UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS
Departamento de Física Atómica, Molecular y Nuclear

TESIS DOCTORAL

Measurement of the neutrino mixing angle θ_{13} in the Double Chooz experiment

Medida del ángulo de mezcla de neutrinos θ_{13} en el experimento Double Chooz

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

José Ignacio Crespo Anadón

Directora
Inés Gil Botella

Madrid, 2016

UNIVERSIDAD
COMPLUTENSE
MADRID

Departamento de Física Atómica, Molecular y Nuclear
Facultad de Ciencias Físicas

Measurement of the neutrino mixing angle θ_{13} in the Double Chooz experiment

Medida del ángulo de mezcla de neutrinos θ_{13}
en el experimento Double Chooz

Tesis presentada por

José Ignacio Crespo Anadón

para optar al título de Doctor

Dirigida por
Dra. Inés Gil Botella

En Madrid, octubre de 2015

A mi familia.

Agradecimientos

En primer lugar quiero agradecer al CIEMAT, y en particular a Inés Gil Botella, mi directora de tesis, el haberme dado la oportunidad de cumplir mi sueño de trabajar en Física de Partículas. A Inés quiero agradecer especialmente su constante apoyo a mi trabajo durante todos estos años, y su apuesta por mi formación enviándome a reuniones de la colaboración, conferencias, escuelas y estancias.

Quiero agradecer al "grupo de neutrinos" por haberme acogido. Gracias a Carmen Palomares por sus perspicaces observaciones sobre mis análisis y esta tesis, y a Pau Novella por todo lo que he aprendido de él, tanto directamente como a través del estudio de sus ejemplares programas. También quiero agradecer a Roberto Santorelli, que me enseñó a dar mis primeros pasos en `ROOT`, y a Antonio Verdugo y Sergio Jiménez, por su ayuda en mi trabajo de fin de máster. Muchas gracias a J. Mariano López Castaño por su ayuda con los sucesos accidentales y su explicación de los sutilísimos factores de corrección. Sin su magnífico trabajo mis estudios de la eficiencia de detección de neutrones no habrían sido posibles. Y muchas gracias por hacer la convivencia siempre fácil en todas las habitaciones que hemos compartido en reuniones y conferencias. Tengo una deuda de gratitud con Mariano y Diana Navas Nicolás por haberse encargado de los últimos turnos de guardia para que yo pudiera centrarme en escribir la tesis.

Ha sido un privilegio compartir estos años de Física con los miembros del departamento de Investigación Básica en reuniones, seminarios y cafés científicos. Me siento muy afortunado de haber sido testigo de la emoción por el lanzamiento del experimento AMS-02, el descubrimiento del bosón de Higgs... He visitado muchos lugares estos años, pero hay algo que hace que trabajar en el CIEMAT sea muy especial: la comunidad de becarios e ingenieros, personas inspiradoras; de todas ellas me llevo algo. Gracias por compartir la hora de la comida juntos y las conversaciones en la sobremesa. Y por los partidos de baloncesto, las excursiones, el cine, el Diplomacy, las cenas de Navidad y el amigo invisible... Atesoro grandes recuerdos de estos años.

Hay unas personas que deben saber que esta tesis es también suya: mi familia. Esta tesis es un paso más en mi vocación científica, que han apoyado regalándome mis primeros libros sobre el espacio y la Física de Partículas, trayéndome boletines de noticias científicas, revistas, acompañándome a conferencias... Muchas gracias por estar a mi lado y darme los medios para hacerla realidad. La tesis es la obra más extensa que he escrito hasta ahora, y en una lengua que no es mi materna. Si ha sido posible es por ti, papá. Muchas gracias por enseñarme la importancia del inglés desde muy pronto. Muchas gracias, mamá, por tu ayuda en los momentos más duros de esta tesis de la mejor forma que podías:

con comida (hecha con muchísimo amor). Estoy convencido de que esta tesis es mejor gracias a ti. Muchas gracias a mi hermano por ser un ejemplo de lucha para conseguir tus metas. Mi motivación al escribir las partes sobre Física de neutrinos fue que algún día las leerías. En el camino que ha llevado a esta tesis he aprendido muchas cosas, no sólo de Física, y todo ha sido junto a ti, Sara. Muchas gracias por querer compartir tu vida conmigo, y por cosas más mundanas como enseñarme a conducir y prácticamente regalarme un coche. Todo el tiempo ganado he podido invertirlo en mejorar mi trabajo, con una libertad de horarios total que has soportado siempre con tu alegría. Muchas gracias por esas veces que te has encargado de todo lo demás para que yo me pudiera concentrar en la tesis.

Para la realización de este trabajo he contado con una ayuda para personal investigador en formación del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas; y financiación de los programas FPA2010-15915 del Ministerio de Ciencia e Innovación y FPA2013-40521 del Ministerio de Economía y Competitividad. Asimismo, quiero agradecer las ayudas económicas de las organizaciones de la XXXIII Reunión Bienal de la Real Sociedad Española de Física en Santander (y el inmejorable alojamiento en el Palacio de la Magdalena), el XL International Meeting on Fundamental Physics en Benasque y la XXXIV Reunión Bienal de la Real Sociedad Española de Física en Valencia, que facilitaron mi asistencia a dichas conferencias.

Acknowledgements

The Double Chooz experiment is the result of a huge amount of work by many people over many years. Although it is not possible to name them all here, I have benefited so much from the good work that others did and they have my sincere gratitude.

During these years I had the pleasure to visit some of the institutions which form the Double Chooz collaboration and learn personally from the physicists in them.

I am thankful to the Double Chooz group at Nevis Laboratories for hosting me some weeks in June and July 2012. My understanding of the oscillation analysis improved enormously there. And I felt privileged to live in the quaint Nevis mansion.

I would like to thank Anatael Cabrera for inviting me to spend some weeks in November 2012 at the APC laboratory and sharing his scientific creativity and contagious enthusiasm with me. Thanks also to Kazuhiro Terao for sharing his knowledge on the neutron detection efficiency estimation for the first hydrogen-based analysis while at Paris.

I am deeply grateful to professors Masahiro Kuze and Masaki Ishitsuka, and to the Foreign Graduate Student Invitation Program of the Department of Physics of the Tokyo Institute of Technology for the opportunity to spend the month of February 2013 there. The warm hospitality of them and the students Yosuke Abe, Tomoyuki Konno, Ralitsa Sharankova made it an unforgettable experience. The seeds of the new methods to estimate the neutron detection efficiency were planted then. It was a pleasure to share this experience with Julia Haser.

I am thankful to Camillo Mariani, for hosting me again at Virginia Tech in April 2013, and to Leonidas Kalousis for creating a welcoming atmosphere. Many thanks to Rachel Carr and Keith Crum for all the discussions about the oscillation analysis and for their superb documentation on the topic.

I am especially grateful to the "neutron detection efficiency working group" (Christian Buck, Anatael Cabrera, Antoine Collin, Zelimir Djurcic, Inés, Julia Haser, Masaki Ishitsuka, Ralitsa Sharankova and Guang Yang). I had the privilege to spend many hours in meetings with them, from which I learnt a lot of physics and improved my analyses as well.

Many thanks to the people at the *château* I had the pleasure to share a dinner with during my on-site shifts. I am especially thankful to Didier Kryn and Dave Reyna for making my first shift week (with the first off-off data) so enjoyable with their delicious food and card games.

I want to acknowledge the financial support to attend the 37^{th} International Conference on High Energy Physics (ICHEP 2014) in Valencia (Spain) and the International Neutrino Summer School (INSS 2014) in St. Andrews (Scotland) by the respective organizations.

Contents

List of Figures

List of Tables

Resumen

Introducción El ángulo de mezcla de neutrinos θ_{13} es uno de los parámetros que determinan la relación entre los autoestados de sabor y los autoestados de masa de los neutrinos. El fenómeno de la oscilación de los neutrinos (el cambio periódico en la probabilidad de medir un cierto sabor del neutrino en función de la distancia recorrida) demuestra que ambos conjuntos de autoestados difieren entre sí, y que los neutrinos poseen una masa no nula. Debido a su pequeño valor, el ángulo θ_{13} fue medido por primera vez en los años 2011 y 2012 por varios experimentos, Double Chooz entre ellos. La importancia de θ_{13} radica no sólo en que es un parámetro fundamental de la naturaleza, sino que sólo en el caso en el que no sea cero es posible que exista violación de la simetría CP en la oscilación de neutrinos.

El experimento Double Chooz mide θ_{13} a partir de la desaparición de antineutrinos electrónicos ($\overline{\nu}_e$) en el flujo de antineutrinos emitido por los reactores de la central nuclear de Chooz (Francia). Los $\overline{\nu}_e$ se detectan usando la reacción de desintegración β inversa (IBD por sus siglas en inglés): $\overline{\nu}_e + p \rightarrow e^+ + n$. Para ello, Double Chooz cuenta con dos detectores idénticos: uno cerca de los reactores (a $\sim 400\,\mathrm{m}$, el Near Detector), que mide el flujo de $\overline{\nu}_e$ donde la desaparición apenas ha comenzado, y otro lejos de los reactores (a $\sim 1050\,\mathrm{m}$, el Far Detector), que lo mide donde la desaparición es casi máxima. El uso de dos detectores idénticos permite reducir las incertidumbres en el flujo de $\overline{\nu}_e$ de los reactores y en la detección de los $\overline{\nu}_e$, mejorando la precisión en la medida de θ_{13}.

La construcción de los detectores se ha realizado de forma escalonada: el Far Detector empezó a tomar datos en abril de 2011, mientras que el Near Detector inició la toma de datos a finales de 2014. Por ello, esta tesis sólo incluye los datos procesados del Far Detector, que corresponden a 489.51 días. En esta fase con un único detector, el flujo de $\overline{\nu}_e$ esperado en el detector se obtiene de una simulación Monte Carlo (MC) del experimento. Por tanto, la fidelidad de esta simulación es clave para obtener una medida exacta de θ_{13}.

Objetivos La tesis tiene como objetivo general realizar una medida de precisión del ángulo de mezcla de neutrinos θ_{13} con el experimento Double Chooz. La mejora de la precisión en θ_{13} con respecto a la primera medida de Double Chooz se organiza en torno a tres objetivos específicos:

- Mejora de la linealidad de la escala de energía del detector. Double Chooz fue el primer experimento en medir θ_{13} a partir de la información del espectro energético de los $\overline{\nu}_e$. En esta medida, la principal incertidumbre

en la escala de energía provenía de la no-linealidad a baja energía. Por ello, reduciendo esta no-linealidad se reduce la incertidumbre en θ_{13}.

