), *"Proposal of a search for sterile neutrinos at ILL: the STEREO experiment"*, (2012)

[83] A. Vacheret for SOLi∂ Collaboration, *"SoLid: Search for short baseline Oscillations with a novel Lithium-6 composite scintillator"*, seminar at APC (2013)

[84] P. A. R. Ade et al. (Planck Collaboration), *"Planck 2015 results. XIII. Cosmological parameters"*, Submitted to A&A (2015)

[85] M. Dittmar, M. Grazzini et al., *"Particle Physics Phenomenology 2"*, course at Institute for Theoretical Physics

[86] P. W. Higgs, *"Broken symmetries, massless particles and gauge fields"*, Phys.Lett.12, 132-133 (1964)

[87] F. Englert and R. Brout, *“Broken Symmetry and the Mass of Gauge Vector Mesons”*, Phys.Rev.Lett.13, 321 (1964)

[88] G. Aad et al. (ALTAS Collaboration), *“Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”*, Phys.Lett.B716, 1-29 (2012)

[89] S. Chatrchyan et al. (CMS Collaboration), *“Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”*, Phys.Lett.B716, 30-61 (2012)

[90] E. Majorana, L. Maiani, *“A symmetric theory of electrons and positrons”*, translated from E. Majorana, *“Teoria simmetrica dellelettrone e del positrone”*, Il Nuovo Cimento, Volume 14, Issue 4, pp 171-184 (1937)

[91] Ch. Kraus et al., *“Final Results from phase II of the Mainz Neutrino Mass Search in Tritium β Decay”*, Eur.Phys.J.C40, 447-468 (2005)

[92] A.I. Belesev et al., *“An upper limit on additional neutrino mass eigenstate in 2 to 100 eV region from ’Troitsk nu-mass’ data”*, JETPLett.97, 67-69 (2013)

[93] A. Osipowicz et al. (KATRIN Collaboration), *“KATRIN: A next generation tritium beta-decay experiment with sub-eV sensitivity for the electron neutrino mass”*, arXiv:hep-ex/0109033 (2001)

[94] K. Assamagan et al., *“Upper limit of the muon-neutrino mass and charged-pion mass from momentum analysis of a surface muon beam”* , Phys.Rev.D53, 6065 (1996)

[95] R. Barate et al. (ALEPH Collaboration), *“An upper limit on the τ neutrino mass from three- and five-prong tau decays”*, Eur.Phys.J.C2, 395-406 (1998)

[96] Y. Kolomensky, *“Neutrinoless Double-Beta Decay”*, NuFact 2015 conference (August 2015)

[97] H.V. Klapdor-Kleingrothaus et al., *“Evidence for Neutrinoless Double Beta Decay”*, Modern Physics Letters A 16 (37): 2409 (2001)

[98] Yu. A. Shitov et al. (SuperNEMO Collaboration), *“A Search for Neutrinoless Double Beta Decay: from NEMO-3 to SuperNEMO”*, arXiv:1006.4775 (2010)

[99] R. Ardito et al., *"CUORE: A Cryogenic Underground Observatory for Rare Events"*, arXiv:hep-ex/0501010 (2005)

[100] A. Gando et al. (KamLAND-Zen Collaboration), *"imit on Neutrinoless $\beta\beta$ Decay of Xe^{136} from the First Phase of KamLAND-Zen and Comparison with the Positive Claim in Ge^{76}"*, Phys.Rev.Lett.110, 062502 (2013)

[101] J.B. Albert et al. (EXO-200 Collaboration), *"Search for Majorana neutrinos with the first two years of EXO-200 data"*, Nature 510, 229-234 (12 June 2014)

[102] E. Arushanova and A. R. Back for the SNO+ Collaboration, *"Probing neutrinoless double beta decay with SNO+"*, Prospects in Neutrino Physics Conference (December 2014)

[103] G. Orebi Gann for the THEIA Collaboration, *"THEIA: Physics Potential of an Advanced Scintillation Detector"*, NuFact 2015 conference (August 2015)

[104] U. F. Katz et al. (KM3NeT Collaboration), *"The ORCA Option for KM3NeT"*, arXiv:1402.1022 (2014)