- Estimación de la eficiencia de detección de neutrones en el volumen del detector con nuevos métodos más precisos. La incertidumbre en la eficiencia de detección de los neutrones producidos en la reacción $\overline{\nu}_e + p \rightarrow e^+ + n$ representa la mayor contribución a la incertidumbre sistemática relacionada con la detección de los $\overline{\nu}_e$. La estimación precisa de esta eficiencia en los datos reales y simulados, en todo el volumen del detector, es esencial para minimizar la incertidumbre en la normalización de los antineutrinos, que afecta directamente a la incertidumbre en θ_{13}.

- Ampliación y desarrollo de la herramienta estadística para el análisis de oscilaciones en la fase con dos detectores. La herramienta para la medida de θ_{13} se desarrolló para la primera fase con el detector lejano únicamente. La inclusión del detector cercano permite realizar un análisis simultáneo del flujo de $\overline{\nu}_e$ medido por los dos detectores, lo que representa la mejora crucial en la precisión de la medida de θ_{13} al producirse la cancelación de las incertidumbres correlacionadas entre detectores.

Metodología El detector de Double Chooz es un calorímetro de centellador orgánico líquido dopado con gadolinio para facilitar la detección del neutrón de la reacción $\overline{\nu}_e + p \rightarrow e^+ + n$. La interacción de un $\overline{\nu}_e$ en el detector se manifiesta por la coincidencia de dos señales, una que sigue inmediatamente a la interacción del $\overline{\nu}_e$ dada por la pérdida de energía cinética del positrón y su aniquilación con un electrón del detector (depositando una energía en el rango $1-9\,\mathrm{MeV}$), y otra que ocurre un tiempo después debido a la captura radiativa del neutrón en un núcleo. El tiempo entre señales y la energía de la segunda señal dependen del núcleo que capture el neutrón: $\sim 30\,\mu$s para el Gd, que libera rayos γ con una energía total de $\sim 8\,\mathrm{MeV}$; y $\sim 200\,\mu$s para el H, que libera un rayo γ de $2.2\,\mathrm{MeV}$. La coincidencia de las dos señales permite discriminar la señal producida por los sucesos de $\overline{\nu}_e$ del ruido de fondo debido a otros procesos.

Se ha desarrollado una simulación MC completa que predice el número de $\overline{\nu}_e$ en el detector. Esta simulación incluye la emisión de los $\overline{\nu}_e$ en las desintegraciones β^- de los productos de fisión en los reactores nucleares, que debe ser modelizada teniendo en cuenta las variaciones en la operación de los reactores y la evolución del combustible nuclear, y la posterior interacción de los $\overline{\nu}_e$ en el detector.

La colaboración Double Chooz ha desarrollado dos selecciones de $\overline{\nu}_e$ utilizando las capturas en Gd e H. La selección en Gd es la selección para la que originalmente se diseñó el detector, y presenta la mejor relación señal−ruido. Esta selección se ha optimizado respecto a aquélla usada para los primeros resultados de θ_{13} de Double Chooz, aumentando la eficiencia de detección de $\overline{\nu}_e$ y reduciendo la contaminación de fondos y las incertidumbres sistemáticas. La selección en H corresponde a un esfuerzo por obtener una medida de θ_{13} usando una muestra independiente y con mayor estadística. Debido a la baja energía liberada en la captura del neutrón, esta selección tiene una gran contaminación de coincidencias accidentales de la radiactividad ambiental. Una estrategia basada en una red neuronal ha permitido reducir ampliamente esta contaminación.

Ambas selecciones de $\overline{\nu}_e$ emplean asiduamente la energía de los sucesos como criterio de selección. Esta energía se obtiene a partir de la recolección de la luz de centelleo por parte de tubos fotomultiplicadores, que la transforman en un pulso de corriente eléctrica que se digitaliza y posteriormente se reconstruye. En dicha reconstrucción se calibra la linealidad, la uniformidad y la estabilidad de la respuesta del detector. Para ello se usan fuentes de calibración radioactivas que se introducen en el detector, la radiactividad natural y la inyección controlada de luz dentro del detector.

La estimación de la eficiencia de detección del neutrón se realiza mediante dos fuentes: ^{252}Cf y los neutrones producidos en la reacción IBD. Esta eficiencia consta de tres componentes: la fracción de capturas de neutrones en el núcleo elegido respecto al total, la eficiencia de selección de capturas de neutrones en el detector, y el cambio en el número de neutrones seleccionados por efectos de migración de los neutrones en el detector. La eficiencia de selección debe estimarse en todo el volumen del detector: en el caso del ^{252}Cf se obtiene a partir de la extrapolación de medidas tomadas en zonas concretas del mismo; mientras que los neutrones IBD, al estar distribuidos homogéneamente, proporcionan una medida directa. Los efectos de migración de neutrones se estiman comparando distintas simulaciones MC.

La medida de θ_{13} usando solamente el Far Detector se obtiene de la comparación del número de $\overline{\nu}_e$ observados con el número esperado dado por la simulación MC. Para mejorar la precisión en θ_{13}, la comparación se realiza de dos formas: analizando la modulación de la frecuencia de los sucesos de $\overline{\nu}_e$ con la potencia de los reactores (análisis RRM por sus siglas en inglés), o analizando la forma y la normalización del espectro de positrones (análisis *Rate+Shape*). En la fase con dos detectores la medida de θ_{13} se obtiene fundamentalmente de la comparación entre los datos del Near Detector con los del Far Detector. Para ello, es primordial determinar el grado de correlación entre las incertidumbres sistemáticas de ambos detectores.

Resultados La reconstrucción de la energía se ha mejorado a todos los niveles, obteniendo una respuesta más lineal, homogénea y constante. En particular, la eficiencia de reconstrucción de señales de baja energía ha aumentado un 4%. Los análisis de la eficiencia de detección de neutrones han reducido la incertidumbre relativa del 0.96% en la selección anterior en Gd a un 0.54% en la actual, y de un 1.25% en la selección anterior en H a un 0.42% en la actual. Ello, junto con las mejoras en las selecciones de $\overline{\nu}_e$ y el aumento en la cantidad de datos, ha permitido aumentar la precisión de la medida de $\sin^2(2\theta_{13})$ más de un 20% en el canal de Gd y más de un 28% en el canal de H respecto a las medidas previas de Double Chooz.

El análisis RRM ha medido $\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035}$ en el canal de Gd, y $\sin^2(2\theta_{13}) = 0.095^{+0.038}_{-0.039}$ en el canal de H. Al analizar simultáneamente ambos canales se obtiene $\sin^2(2\theta_{13}) = 0.088 \pm 0.033$.

El análisis *Rate+Shape* ha medido $\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029}$ en el canal de Gd, y $\sin^2(2\theta_{13}) = 0.124^{+0.030}_{-0.039}$ en el canal de H. Además se ha encontrado una distorsión en el espectro de $\overline{\nu}_e$ que sugiere la existencia de una nueva componente en el espectro de los $\overline{\nu}_e$ provenientes del reactor no tenida en cuenta por los modelos del flujo actuales. No obstante, esta distorsión no afecta al valor de θ_{13}, como demuestra el buen acuerdo con el resultado del análisis RRM.

Gracias a las mejoras en la incertidumbre sistemática ya obtenidas y al uso simultáneo de los detectores cercano y lejano, se prevé alcanzar una precisión del $10 - 15\%$ en $\sin^2(2\theta_{13})$ en el canal de Gd tras tres años de toma de datos con ambos detectores.

Conclusiones Los resultados presentados en esta tesis corresponden a las medidas de θ_{13} más precisas realizadas por el experimento Double Chooz hasta la fecha. Las mejoras en la escala de energía, la selección de $\overline{\nu}_e$, y la estimación de la eficiencia de detección de $\overline{\nu}_e$ descritas en esta tesis demuestran la destacable precisión alcanzable con un único detector, y sientan la base sobre la que ejecutar la medida con dos detectores. La medida con dos detectores mejorará significativamente la precisión en θ_{13} debido a la supresión de las incertidumbres sistemáticas correlacionadas entre detectores, especialmente la del flujo del reactor que predomina sobre todas las demás.

Abstract

Introduction The neutrino mixing angle θ_{13} is one of the parameters that relate the neutrino flavor eigenstates to the mass eigenstates. The phenomenon of neutrino oscillations (the periodic change in the probability of measuring a certain neutrino flavor as a function of the distance traveled by the neutrino) proves that the flavor eigenstates are not aligned with the mass eigenstates, and that neutrinos are massive particles. As a result of θ_{13} being a small angle, it was measured for the first time in 2011 and 2012 by several experiments, Double Chooz among them. The value of θ_{13} is important not only because it is a fundamental parameter of nature, but also because only if it is non-null CP violation can occur in neutrino oscillations.

The Double Chooz experiment measures θ_{13} from the disappearance of electron antineutrinos ($\overline{\nu}_e$) in the antineutrino flux emitted by the reactor cores of the Chooz nuclear power plant (France). The $\overline{\nu}_e$ are detected using the inverse β-decay reaction (IBD): $\overline{\nu}_e + p \rightarrow e^+ + n$. Two identical detectors are used: one close to the reactor cores ($\sim 400\,\text{m}$, the Near Detector) which measures the $\overline{\nu}_e$ flux where the disappearance has barely begun, and another away from the reactor cores ($\sim 1050\,\text{m}$, the Far Detector) which measures the $\overline{\nu}_e$ flux where the disappearance is almost maximal. As a consequence of the identicalness between detectors, the uncertainties on the reactor $\overline{\nu}_e$ flux and the detection of the $\overline{\nu}_e$ are reduced, increasing the precision of the measurement of θ_{13}.

The detectors were built in a phased approach: the Far Detector begun the data taking in April 2011, and the Near Detector started the data taking by the end of 2014. Hence, only the Far Detector data are covered in this thesis, totaling 489.51 run-days. In this one-detector phase, the expected $\overline{\nu}_e$ flux in the Far Detector is obtained from a Monte Carlo (MC) simulation of the experiment. Therefore, the faithfulness of the simulation is crucial in order to measure θ_{13} accurately.

Objectives The general objective of the thesis is to provide a precision measurement of the neutrino mixing angle θ_{13} with the Double Chooz experiment. The improvement on the precision on θ_{13} with respect to the first Double Chooz measurement is organized around three specific objectives:

- Improving the linearity of the energy scale of the detector. Double Chooz was the first experiment which measured θ_{13} using the energy spectrum of the $\overline{\nu}_e$. The main uncertainty on the energy scale of this measurement was due to the energy scale non-linearity at low energy. Therefore, the uncertainty on θ_{13} can be brought down if the non-linearity is reduced.

- Precise estimation of the neutron detection efficiency in the detector volume using new methods. The uncertainty on the detection of the neutrons produced in the reaction $\overline{\nu}_e + p \rightarrow e^+ + n$ is the dominant systematic uncertainty on the detection of the $\overline{\nu}_e$. The precise estimation of the neutron detection efficiency in the real and simulated data, in the full detector volume, is a fundamental step to minimize the uncertainty on the $\overline{\nu}_e$ normalization, which affects directly the uncertainty on θ_{13}.

- Expansion and development of the statistical tool for the oscillation analysis in the two-detector phase. The statistical tool for measuring θ_{13} was developed for the first phase of the experiment, in which only the Far Detector data is available. The addition of the Near Detector data allows to measure simultaneously the $\overline{\nu}_e$ flux with the two detectors. This grants the experiment the pivotal improvement on the precision of the measurement of θ_{13}, since the correlated uncertainties between detectors are canceled.

Methods The Double Chooz detector is an organic liquid scintillator calorimeter loaded with gadolinium to enhance the detection of the neutron from the reaction $\overline{\nu}_e + p \rightarrow e^+ + n$. The interaction of one $\overline{\nu}_e$ in the detector is revealed from the coincidence of two signals, a prompt one following the $\overline{\nu}_e$ interaction given by the kinetic energy loss of the positron and its annihilation with an electron from the detector (depositing an energy in the range $1 - 9\,\mathrm{MeV}$), and a delayed one given by the radiative neutron capture on a nucleus. The time interval between signals and the energy of the second signal depend on the nucleus which captures the neutron: $\sim 30\,\mu\mathrm{s}$ for Gd, which releases γ rays with a total energy of $\sim 8\,\mathrm{MeV}$; and $\sim 200\,\mu\mathrm{s}$ for H, which releases a single γ ray of $2.2\,\mathrm{MeV}$. The coincidence of these two signals allows to discriminate the $\overline{\nu}_e$ events from other background events.

A full MC simulation of the experiment predicts the number of $\overline{\nu}_e$ in the detector. This simulation includes the emission of the $\overline{\nu}_e$ in the β^- decays of the fission products in the reactor cores, which must be modeled taking into account the variations due to the reactor operation and the nuclear fuel evolution, and the $\overline{\nu}_e$ interaction within the detector.

The Double Chooz collaboration has developed two $\overline{\nu}_e$ selections using the neutron captures on Gd and on H. The Gd-based selection is the one for which the detector was originally designed, and features the best signal-to-background ratio. This selection has been optimized with respect to the one used in the first measurement of θ_{13} by Double Chooz, increasing the $\overline{\nu}_e$ detection efficiency and reducing the background contamination and the systematic uncertainties. The H-based selection is an endeavor to measure θ_{13} using an independent data sample with larger statistics. However, this selection has a sizable contamination of accidental coincidences of the natural background radiation because of the low energy released in the neutron capture. A selection strategy based on an artificial neural network has succeeded in reducing this contamination.

Both $\overline{\nu}_e$ selections use the visible energy of the events as one of the most important discriminating variables. The visible energy is obtained from the scintillation light collected by photomultiplier tubes, which transform it into an electric current pulse which is digitized and then reconstructed offline. As a part of the reconstruction, the linearity, uniformity and stability of the detector response are calibrated. In order to do so, radioactive sources deployed inside

of the detector, the natural radioactivity and controlled light injection in the detector are used as calibration sources.

The neutron detection efficiency estimation is performed using two neutron sources: ^{252}Cf and the neutrons produced in the IBD reaction. This efficiency consists of three components: the fraction of neutron captures in the chosen nucleus with respect to the total neutron captures, the selection efficiency of neutron captures within the detector, and the change in the number of selected neutrons due to neutron migration effects in the detector. The selection efficiency must be evaluated in the full detector volume: in the case of ^{252}Cf, it is inferred from the extrapolation of measurements at the source deployment positions; whereas the IBD neutrons produce a direct measurement since they are homogeneously distributed. The neutron migration effects are estimated from the comparison between MC simulations.

The measurement of θ_{13} using only the Far Detector is obtained from the comparison of the number of $\overline{\nu}_e$ observed to the one expected according to the MC simulation. In order to improve the precision on θ_{13}, the comparison is made in two different ways: in the Reactor Rate Modulation (RRM) analysis, the $\overline{\nu}_e$ interaction rate is studied as a function of the reactor power; in the Rate+Shape analysis, the shape and the normalization of the positron spectrum are studied. In the two-detector phase, the measurement of θ_{13} is practically obtained from the comparison between the Near Detector data with the Far Detector data. Therefore, it is primordial to establish the degree of correlation between the systematic uncertainties of both detectors.

Results The energy reconstruction has been improved in every aspect, achieving a response more linear, homogeneous and constant. In particular, the efficiency of the reconstruction of low energy signals has increased by 4%. The analyses of the neutron detection efficiency have reduced the relative uncertainty from 0.96% in the previous Gd-based selection to 0.54% in the current one, and from 1.25% in the previous H-based selection to 0.42% in the current one. These improvements together with the others in the $\overline{\nu}_e$ selections and the doubled statistics have made possible to increase the precision on $\sin^2(2\theta_{13})$ by more than 20% in the Gd channel and more than 28% in the H channel with respect to the previous measurements by Double Chooz.

The RRM analysis founds $\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035}$ in the Gd channel, and $\sin^2(2\theta_{13}) = 0.095^{+0.038}_{-0.039}$ in the H channel. When the two channels are analyzed simultaneously, the best fit is found at $\sin^2(2\theta_{13}) = 0.088 \pm 0.033$.

The Rate+Shape analysis founds $\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029}$ in the Gd channel, and $\sin^2(2\theta_{13}) = 0.124^{+0.030}_{-0.039}$ in the H channel. Moreover, a distortion in the $\overline{\nu}_e$ spectrum has been revealed, which suggests the existence of a new component in the spectrum of the reactor $\overline{\nu}_e$ which is not accounted for in the current reactor-flux models. Nevertheless, this distortion does not affect the θ_{13} value, as it is demonstrated by the good agreement with the result from the RRM analysis.

As a consequence of the improvements on the systematic uncertainties already obtained and the simultaneous use of the near and far detectors, a $10-15\%$ precision on $\sin^2(2\theta_{13})$ in the Gd channel is expected after three years of data taking with both detectors.

Conclusions The results presented in this thesis correspond to the most precise measurements of θ_{13} made by the Double Chooz experiment so far. The improvements on the energy scale, the $\overline{\nu}_e$ selections, and the estimation of the $\overline{\nu}_e$ detection efficiency described in this thesis demonstrate the remarkable precision which has been achieved with only one detector, and pave the way to measure θ_{13} with two detectors. The measurement of θ_{13} in the two-detector phase will improve significantly the precision on θ_{13} since the systematic uncertainties correlated between detectors are suppressed, especially the one on the reactor flux, which prevails over all the others.

Introduction

Since the beginning of philosophy, the $\alpha\rho\chi\acute{\eta}$, the ultimate substance of nature, has been a central question in our attempt to understand rationally the reality. Nowadays, the most precise answer to that question is given by the Standard Model of Elementary Particle Physics. The Standard Model has been able to explain and predict phenomena in agreement with experimental measurements to an amazing precision. However, it is an incomplete theory since it does not include gravity. Moreover, the recent astrophysical and cosmological observations require the existence of two mysterious agents known as dark matter and dark energy, for which the Standard Model offers no explanation. In addition, the Standard Model requires a relatively large number of input parameters which must be determined exclusively by experiments. This has led the Standard Model to be regarded as an effective theory of nature, with new physics awaiting to be discovered at higher energies.

So there is the incontestable success of the Standard Model over many years of experimental tests, and there is the conviction that there is physics beyond it; and in between, the neutrinos. As it will be explained in chapter 1, neutrinos were included in the Standard Model as massless particles from the available data when the theory was developed. However, the observed phenomenon of *neutrino oscillations*, in which the different types (flavors) of neutrinos transmute into each other periodically, requires them to be massive and mix among themselves; so that the original Standard Model implementation must be modified. Even though the Standard Model is able to give masses to neutrinos in the same way it gives masses to the other elementary particles, the fact that neutrino masses are much smaller than the others suggests that a different mechanism might be behind them. In any case, in order to describe the neutrino oscillations, the number of fundamental parameters of the Standard Model must be enlarged. The neutrino mixing angle θ_{13} is one of these new parameters. What makes θ_{13} special is that, unlike the other two neutrino mixing angles θ_{12} and θ_{23}, it has been measured to be relatively small. In fact, when the research that has led to this thesis started, only an upper limit on its value had been established. The measurement of θ_{13} is important not only because it is a fundamental parameter of nature, but also because only if it is non-null CP violation can take place in neutrino oscillations in vacuum, which might be related to the prevalence of matter over antimatter in our Universe, one of the observations for which the Standard Model does not provide a satisfactory answer.

The Double Chooz experiment, introduced in chapter 2, was designed to measure θ_{13} using the electron antineutrinos ($\overline{\nu}_e$) emitted from the two reactor cores of the Chooz nuclear power plant in France. For distances of the order of a kilometer, the fraction of $\overline{\nu}_e$ which oscillate into undetectable flavors is

proportional to $\sin^2(2\theta_{13})$. In order to measure the disappearance of $\overline{\nu}_e$, two detectors are used: the Near Detector at $\sim 400\,\mathrm{m}$ from the reactors, where the oscillation is almost undeveloped, and the Far Detector at $\sim 1050\,\mathrm{m}$, where the oscillation is sizable. The $\overline{\nu}_e$ are detected via the inverse-β decay reaction (IBD), $\overline{\nu}_e + p \rightarrow e^+ + n$, in a liquid scintillator calorimeter loaded with gadolinium to enhance the detection of the neutron captures. The experiment followed a staged approach: the Far Detector started the data taking first in April 2011, and the Near Detector joined the data taking in December 2014. Only the Far Detector data are covered in this thesis. In order to be able to measure θ_{13} precisely during the single-detector phase of the experiment, a detailed simulation to predict the number of $\overline{\nu}_e$ in the detector was developed.

Chapter 3 describes the event reconstruction in Double Chooz. A special emphasis is put on the energy reconstruction, since this variable plays a fundamental role in the selection of the $\overline{\nu}_e$ events and the measurement of θ_{13}. In addition, the improvement of the linearity at low energies is reviewed, since this was one of my first contributions to the Double Chooz measurement of θ_{13} published in 2012 [1].

The selection of $\overline{\nu}_e$ events is explained in detail in chapter 4. The Double Chooz collaboration has developed two selections depending on whether the IBD neutron is captured on a gadolinium or a hydrogen nucleus. In this chapter the estimation of the different backgrounds contaminating the selected samples is described too.