[105] M. G. Aartsen et al. (IceCube-PINGU Collaboration), *"Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU)"*, arXiv:1401.2046 (2014)

[106] ICAL Collaboration, *"Physics Potential of the ICAL detector at the India-based Neutrino Observatory (INO)"*, arXiv:1505.07380 (2015)

[107] J. Brunner for KM3NeT-ORCA Collaboration, *"KM3NeT - ORCA: Measuring neutrino oscillations and the mass hierarchy in the Mediterranean"*, ICRC 2015 conference (August 2015)

[108] F. An et al. (JUNO Collaboration), *"Neutrino Physics with JUNO"*, arXiv:1507.05613 (2015)

[109] K.K. Joo for the RENO Collaboration, *"RENO/RENO-50"*, NuFact 2015 conference (August 2015)

[110] J. Kenneth Shultis, Richard E. Faw *"Fundamentals of Nuclear Science and Engineering"*, Taylor & Francis (2002), (table 6.2), ISBN 9780824708344

[111] C. Aberle Ph.D. thesis, *"Optimization, simulation and analysis of the scintillation signals in the Double Chooz experiment"* (2011)

[112] Cristiano Galbiati, John. F. Beacom, *"Measuring the Cosmic Ray Muon-Induced Fast Neutron Spectrum by (n,p) Isotope Production Reactions in Underground Detectors"*, Phys.Rev. C72 (2005) 025807

[113] F. Ardellier et al., *"Letter of Intent for Double-CHOOZ: a Search for the Mixing Angle Theta13"*, arXiv:hep-ex/0405032 (2004)

[114] Icarus Collaboration, *"Measurement of the muon decay spectrum with the ICARUS liquid Argon TPC"*, Eur.Phys.J.C33:233-241 (2004)

[115] P C Divari et al, *"Coherent scattering of neutral-current neutrinos as a probe for supernova detection"*, J. Phys. G: Nucl. Part. Phys. 39 095204 (2012)

[116] Y. Abe et al, *"The waveform digitiser of the Double Chooz experiment: performance and quantisation effects on photomultiplier tube signals"*, JINST 8 P08015 (2013)

[117] K. Terao, *"Spallation Neutrons and Detector Response After Muon"*, Double Chooz internal talk, Doc-DB 3865-v7 (2012)

[118] Dong-Ming Mei, *"Measurement of Cosmic-Ray Muon induced processes at the KamLAND neutrino experiment"*, PhD thesis (2003)

[119] Y-F. Wang, V. Balic, G. Gratta, A. Fasso, S. Roesler, A. Ferrari, *"Predicting Neutron Production from Cosmic-ray Muons"*, Phys.Rev. D64 (2001) 013012

[120] P. Novella, *"RecoPulse: Software and algorithms for pulse reconstruction"*, Double Chooz internal note, Doc-DB 649-v2 (2009)

[121] P. Novella, *"RecoPulse configuration for data reprocessing"*, Double Chooz internal note, Doc-DB 2997 (2011)

[122] R. Gupta, H.P. Singh, C.A.L. Bailer-Jones, *"An introduction to artificial neural networks"*, Automated Data Analysis in Astronomy, Narosa Publishing House, New Delhi, India, 2001

[123] Andreas Zell et al., *"Stuttgard Neural Network Simulator, User manual"*, Version 4.2

[124] D.E. Rumelhart, J.L. McClelland, the PDP Research Group, eds, *"Parallel Distributed Processing: Explorations in the Microstructure of Cognition"*, MIT Press: Boston, p. 318 (1986)

[125] D.E. Rumelhart, G.E. Hinton, R.J. Williams, *"Learning representations by back-propagating errors"*, Nature 323, 533 - 536 (09 October 1986)

[126] Martin Riedmiller and Heinrich Braun, *"RPROP - A Fast Adaptive Learning Algorithm"*, Proc. of ISCIS VII (1992)

[127] Martin Riedmiller, *"Advanced supervised learning in multi-layer perceptrons From backpropagation to adaptive learning algorithms"*, Computer Standards & Interfaces, Volume 16, Issue 3, Pages 265-278 (July 1994)

[128] G. Pronost, *"NNIVReco off-time correction for events reconstructed outside IV"*, Double Chooz internal talk, Doc-DB 5027-v1 (2013)