Chapter 5 reviews the estimation of the neutron detection efficiency in the experimental data and the Monte Carlo simulations. The neutron detection is the dominant component of the $\overline{\nu}_e$ detection systematic uncertainty for the Gd-based selection, and the next-to-leading uncertainty for the H-based selection. Therefore, a precise estimation is mandatory, since it conditionates the precision on the measurement of θ_{13}. The estimations using the IBD neutrons represent my central contributions to the Double Chooz measurements of θ_{13} in the Gd channel, published in 2014 [2], and in the H channel, which is about to be submitted for publication.

The two statistical analyses used to extract the value of θ_{13} from the data are described in chapter 6. It also includes the projected sensitivity of Double Chooz once the data from both Near and Far Detectors is analyzed, using the extension to two detectors of the statistical analysis which I contributed to develop.

In chapter 7 the conclusions of this work are presented.

Finally, some conventions used throughout this thesis:

- Natural units ($\hbar = c = 1$) are used unless stated otherwise.
- When the differentiation between actual data taken whith detector and the corresponding one obtained from the Monte Carlo simulation is important, the capitalized DATA and MC are used, respectively.
- The term "neutrino(s)" is used sometimes to denote both neutrino(s) and antineutrino(s) when the distinction is not relevant or both apply.
- Uncertainties are given at 68% C.L. unless noted otherwise.

Chapter 1

Neutrino physics

This chapter is an overview of the basics of neutrino physics in order to contextualize the measurement of the θ_{13} parameter pursued by the Double Chooz experiment. In section 1.1 neutrinos are introduced in the theoretical framework of the Standard Model. Section 1.2 discusses the phenomenon of neutrino flavor oscillations, a consequence of the neutrinos being massive particles which mix between themselves. In this section, the experimental determination of the parameters which drive the oscillations is also reviewed. Section 1.3 describes the different methods being used in order to measure the mass of the neutrinos. Finally, section 1.4 lists the most important questions in neutrino oscillations yet to be answered.

1.1 Neutrinos in the Standard Model

The Standard Model is a renormalizable relativistic quantum field theory defined by the local gauge symmetry $\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ (see [3] for an introduction). The symmetry group $\mathrm{SU(3)_C}$, where the subscript C denotes the color charge, determines the interaction between color-charged fields, known also as the strong interaction, which is described by the theory of quantum chromodynamics (QCD). The strong interaction is mediated by 8 massless bosons, the gluons, with $J^P = 1^-$, where J is the spin in natural units and P is the parity. The symmetry group $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$, where L denotes the weak isospin of the left-handed fields and Y the weak hypercharge, describes the dynamics of the electroweak interaction [4, 5, 6]. The electroweak symmetry $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ is spontaneously broken through the so-called Higgs mechanism [7, 8, 9] into $\mathrm{U(1)_{EM}}$, the symmetry group of the electromagnetic interaction, described by the theory of quantum electrodynamics (QED). The electroweak symmetry breaking results in the intermediate vector bosons ($J^P = 1^-$) of the weak interaction, W^+, W^-, Z, acquiring mass ($m_{W^\pm} = 80.385 \pm 0.015\,\mathrm{GeV}$, $m_Z = 91.1876 \pm 0.0021\,\mathrm{GeV}$ [10]); while the photon, γ, the mediator of the electromagnetic interaction, remains massless. Evidence of the breaking mechanism has been found recently in the form of the predicted Higgs scalar boson ($J^P = 0^+$), discovered with a mass of $125.7 \pm 0.4\,\mathrm{GeV}$ [10] by the ATLAS and CMS experiments at the LHC at CERN [11, 12].

The fermion fields ($J = 1/2$) are not predicted by the mathematical structure

Q	Generation 1st	2nd	3rd	
$+\frac{2}{3}$	u ($2.3^{+0.7}_{-0.5}$ MeV)	c (1275 ± 25 MeV)	t ($173.21 \pm 0.51 \pm 0.71$ GeV)	Quarks
$-\frac{1}{3}$	d ($4.8^{+0.5}_{-0.3}$ MeV)	s (95 ± 5 MeV)	b (4.18 ± 0.03 GeV)	
0	ν_e (< 2 eV)	ν_μ (< 0.19 MeV)	ν_τ (< 18.2 MeV)	Leptons
-1	e (0.511 MeV)	μ (105.658 MeV)	τ (1776.82 ± 0.16 MeV)	

Table 1.1: Elementary fermion particles of the Standard Model, indicating their electric charge Q (in units of the proton charge), their mass (below the particle symbol and within parentheses, in natural units, from [10]) and the generation which belong to. The antiparticles have the same mass as their particle counterparts but the opposite charge, and are denoted with a bar over the symbol.

of the Standard Model, and must be discovered experimentally. In table 1.1, the elementary fermions are introduced together with their mass and charge. Depending on whether they experience the strong interaction or not, fermions are grouped into quarks and leptons, respectively. Each fermion is characterized by a quantum number called flavor; in addition, all quarks are assigned a baryon number $+1/3$ and all leptons are assigned a lepton number $+1$ (the antiparticles have the opposite numbers). In the Standard Model, the total baryon and lepton numbers are conserved in the interactions. They are referred to as accidental symmetries, since they are not imposed on the model. It has been observed that the fermions can be arranged in generations or families, consisting of two quarks with charges $+2/3$ and $-1/3$, and two leptons, one being neutral and the other negatively charged. Three generations have been found, each one being a copy of the preceding one but with heavier elements. If neutrinos did not mix, a lepton family number for each generation could be introduced, which would be conserved too. However, neutrino oscillations (described in the next section) prove this is not the case.

For every fermion field ψ, the left $\psi_{\mathrm{L}} = P_{\mathrm{L}}\psi$ and right $\psi_{\mathrm{R}} = P_{\mathrm{R}}\psi$ chirality projections (where $P_{\mathrm{L}} \equiv (1-\gamma^5)/2$ and $P_{\mathrm{R}} \equiv (1+\gamma^5)/2$ are the projectors), are assigned to irreducible representations of the symmetry group $\mathrm{SU(3)_C \times SU(2)_L \times U(1)_Y}$ in table 1.2. Notice the absence of the right-handed neutrinos. The reason is that when the Standard Model was formulated, the lightness of the neutrinos and the fact that only experimental evidence of negative helicity neutrinos (or positive helicity antineutrinos) had been found [13], led to consider them as massless; and hence described by pure chiral fields (left-handed for the neutrinos, right-handed for the antineutrinos). However, the observed phenomenon of neutrino oscillation requires them to be massive.

In the Standard Model, after the electroweak spontaneous symmetry breaking has occurred, the mass of the charged fermions is generated through their Yukawa interaction with the Higgs field, which is written in the Standard Model

Generation			Representation		
1^{st}	2^{nd}	3^{rd}	$\text{SU}(3)_\text{C}$	$\text{SU}(2)_\text{L}$	Y
$\begin{pmatrix} u_i \\ d_i \end{pmatrix}_\text{L}$	$\begin{pmatrix} c_i \\ s_i \end{pmatrix}_\text{L}$	$\begin{pmatrix} t_i \\ b_i \end{pmatrix}_\text{L}$	3	2	$\frac{1}{3}$
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_\text{L}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_\text{L}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_\text{L}$	1	2	-1
$u_{i,\text{R}}$	$c_{i,\text{R}}$	$t_{i,\text{R}}$	3	1	$\frac{4}{3}$
$d_{i,\text{R}}$	$s_{i,\text{R}}$	$b_{i,\text{R}}$	3	1	$-\frac{2}{3}$
e_R	μ_R	τ_R	1	1	-2

Table 1.2: Left and right chirality projections of the elementary fermion fields classified according to how they transform under the symmetry group $\text{SU}(3)_\text{C} \times \text{SU}(2)_\text{L} \times \text{U}(1)_\text{Y}$. The subscript i in the quark fields denotes the color charge which has three possible values, usually named *red*, *green* and *blue*. The weak hypercharge is defined as $Y = 2(Q - T_3)$, where T_3 is the third component of the weak isospin.

lagrangian (density) as (for simplicity, only one fermion is considered)

$$\mathcal{L}_\text{mass}^\text{Dirac} = -\overline{\psi_\text{R}} M_\text{D}^\psi \psi_\text{L} + \text{H.c.}, \tag{1.1}$$

where the mass of the ψ fermion is given by $M_\text{D}^\psi = \frac{1}{\sqrt{2}} y^\psi v$, with y^ψ the Yukawa coupling of the ψ fermion to the Higgs field, and v the vacuum expectation value of the Higgs field (246 GeV), respectively (H.c. stands for Hermitian conjugate). The mass of the neutrinos can also be generated in this way, known as Dirac mass, provided that the set of right-handed neutrino fields missing in table 1.2 is added, extending minimally the original Standard Model to allow for massive neutrinos. According to the classification of table 1.2, the right-handed neutrino fields would have the following representation $(1, 1, 0)$, which make them singlets under all the Standard Model interactions, so they are known as *sterile*. However, this mechanism raises the question about why the neutrino masses are so small compared to the other fermions [1]. In section 1.3.2, another mass generation mechanism is introduced which relies on the existence of Majorana neutrino fields.

Neutrinos, as neutral leptons, can only undergo weak interactions in the Standard Model. There are two types of weak interactions, the charged current or the neutral current. The charged currents are mediated by the $W^\pm$ bosons, and are written

$$\mathcal{L}_\text{CC} = -\frac{g}{\sqrt{2}} \overline{\nu_{l\text{L}}} \gamma^\mu l_\text{L} W_\mu + \text{H.c.}, \tag{1.2}$$

where g is the weak interaction coupling constant, $\overline{\nu_{l\text{L}}} = \nu_{l\text{L}}^\dagger \gamma^0$, $\nu_{l\text{L}}$ and l_L are the left-chirality projections of the neutrino and charged lepton fields, respectively; γ^μ is the μ-th gamma matrix and W_μ is the W boson gauge field. Notice that

[1] It is legitimate to argue that putting the neutrinos aside, the Higgs-induced masses of the charged fermions require Yukawa couplings which differ in 6 orders of magnitude. However, the neutrino masses require at least another 6 orders of magnitude more.

the flavor of the neutrino ν_l, with $l = e, \mu, \tau$, is assigned according to the charged lepton l which accompanies it in the charged current interaction.

Neutral currents are mediated by the Z boson, and are written

$$\mathcal{L}_{\mathrm{NC}} = -\frac{g}{2\cos\theta_W}\overline{\nu_{l\mathrm{L}}}\gamma^\mu\nu_{l\mathrm{L}}Z_\mu, \tag{1.3}$$

where θ_W is the weak mixing angle parameter, and Z_μ is the Z boson gauge field.

The existence of the (electron) neutrino was first postulated by Pauli in 1930 [14]. According to Pauli, in order to explain the integer spin of nuclei ^{14}N and ^{6}Li, and at the same time the apparent non-conservation of energy in the β decay; a light neutral fermion with spin 1/2 existed within the nucleus, which was emitted along the electron in the β decay. Pauli's proposal turned out to be two different particles, the neutron, discovered by Chadwick in 1932 [15]; and the neutrino, "little neutral one" in Italian, named by Fermi, who formulated the first theory of β decay [16], which included this new particle. The discovery of the neutrino (to be precise, the electron antineutrino) was made by Reines and Cowan in 1956, when they detected the electron antineutrinos emitted by the Savannah River nuclear reactor [17]. The existence of a second neutrino flavor different from the electron one, the muon neutrino, was hypothesized to explain the muon decay modes [18, 19]. It was found in 1962 by Lederman, Schwartz, Steinberger *et al.* in an experiment at Brookhaven National Laboratory using a neutrino beam produced by the decay in flight of pions [20]. The tau neutrino was not directly observed until the year 2000 by the DONUT experiment at Fermi National Accelerator Laboratory [21]. However, given the recurrent neutrino-charged lepton symmetry, its existence could be anticipated since the discovery of the tau lepton in 1975 by Perl *et al.* at the SPEAR e^+e^- collider at SLAC [22]. Moreover, the three-flavor neutrino paradigm was firmly established since the analysis of the decay width of the Z boson by the ALEPH, DELPHI, L3 and OPAL detectors at LEP at CERN, which determined from the invisible partial width that the number of neutrinos [2] is 2.9840 ± 0.0082 [23].

1.2 Neutrino oscillations

In the three-flavor framework of massive neutrinos, the relationship between the neutrino flavor eigenstates that participate in the weak interaction, $|\nu_l\rangle$, with $l = e, \mu, \tau$, and the mass eigenstates, $|\nu_j\rangle$, which describe the propagation in vacuum can be written

$$|\nu_l\rangle = \sum_{j=1}^{3} U_{lj}^* \, |\nu_j\rangle \,, \tag{1.4}$$

where the sum in j is over the three mass eigenstates with mass m_j. U is a 3×3 unitary matrix which will be non-diagonal if neutrinos mix, playing an equivalent role to the CKM (Cabibbo-Kobayashi-Maskawa) matrix in the quark sector [24, 25]. U is known as the Pontecorvo-Maki-Nakagawa-Sakata matrix, or PMNS for short [26, 27, 28].

[2] Specifically, the number of light neutrinos ($m_\nu < m_Z/2$) which experience weak interactions, also known as *active* neutrinos.

Among the 9 parameters that define a 3×3 unitary matrix like U, 5 (3) phases can be absorbed in the redefinition of the neutrino fields if they are Dirac (Majorana), so only 4 (6) remain as independent. The flavor and mass eigenstates can be regarded as two bases of a vector space of dimension 3, so U would be the matrix to change of basis. Therefore, U can be parametrized in terms of three Euler angles, θ_{12}, θ_{23} and θ_{13}, known as the neutrino mixing angles. In the case of Dirac neutrinos, the remaining parameter is a CP-violating phase, δ, and the most general parametrization of U is:

$$\begin{aligned} U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}\mathrm{e}^{-\mathrm{i}\delta} \\ 0 & 1 & 0 \\ -s_{13}\mathrm{e}^{\mathrm{i}\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-\mathrm{i}\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{1.5}$$

where $c_{ij} \equiv \cos(\theta_{ij})$, $s_{ij} \equiv \sin(\theta_{ij})$, with $\theta_{ij} \in [0, \pi/2]$ and $\delta \in (-\pi, \pi]$. In the case of Majorana neutrinos, two additional CP-violating phases, α_1 and α_2 are needed; so U must be multiplied by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\alpha_1} & 0 \\ 0 & 0 & \mathrm{e}^{\mathrm{i}\alpha_2} \end{pmatrix} \tag{1.6}$$

However, these two phases do not affect the neutrino oscillations as we will see.

Let $|\nu_l(t)\rangle$ be a neutrino after a time t, which originally was created with flavor l, $|\nu_l(t=0)\rangle = |\nu_l\rangle$. According to equation 1.4, it can be expressed as

$$|\nu_l(t)\rangle = \sum_{j=1}^{3} U_{lj}^{*} \, |\nu_j(t)\rangle \,, \tag{1.7}$$

where $|\nu_j(t)\rangle$ is the time-evolution of the mass eigenstate $|\nu_j\rangle$. Following [29], we describe the mass eigenstate as a plane wave [3]; its time-evolution is given by the Scrödinger equation:

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t} |\nu_j(t)\rangle = E_j \, |\nu_j\rangle \,, \tag{1.8}$$

where $E_j = \sqrt{|\vec{p_j}|^2 + m_j^2}$. Therefore, equation 1.7 is written as

$$|\nu_l(t)\rangle = \sum_{j=1}^{3} U_{lj}^{*} \mathrm{e}^{-\mathrm{i}E_j t} \, |\nu_j\rangle \,. \tag{1.9}$$

If the mass eigenstate $|\nu_j\rangle$ is expressed as a sum of the flavor eigenstates by inverting equation 1.4,

$$|\nu_l(t)\rangle = \sum_{l'=e,\mu,\tau} \left(\sum_{j=1}^{3} U_{lj}^{*} \mathrm{e}^{-\mathrm{i}E_j t} U_{l'j} \right) |\nu_{l'}\rangle \,, \tag{1.10}$$

[3] More sophisticated descriptions based on wave packets (e.g. [30]) or quantum field theory (e.g. [31]) arrive at the same neutrino oscillation probability, but the plane wave is instructive because of its simplicity.

which shows that after a time t, $|\nu_l(t)\rangle$ is found in a superposition of flavor eigenstates, with coefficients given by the factor within parentheses. The probability to detect a particular flavor l', $P_{ll'}$, is given by

$$P_{ll'}(t) = \left| \langle \nu_{l'} | \, \nu_l(t) \rangle \right|^2 = \sum_{j,k=1}^{3} U_{lj}^* U_{l'j} U_{lk} U_{l'k}^* \mathrm{e}^{-\mathrm{i}(E_j - E_k)t}. \tag{1.11}$$

Since all experiments made so far use ultrarelativistic neutrinos, the energy of the mass eigenstate $|\nu_j\rangle$ can be approximated by

$$E_j \simeq E + \frac{m_j^2}{2E}, \tag{1.12}$$

with $E \simeq |\vec{p_j}|$ the neutrino energy neglecting the mass contribution. Then,

$$E_j - E_k \simeq \frac{\Delta m_{jk}^2}{2E}, \tag{1.13}$$

where

$$\Delta m_{jk}^2 \equiv m_j^2 - m_k^2. \tag{1.14}$$

Moreover, in the experiments the production time of the neutrino is not known (at least with enough precision), but the distance L between the source and the detector can be measured precisely. For ultrarelativistic neutrinos, $t \simeq L$; so the probability of equation 1.11 becomes

$$P_{ll'}(L, E) = \sum_{j,k=1}^{3} U_{lj}^* U_{l'j} U_{lk} U_{l'k}^* \exp\left(-\mathrm{i} \frac{\Delta m_{jk}^2 L}{2E} \right). \tag{1.15}$$

Because $\mathrm{e}^{-\mathrm{i}x} = \cos x - \mathrm{i} \sin x$, the probability of detecting a neutrino of flavor l' oscillates as a function of the distance traveled L or the neutrino energy E, hence the name *neutrino oscillations*. For a fixed L/E, the oscillation frequency is determined by Δm_{jk}^2. Therefore, the observation of neutrino oscillation implies that $\Delta m_{jk}^2 \neq 0$, that is, neutrinos are massive and non-degenerated ($m_j \neq m_k$). The amplitude of the oscillations depends on the magnitude of the U matrix elements; however, it can be shown that in the product $U_{lj}^* U_{l'j} U_{lk} U_{l'k}^*$, the Majorana phases of equation 1.6 cancel out, so the oscillation amplitude is only a function of the three mixing angles, θ_{12}, θ_{23}, θ_{13}, and the CP-violating phase δ.

The neutrino oscillation experiments are designed so that L and E can be known (within uncertainties); so the unknown physical parameters (mixing angles, squared-mass differences, CP-violating phase) can be inferred from the measurement of the probability of detecting a flavor l' in a neutrino flux created with a flavor l. If $l = l'$, the probability is called *survival probability*; if $l \neq l'$, it is called *transition probability*.

So far, only neutrinos have been considered. The derivation for antineutrinos follows along the sames lines of the neutrino case, except that the coefficients of the antineutrino mass eigenstates in an expression analogous to equation 1.4 are the complex conjugate of the neutrino ones. Consequently, it is useful to

rewrite the oscillation probability in equation 1.15 as

$$\begin{aligned} P_{ll'}(L,E) = \delta_{ll'} - 4\sum_{j>k} \mathrm{Re}\left(U^*_{lj}U_{l'j}U_{lk}U^*_{l'k}\right)\sin^2\left(\frac{\Delta m^2_{jk}L}{4E}\right) \\ \pm 2\sum_{j>k} \mathrm{Im}\left(U^*_{lj}U_{l'j}U_{lk}U^*_{l'k}\right)\sin\left(\frac{\Delta m^2_{jk}L}{2E}\right), \end{aligned} \tag{1.16}$$

where the sign of the imaginary part is + for neutrinos and − for antineutrinos. Hence, the CP violation will be determined by magnitude of the imaginary part. Furthermore, since CPT symmetry requires the probabilities $P(\nu_l \to \nu_l)$ and $P(\overline{\nu}_l \to \overline{\nu}_l)$ to be equal, CP violation effects will appear only in the transition probabilities.