[129] Anthony Onillon, *"Prédiction des taux de fission des curs de Chooz et estimation des incertitudes associées dans le cadre de lexpérience Double Chooz"*, PhD thesis (7 may 2014)

[130] M. Fallot, S. Cormon, M. Estienne, A. Algora, V.M. Bui, et al., *"New antineutrino energy spectra predictions from the summation of beta decay branches of the fission products"*, Phys. Rev. Lett., 109 : 202504 (2012)

[131] Abdoul-Aziz Zakari, *"Etude des propriétés de désintégration bta de produits de fission d'intért pour les spectres antineutrinos et la puissance résiduelle des réacteurs nucléaires"*, PhD thesis (7 may 2015)

[132] N. Haag et al., *"Experimental Determination of the Antineutrino Spectrum of the Fission Products of ^{238}U"*, Phys. Rev. Lett. 112, 122501 (2014)

[133] Th. A. Mueller et al., *"Improved Predictions of Reactor Antineutrino Spectra"*, Phys. Rev. C83, 054615, (2011)

[134] P. Huber, *"On the determination of anti-neutrino spectra from nuclear reactors"*, Phys. Rev. C84, 024617 (2011).

[135] E. Chauveau, *"ND commisionning update (1)"*, Double Chooz internal talk, Doc-DB 5806-v1 (2014)

[136] C. Aberle et al, *"Large scale Gd-beta-diketonate based organic liquid scintillator production for antineutrino detection"*, JINST 7 P06008 (2012)

[137] Ilja Bekman et al, *"Trigger Configuration Status"*, Double Chooz internal talk, Doc-DB 6060-v3 (2015)

[138] E. Chauveau, *"FD+ND commissioning status (4)"*, Double Chooz internal talk, Doc-DB 5970-v1 (2015)

[139] Private discussion with Ilja Bekman, June 12th 2015

[140] P. Vogel and J. F. Beacom, *"Angular distribution of neutron inverse beta decay,* $\overline{\nu}_e + p \rightarrow e^+ + n$*"*, Phys. Rev. D 60, 053003 (1999)

[141] D.H. Wilkinson, *"Analysis of neutron* β*-decay"*, Nucl. Phys. A377, 474 (1982)

[142] D.H. Wilkinson, *"Phase Space for Neutron* β*-Decay: An update"*, Nucl. Instrum. Methods Phys. Res. A 404, 305 (1998)

[143] A. Pichlmaier et al., *"Neutron lifetime measurement with the UCN trap-in-trap MAMBO II"*, Phys.Lett. B693, 221-226 (2010)

[144] Y. Declais et al., *"Study of reactor antineutrino interaction with proton at Bugey nuclear power plant"*, Phys.Lett. B338, 383-389 (1994)

[145] A. Cabrera et al., *"Readout Non-Linearities"*, Double Chooz internal talk, Doc-DB 3574 (2012)

[146] Brookhaven National Laboratory, National Nuclear Data Center, *http : //www.nndc.bnl.gov/*

[147] Y. Abe et al. (Double Chooz Collaboration), *"Reactor electron antineutrino disappearance in the Double Chooz experiment"*, Phys. Rev. D 86, 052008 (2012)

[148] A.C. Hayes et al., *"The Possible Origin and Implications of the Shoulder in Reactor Neutrino Spectra"*, Phys.Rev.D92, 033015 (2015) (2015)

[149] G. Pronost and R. Santorelli, *"Inner Veto Vertex Reconstruction - A Neural Network approach (Update and Summary)"*, Double Chooz internal talk, Doc-DB 4550 (2013)

[150] G. Pronost, A. Cabrera and B. Reinhold, *"Update on energy uniformity systematic uncertainty"*, Double Chooz internal talk, Doc-DB 5597 (2014)

[151] A. Cabrera, R. Carr, M. Ishitsuka and G. Pronost, *"New LNL Energy Treatment (R+S integrated)"*, Double Chooz internal talk, Doc-DB 6025 (2015)

[152] R. Carr, *"Update to H-III energy scale inputs for R+S fit"*, Double Chooz internal talk, Doc-DB 6093 (2015)

[153] C. Bauer et al., *"Qualification Tests of 474 Photomultiplier Tubes for the Inner Detector of the Double Chooz Experiment"*, JINST 6:P06008 (2011)