Under certain circumstances, such as when there are two squared-mass differences, Δm^2_{jk} and Δm^2_{ik} with $i \neq j$, which differ greatly in magnitude, $\Delta m^2_{jk} >> \Delta m^2_{ik}$, so that is possible to find a L/E range in which one oscillation has developed ($\Delta m^2_{jk} \sim 1$) while the other not, the three-neutrino mixing can be approximated by a two-flavor mixing. A two-flavor approximation can also be used if one of the mixing angles is much smaller than the others; or if the detector is only sensitive to one flavor l, so the other two flavors can be combined in an effective indistinguishable flavor x (e.g. if the detector only "sees" ν_e, the effective flavor $\nu_x = \epsilon\,\nu_\mu + \sqrt{1-\epsilon^2}\,\nu_\tau$ can be defined, with $\epsilon \in [0,1]$). In this case, the relationship between flavor eigenstates and mass eigenstates is reduced to

$$\begin{pmatrix} |\nu_l\rangle \\ |\nu_x\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \end{pmatrix}, \tag{1.17}$$

where only one mixing angle θ exists. Defining the unique squared-mass difference

$$\Delta m^2 \equiv m^2_2 - m^2_1, \tag{1.18}$$

where m_1 is taken as the lightest mass, so $\Delta m^2 > 0$; the transition probability reads

$$P^{2\nu}_{lx\,(l\neq x)}(L,E) = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right), \tag{1.19}$$

and the survival probability

$$P^{2\nu}_{ll}(L,E) = 1 - P_{lx\,(l\neq x)}(L,E). \tag{1.20}$$

The oscillation probabilities just derived apply to the case in which the neutrinos propagate in vacuum. If the neutrinos traverse a medium dense enough so that matter effects become important, the propagation is affected by the coherent forward elastic scattering with the particles composing medium [32]. For electron neutrinos and antineutrinos, the charged current weak interaction with a homogeneous gas of unpolarized electrons ($\overset{(-)}{\nu}_e\, e^- \xrightarrow{\text{CC}} \overset{(-)}{\nu}_e\, e^-$) induces a potential

$$V_{\text{CC}} = \pm\sqrt{2}G_{\text{F}}N_e, \tag{1.21}$$

where the sign is + for ν_e (− for $\overline{\nu}_e$), $G_{\text{F}} = \sqrt{2}\frac{g^2}{8m^2_W}$ is the Fermi coupling constant, and N_e is the electron density in the medium. In addition, for the

three flavors, the neutral current weak interaction with the f fermions (electrons, protons and neutrons) ($\overset{(-)}{\nu}_l f \xrightarrow{\text{NC}} \overset{(-)}{\nu}_l f$) of a neutral medium induces the potential

$$V_{\text{NC}} = \mp -\frac{1}{2}\sqrt{2}G_{\text{F}}N_n, \tag{1.22}$$

where the sign is $-$ for neutrinos ($+$ for antineutrinos), and only depends on N_n, the neutron density in the medium, since the electron and proton contributions cancel each other out. As a result of these potentials, the mass eigenstates in matter differ from those in the vacuum, and thereby the mixing angles do too.

Working in the two-flavor approximation, one flavor being e and another μ or τ, the mixing angle in vacuum, θ, is replaced by the mixing angle in matter θ_m, which is related to the former by

$$\tan 2\theta_m = \frac{\tan 2\theta}{1 \mp \frac{2\sqrt{2}EG_{\text{F}}N_e}{\Delta m^2 \cos 2\theta}}, \tag{1.23}$$

and their squared-mass difference becomes

$$\Delta m^2_m = \sqrt{(\Delta m^2 \cos 2\theta \mp 2\sqrt{2}EG_{\text{F}}N_e)^2 + (\Delta m^2 \sin 2\theta)^2}, \tag{1.24}$$

where the sign is $-$ for neutrinos ($+$ for antineutrinos). No dependence on the neutral current potential appears since it affects both flavors equally. An important consequence of the effect of matter on the neutrino oscillation is that there is a resonance on equation 1.23 if the electron density is

$$N_e^{\text{res}} = \pm\frac{\Delta m^2 \cos 2\theta}{2\sqrt{2}EG_{\text{F}}}, \tag{1.25}$$

where the sign is $+$ for neutrinos ($-$ for antineutrinos). This leads to a maximal mixing angle in matter ($\theta_m = \pi/4$) regardless of the value of the mixing angle in vacuum, a phenomenon known as the MSW effect, after Mikheev, Smirnov and Wolfenstein [32, 33, 34].

To conclude this section, returning to the three-neutrino framework, three squared-mass differences exist: Δm^2_{21}, Δm^2_{32}, Δm^2_{31}. However, only two are independent: $\Delta m^2_{31} = \Delta m^2_{32} + \Delta m^2_{21}$. As it will be explained in the next section, experiments using the matter effects in the Sun show $\Delta m^2_{21} > 0$. Consequently, there are two possibilities left depicted in figure 1.1: m_1 is the lightest neutrino mass ($\Delta m^2_{32} > 0$), known as *normal mass hierarchy*; or m_3 is the lightest mass ($\Delta m^2_{32} < 0$), known as the *inverted mass hierarchy*.

The next sections discuss the measurements of the three mixing angles and the two squared-mass differences.

1.2.1 Measurement of θ_{12} and Δm^2_{21}

The first experiment that detected (unawarely) the consequences of neutrino oscillation was Bahcall and Davis' experiment [36, 37], proposed in 1964. It was located in the Homestake mine at Lead (South Dakota, USA), and aimed to detect the neutrinos produced within the Sun in the thermonuclear fusion of protons into helium nuclei:

$$4p + 2e^- \rightarrow {}^4\text{He} + 2\nu_e \tag{1.26}$$

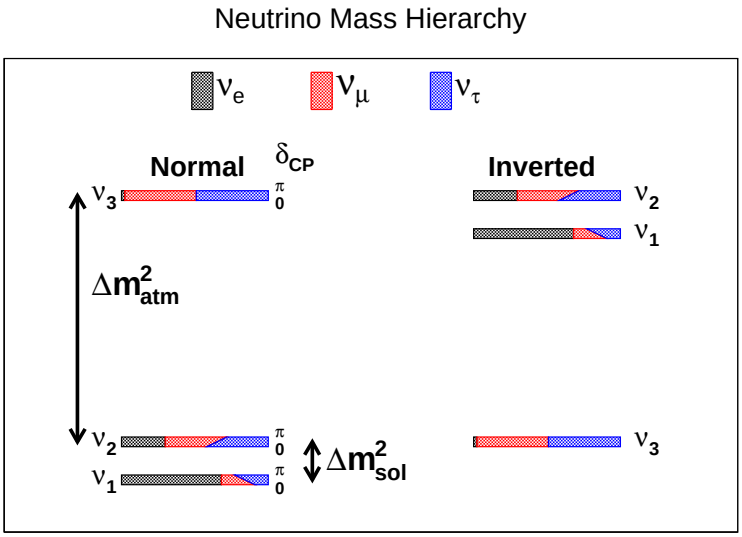


Figure 1.1: The two possible neutrino mass hierarchies. For each mass eigenstate, the flavor composition is shown using $\sin^2(\theta_{12}) = 0.30$, $\sin^2(\theta_{23}) = 0.45$, $\sin^2(\theta_{13}) = 0.02$, and as a function of the CP-violating phase δ_{CP}. The naming convention $\Delta m^2_{21} \equiv \Delta m^2_{\rm sol}$, $\Delta m^2_{32} \equiv \Delta m^2_{\rm atm}$ is used because of the typical neutrino origin used to measure them: solar and atmospheric. From [35].

which has a Q-value of 26.731 MeV (see figure 1.2 for the two different reaction routes, and figure 1.3 for the predicted spectrum of the ν_e). The neutrinos were detected using the reaction:

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow e^- + {}^{37}\mathrm{Ar} \tag{1.27}$$

which has a neutrino energy threshold of $E_{\nu,\rm th} = 0.814\,\mathrm{MeV}$. The detector was a steel tank filled with C_2Cl_4. The resulting ^{37}Ar is a radioisotope which decays back to ^{37}Cl via electron capture (half-life of 35.011 days [38]), so in order to measure the rate of ν_e interactions, the resulting ^{37}Ar was extracted through chemical means every two months approximately, and left to decay in miniature proportional counters.

Since the beginning, the first data pointed to a deficit of ν_e with respect to the Standard Solar Model (SSM) prediction, which was supported by the helioseismological data (i.e. the measurements of the propagation of acoustic waves through the Sun). This discrepancy was called the *solar neutrino problem*. As the years passed by, both the experiment and the theoretical prediction were improved, but the deficit remained. After 24 years of measurements (1970 − 1994), it was found

$$\frac{R_{\rm Cl}}{R_{\rm SSM}} = 0.301 \pm 0.027 \tag{1.28}$$

where $R_{\rm Cl}$ is the observed rate [41] and $R_{\rm SSM}$ is the predicted rate according to

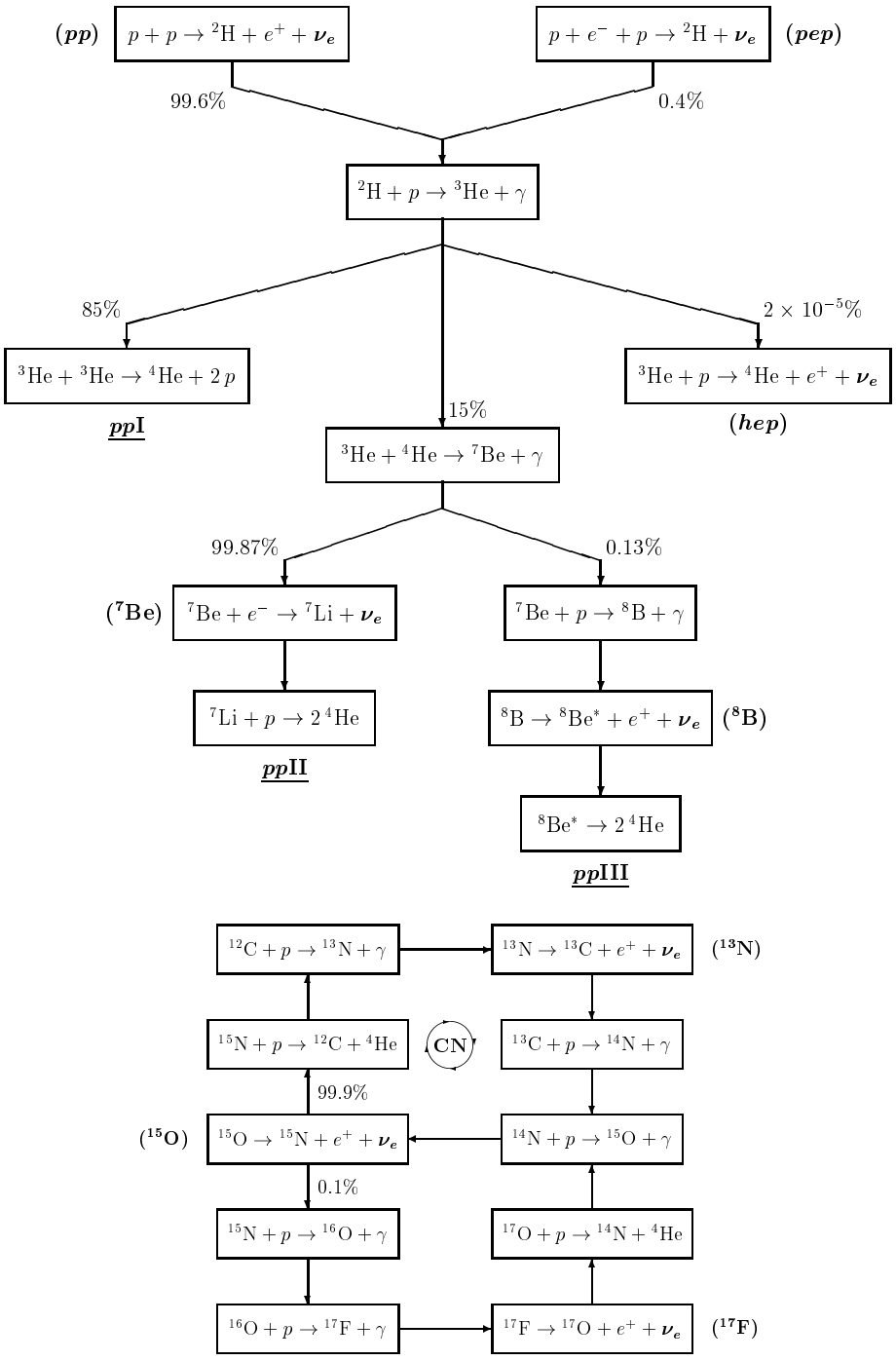


Figure 1.2: The stellar thermonuclear fusion reactions that convert p into ^{4}He: *pp* chain (top, producing 98.4% of the solar energy) and CNO cycle (bottom, producing 1.6% of the solar energy). The produced neutrinos are indicated in boldface, with their traditional name in parentheses. From [39].

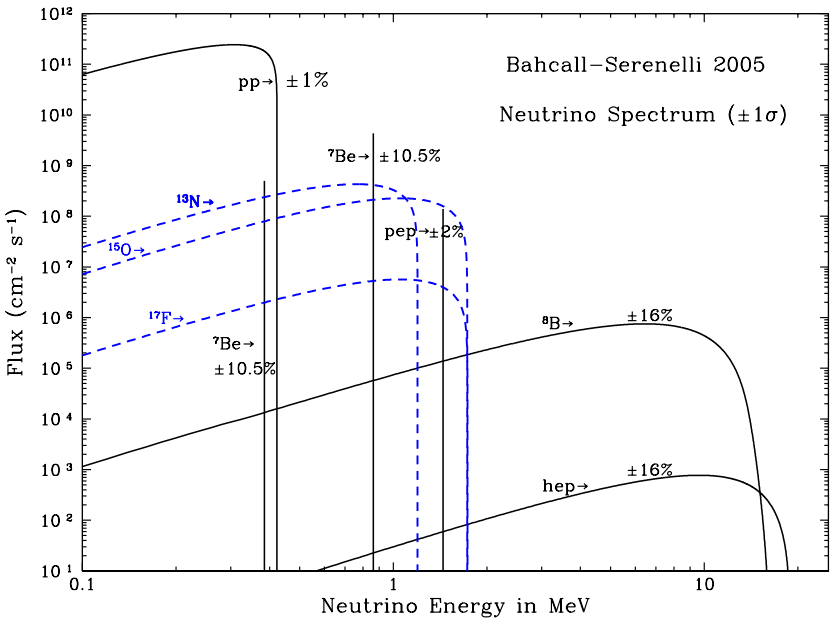


Figure 1.3: Predicted energy spectra of the neutrino fluxes emitted by the Sun. The fluxes are named after the reactions that produce them (see figure 1.2). Units are $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{MeV}^{-1}$ for continuous and $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for line spectra. Percentages show the theoretical uncertainty. From [40].

reference [42] [4].

In order to investigate further the solar neutrino flux, two experiments were proposed: GALLEX (later GNO), located in the Gran Sasso laboratory (Italy), and SAGE, located in the Baksan Neutrino Observatory (Russia). Both used ^{71}Ga as target, which undergoes the following reaction:

$$\nu_e + {}^{71}\mathrm{Ga} \rightarrow e^- + {}^{71}\mathrm{Ge} \tag{1.29}$$

The advantage of this reaction is its lower energy threshold, $E_{\nu,\mathrm{th}} = 0.233\,\mathrm{MeV}$, so the experiments are sensitive to the lower energy part of the spectrum (cf. figure 1.3). ^{71}Ge is also radioactive, and decays back to ^{71}Ga via electron capture (half-life of 11.43 days [38]); so the determination of the rate of ν_e interactions was done similar to the Homestake experiment: ^{71}Ge was chemically extracted and allowed to decay within miniature proportional counters. Both experiments measured a deficit [43, 44] with respect to the prediction,

$$\frac{R_{\mathrm{GALLEX/GNO}}}{R_{\mathrm{SSM}}} = 0.529 \pm 0.042 \tag{1.30}$$

$$\frac{R_{\mathrm{SAGE}}}{R_{\mathrm{SSM}}} = 0.540 \pm 0.040 \tag{1.31}$$

in good agreement with each other, but different from the chlorine experiment (equation 1.28).

The previous experiments rely on radiochemical methods to detect the accumulation of neutrino interactions. However, an alternative technique is that of the water Cherenkov detectors, in which the neutrinos are detected in real time through the Cherenkov radiation emitted by the electron upon elastic scattering (ES):

$$\nu_l + e^- \rightarrow \nu_l + e^- \quad \text{(ES)}. \tag{1.32}$$

This process is sensitive to the three neutrino flavors via neutral-current interaction, but the cross-section for ν_e is ~ 6 times larger because of the additional contribution of the charged current. The electron is scattered in the forward direction along the line defined by the Sun and the detector, allowing to use the directionality to discriminate signal events from the isotropic background.

The first experiment that detected solar neutrinos using this technique was the Kamiokande experiment, located in the Kamioka mine (Japan). Originally it was built to search for the nucleon decay, with 3000 ton of water in a steel tank observed by about 1000 large photomultiplier tubes (50 cm diameter). After an upgrade, it began to detect solar neutrinos with an energy cut of $E_\nu > 9.0\,\mathrm{MeV}$ to reject background events, which was reduced in subsequent phases of the experiment to 7.2 MeV, and finally to 6.7 MeV. The ratio of the observed flux [45] to the predicted in case of no oscillation was found to be

$$\frac{\Phi_{\mathrm{Kamiokande}}}{\Phi_{\mathrm{SSM}}} = 0.484 \pm 0.066. \tag{1.33}$$

Kamiokande was succeeded by Super-Kamiokande (SK), an enlarged version of the former which used 50 kton of water. During its first phase, it took data

[4] Several predictions are available which differ slightly depending on the solar input data used. Since the purpose of the observed to predicted ratios shown here is to illustrate the deficit of ν_e rather than a systematic study, reference [42] (known as BP04) is taken as the prediction, in order to quote the ratios with respect to it from reference [29].

with an initial detector threshold set at $E_\nu > 6.2\,\mathrm{MeV}$, which was lowered to 4.7 MeV later. The results of this phase [46] yielded a ratio to the prediction of

$$\frac{\Phi_{\mathrm{SK-I}}}{\Phi_{\mathrm{SSM}}} = 0.406 \pm 0.014, \tag{1.34}$$

which confirmed the deficit of ν_e.

The solar neutrino problem was solved by the SNO experiment. It was located in a mine near Sudbury (Ontario, Canada), and as Kamiokande and Super-Kamiokande, it was a water Cherenkov detector. What made SNO special is that it used heavy water (D_2O) instead of regular water (H_2O). Therefore, in addition of the elastic scattering (see equation 1.32), SNO could detect electron neutrinos using the charged-current reaction

$$\nu_e + \mathrm{D} \rightarrow e^- + p + p \quad (\mathrm{CC}), \tag{1.35}$$

which has a neutrino energy threshold of 1.442 MeV; although an analysis cut on the electron kinetic energy 5 MeV was used to reject the background, so the effective neutrino thresholds were $E_\nu > 5.7\,\mathrm{MeV}$ for the ES channel and 6.9 MeV for the CC channel. Moreover, SNO could detect neutrino interactions through the neutral-current reaction

$$\nu_l + \mathrm{D} \rightarrow \nu_l + p + n \quad (\mathrm{NC}), \tag{1.36}$$

which has a neutrino energy threshold of 2.224 MeV. This reaction is of utmost importance because it is equally sensitive to the three flavors, so it allows to measure the solar neutrino flux regardless of whether neutrinos have oscillated into other flavors or not.

In the first phase of the experiment the neutron was detected upon radioactive capture on deuterium, $n + \mathrm{D} \rightarrow \mathrm{T} + \gamma$, which produces a tritium nucleus and a single 6.25 MeV γ-ray which is detected through the Cherenkov radiation from secondary Compton electrons or e^+e^- pairs. In order to enhance the neutron detection, salt (NaCl) was added to the heavy water in the second phase of the experiment, so the neutrons could be captured on ^{35}Cl: $n + {}^{35}\mathrm{Cl} \rightarrow {}^{36}\mathrm{Cl} + \sum_i \gamma_i$. This process has a thermal cross-section much larger than the deuterium one (43.6 b versus 0.5 mb); and as a result of the higher energy released in multiple γ rays totalling 8.579 MeV [38], it allows to better discriminate neutron captures from background events and the Cherenkov electron signal. In the third phase of the experiment, the salt was removed and ^{3}He proportional counters were deployed in the heavy water to detect the neutrons using the breakup reaction $^3\mathrm{He}(n,p)^3\mathrm{H}$, which releases a total of 765 keV as kinetic energy of the products and has an even larger thermal cross-section (5330 b) and different systematic uncertainties.

The three phases produced consistent results [47, 48, 49]. Quoting here the ratios of the observed flux in the salt phase to the prediction in the null-oscillation hypothesis:

$$\frac{\Phi^{\mathrm{ES}}_{\mathrm{SNO}}}{\Phi_{\mathrm{SSM}}} = 0.406 \pm 0.046 \tag{1.37}$$

$$\frac{\Phi^{\mathrm{CC}}_{\mathrm{SNO}}}{\Phi_{\mathrm{SSM}}} = 0.290 \pm 0.017 \tag{1.38}$$

$$\frac{\Phi^{\mathrm{NC}}_{\mathrm{SNO}}}{\Phi_{\mathrm{SSM}}} = 0.853 \pm 0.075 \tag{1.39}$$

The $\Phi^{\mathrm{ES}}_{\mathrm{SNO}}$ result is in good agreement with those from Kamiokande (eq. 1.33) and Super-Kamiokande (eq. 1.34), and confirms once again the ν_e deficit. $\Phi^{\mathrm{CC}}_{\mathrm{SNO}}$, which is only sensitive to the electron flavor, indicates also a deficit of different magnitude which is consistent with the chlorine experiment (equation 1.28). However, the NC channel result is compatible with the prediction within uncertainty. Furthermore, the ratio [48]

$$\frac{\Phi^{\mathrm{CC}}_{\mathrm{SNO}}}{\Phi^{\mathrm{NC}}_{\mathrm{SNO}}} = 0.340 \pm 0.023(\mathrm{stat})^{+0.029}_{-0.031}(\mathrm{syst}) \tag{1.40}$$

demonstrates in a solar-model-independent way that there is a disappearance of the ν_e flavor in the solar neutrinos on their way from the core of the Sun to the Earth.