[154] *"Double Chooz new reconstruction methods"* meeting at APC (Paris, France), April 4th 2015

[155] F. Kaether and C. Langbrandtner, *"Transit Time and Charge Correlations of Single Photoelectron Events in R7081 PMTs"*, JINST 7 P09002 (2012)

[156] C. Bemporad, G. Gratta and P. Vogel, *"Reactor-based neutrino oscillation experiments"*, Rev.Mod.Phys.74, 297 (2002)

[157] J. M. Lopez *"Accidental background and composition"*, Double Chooz internal talk, Doc-DB 5282 (2013)

[158] R. Santorelli et al. (CIEMAT Madrid group), *"Toward a better understanding of the glowing: a possible explanation of the light emission process"*, Double Chooz internal note, Doc-DB 3234 (2011)

[159] C. Palomares, *"Study of multiplicity cut for n-Gd and n-H analyses"*, Double Chooz internal talk, Doc-DB 5112 (2013)

[160] A. Cabrera, *"FV veto update"*, Double Chooz internal talk, Doc-DB 5133 (2013)

[161] A. Stueken, *"Results of improved trigger efficiency analysis"*, Double Chooz internal talk, Doc-DB 4486-v2 (2012)

[162] R. Sharankova, *"HIII: accidental BG reduction using ANN"*, Double Chooz internal talk, Doc-DB 6037 (2015)

[163] C. Buck and B. Reinhold, *"Proton numbers in the Double Chooz Detector"*, Double Chooz internal note, Doc-DB 4211-v7 (2012)

[164] J.C. Sublet, J. Eastwood and J. Guy Morgan, *"The FISPACT-II User Manual"*, *http://www.ccfe.ac.uk/assets/Documents/easy/CCFE − R(11)11.pdf* (2014)

[165] *"RooFit Documentation"*, *http://roofit.sourceforge.net/docs/index.html*

[166] F. James and M. Winkler, *"Minuit User's guide"*, *http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf* (2004)

[167] P. Adamson et al. (MINOS Collaboration), *"Measurement of Neutrino and Antineutrino Oscillations Using Beam and Atmospheric Data in MINOS"*, Phys.Rev.Lett.110, 251801 (2013)

[168] A. Algora et al., *"Beta-decay studies using total absorption spectroscopy"*, Eur.Phys.J.A20:199-202 (2003)

[169] *"The Status of Reactor Antineutrino Flux Modelling"*, Workshop helds in SUBATECH on January 21-23, 2015

[170] Y. Abe et al. (Double Chooz collaboration), *"Ortho-positronium observation in the Double Chooz Experiment"*, JHEP 1410 (2014) 032

[171] A. Cabrera, *"DC-III IV-Tag/Veto"*, Double Chooz internal talk, Doc-DB 5038-v2 (2013)

[172] G. Pronost and A. Cabrera, *"IV Veto proposal for H-III analysis"*, Double Chooz internal talk, Doc-DB 5430-v1 (2014)

[173] G. Pronost, A. Cabrera and F. Yermia, *"IVV(H) distance cut harmonisation"*, Double Chooz internal talk, Doc-DB 5436-v2 (2014)

[174] G. Pronost, *"Inner Veto Veto for H selection - Update and tunning after ANN cut"*, Double Chooz internal talk, Doc-DB 5813-v1 (2014)

[175] A. Cabrera for the Double Chooz Collaboration, *"Double Chooz III : First results"*, Seminar in LAL (France), May 22th, 2014

[176] S.-H. Seo for the RENO Collaboration, *"New Results from RENO"*, Neutrino 2014 conference (June 2015)

[177] W. Wang for the Daya Bay Collaboration, *"The latest oscillation results from the Daya Bay reactor neutrino experiment"*, ICHEP 2014 conference (July 2014)

[178] *"Internal note of DC-III neutrino oscillation analyses using energy spectrum"*, Double Chooz internal note, Doc-DB 5162 (2013)

[179] M. Toups, *"IBD Muon Veto Study"*, Double Chooz internal note, Doc-DB 3150 (2011)

[180] A. Remoto Ph.D. thesis, *"Correlated background and impact on the measurement of* θ_{13} *with the Double Chooz detector"* (2012)

[181] T. Kawasaki and M. Kitazawa, *"Correlated BG rate estimation"*, Double Chooz internal note, Doc-DB 5876-v2 (2014)

[182] A. Hourlier, *"HIII FVV and SM contamination"*, Double Chooz internal note, Doc-DB 6039 (2015)

[183] A. Minotti, *"Stopping muons estimation using CPS on HIII"*, Double Chooz internal note, Doc-DB 6026-v3 (2015)

[184] J.B. Birks, *"Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations"*, Proc.Phys.Soc.A64: 874 (1951)

l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Guillaume PRONOST

Études pour la mesure de l'angle θ_{13} avec l'expérience Double-Chooz

Résumé

Double Chooz (DC) est une expérience s'intéressant aux oscillations des antineutrinos électroniques produis par la centrale nucléaire de Chooz (France). L'objectif de cette expérience est de mesurer l'angle de mélange θ_{13}, dont la valeur n'était pas connue jusqu'à novembre 2011. L'expérience dispose de deux détecteurs identiques: un lointain, fonctionnant depuis avril 2011, et un proche, fonctionnant depuis décembre 2014, à ~ 1 km et ~ 400 m des cœurs de réacteur, respectivement. DC mesure θ_{13} en effectuant une analyse en rate+shape sur la distorsion du spectre en énergie et sur le déficit d à la disparition des neutrinos. Le détecteur proche mesure le flux de neutrino et leur spectre en énergie avant oscillations, ce qui permet de normaliser ceux mesurés après oscillations par le détecteur lointain. Afin d'obtenir une mesure précise de θ_{13}, DC doit avoir une erreur systématique aussi basse que possible. Cette thèse présente l'expérience Double Chooz, ainsi que son analyse et ses résultats. Une attention particulière a été apporté à la réduction de certaines systématiques. La calibration en énergie du détecteur est détaillé, celle-ci permettant d'avoir une réponse du détecteur uniforme et stable. La réjection et la mesure du bruit de fond sont aussi exposées, particulièrement celles du bruit de fond corrélé. Une nouvelle méthode de réjection est présentée pour les bruits de fond corrélé (analyses Gd et H) et accidentel (analyse H). Enfin, une nouvelle méthode de calorimétrie est décrite, utilisant l'expérience acquise sur l'électronique utilisé, elle permet d'améliorer la reconstruction des photoelectrons dans le signal des photomultiplicateurs.

Mots clés
Double Chooz ; Angle de mélange θ_{13} ; Antineutrino ; Neutrinos – Interactions ; Neutron rapide ; Calorimétrie ; Oscillations ; Particules (physique nucléaire)

Abstract

Double Chooz (DC) is an experiment looking at oscillations of electron antineutrinos produced in the Chooz nuclear power plant (France). It aims to measure the θ_{13} mixing angle, which was, until November 2011, the last non-known mixing angle. The experiment consists in two identical detectors, the FAR detector, running since April 2011 and the NEAR detector, running since December 2014, respectively at ~ 1 km and ~ 400 m from the reactor cores. DC performs a rate+shape analysis to extract θ_{13} from the energy spectrum distortion and the deficit due to neutrino disappearance. The NEAR detector measures the neutrino flux and spectrum before oscillations, allowing to normalize the ones after oscillation measured with the FAR detector. In order to provide a precise measurement of θ_{13}, DC needs to control accurately the detector response and the different backgrounds, getting associated systematics as low as possible. This thesis presents the Double Chooz experiment, as well as its analysis and its results, with a specific concern about some systematics reductions. The energy calibration of the detector is detailed; this calibration allows to ensure the mandatory uniformity and stability of the detector response. The background rejection and its measurement are also detailed, with a particular focus on the correlated background. A new rejection method is presented for the correlated background (Gd and H analysis) and for the accidental background (H analysis). Finally, a new calorimetry method is described, taking advantage of the good knowledge of the DC electronics, which improves the photoelectron reconstruction from the photomultipliers signal.

Key Words
Double Chooz ; θ_{13} mixing angle ; Antineutrino ; Neutrinos – Interactions ; Fast neutron ; Calorimetry ; Oscillations ; Particles (nuclear physics)