Moreover, the unknown electron $\Phi^{\nu_e}_{\mathrm{SNO}}$ and muon+tau flavor $\Phi^{\nu_{\mu,\tau}}_{\mathrm{SNO}}$ fluxes can be found out from the overdetermined system of equations:

$$\Phi^{\nu_e}_{\mathrm{SNO}} + r^{\mathrm{ES}}\,\Phi^{\nu_{\mu,\tau}}_{\mathrm{SNO}} = \Phi^{\mathrm{ES}}_{\mathrm{SNO}} \tag{1.41a}$$

$$\Phi^{\nu_e}_{\mathrm{SNO}} = \Phi^{\mathrm{CC}}_{\mathrm{SNO}} \tag{1.41b}$$

$$\Phi^{\nu_e}_{\mathrm{SNO}} + \Phi^{\nu_{\mu,\tau}}_{\mathrm{SNO}} = \Phi^{\mathrm{NC}}_{\mathrm{SNO}} \tag{1.41c}$$

where r^{ES} is the ratio of the elastic scattering cross sections for the muon+tau and electron flavors, $r^{\mathrm{ES}} \equiv \sigma^{\mathrm{ES}}_{\nu_{\mu,\tau}}/\sigma^{\mathrm{ES}}_{\nu_e} \approx 0.1553$. This system is represented graphically in figure 1.4, which shows that a solution can be found within uncertainty. The non-null value of $\Phi^{\nu_{\mu,\tau}}_{\mathrm{SNO}}$ is an evidence of the occurrence of neutrino conversion to these flavors.

When the data from all solar neutrino experiments are used (see figure 1.5), the observed ν_e deficit is interpreted as a consequence of neutrino oscillation driven by $\sin^2(2\theta_{12}) \sim 0.85$ and $\Delta m^2_{21} \sim 7.5 \cdot 10^{-5}\,\mathrm{eV}^2$ (so $m_2 > m_1$). This is known as the large mixing angle (LMA) solution, which is influenced by matter effects within the Sun (the MSW effect referred previously). Electron neutrinos are created within the core of the Sun, where the electron density is the highest, and propagate outwards in an electron-density decreasing medium which can be approximated by an exponential. Since the mixing angle within the Sun (see equation 1.23) depends both on energy and a density which varies, this leads to different oscillation probabilities for different energy ranges: for low energy ($< 1\,\mathrm{MeV}$) neutrinos, the resonant density (equation 1.25) is higher than the density in the the core of the Sun, and the oscillations proceed practically as in vacuum. For higher energies ($> 5\,\mathrm{MeV}$), the density of the solar core is above the resonant density, so neutrinos in their way out of the Sun traverse a resonant region where the oscillation is enhanced. This provides an explanation for the discrepancies in the values of the deficit between the experiments, which have different energy thresholds to detect the neutrinos.

The MSW-LMA solution of the solar neutrino problem has been tested by the Borexino experiment located at Gran Sasso [52]. Neutrinos are detected via elastic scattering on electrons (equation 1.32) using a $\sim$ 300 kton liquid scintillator detector of exceptional radiopurity, which allows to explore both the high-energy and the low-energy neutrinos (see figure 1.6).

In order to confirm the previous interpretation of the solar neutrino deficit, the KamLAND experiment was designed. The KamLAND detector is a 1 kton liquid scintillator detector located at the old Kamiokande site, and it detects the

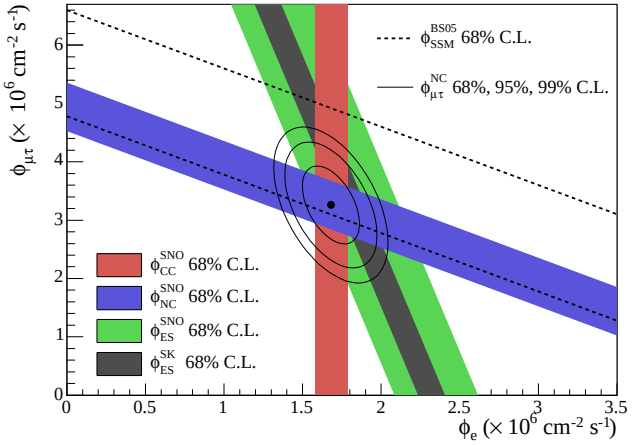


Figure 1.4: The horizontal axis shows the electron neutrino flux and the vertical axis shows the muon+tau neutrino flux. The SNO flux measurements in the salt phase are shown as filled bands (of width the $\pm 1\sigma$ uncertainty): elastic scattering (ES, green), charged-current (CC, red) and neutral-current (NC, blue). The elastic scattering from Super-Kamiokande (SK) [50] is also shown as a dark gray band. The ^{8}B flux prediction from the standard solar model [40] is delimited by dashed lines. The point shows the interception of the SNO CC and NC measurements, with the 68%, 95% and 99% C.L. contours. Figure from [48].

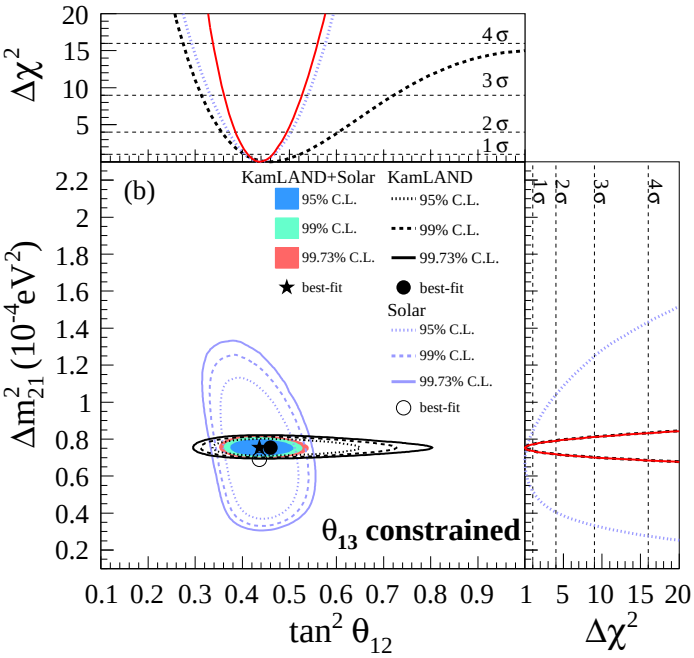


Figure 1.5: Allowed regions on the $\tan^2\theta_{12} - \Delta m^2_{21}$ plane from solar (violet contours), KamLAND (black contours) and combination of solar and KamLAND data (shaded regions). The upper and right panels show the projections of the $\Delta\chi^2$ profiles onto the $\tan^2\theta_{12}$ and Δm^2_{21} axes, respectively. θ_{13} is constrained by the measurements from the experiments described in sec. 1.2.3. From [51].

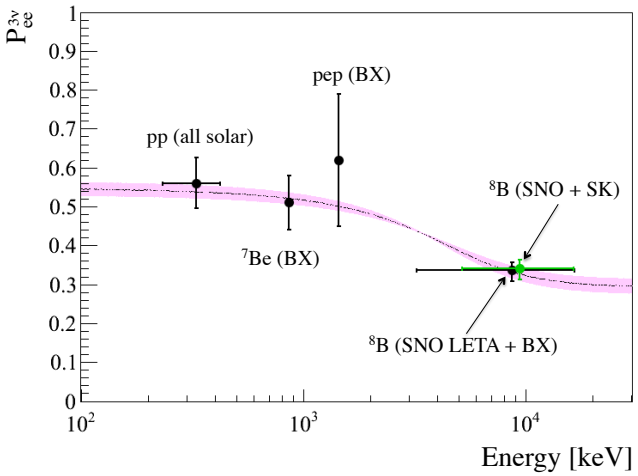


Figure 1.6: Solar electron neutrino survival probability as a function of the neutrino energy. The band represents the MSW-LMA solution with 1σ uncertainty. The points show measurements from the Borexino (BX), SNO, Super-Kamiokande (SK) and the combination of all solar experiments (radiochemical included). From [52].

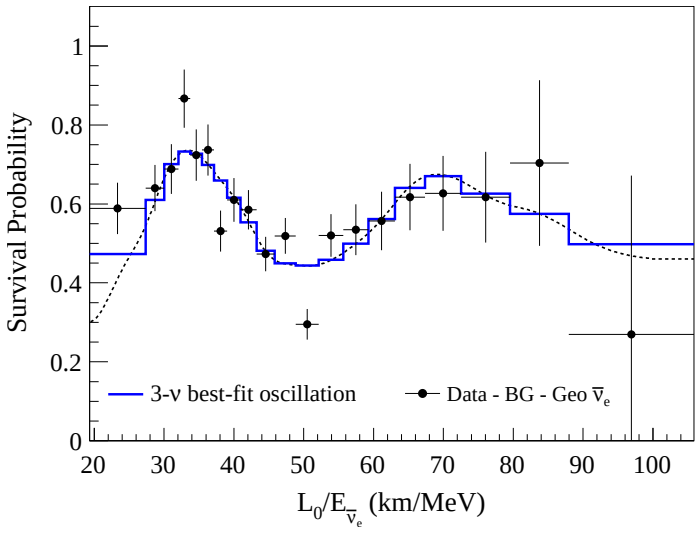


Figure 1.7: Ratio of the observed data in the KamLAND experiment (after subtraction of backgrounds and $\overline{\nu}_e$ produced in the Earth) to the no-oscillation prediction, as a function of L_0/E; where $L_0 = 180\,\text{km}$ is the flux-weighted average baseline to the reactor cores and E is the $\overline{\nu}_e$ energy. The blue histogram shows the best fit to a three-flavor survival probability. From [51].

electron antineutrinos emitted from the Japanese nuclear reactors, at an average distance of 180 km (with a $\sim 2.5\%$ contribution of Korean reactors and $< 1\%$ of other foreign reactors). The $\overline{\nu}_e$ are detected using the reaction $\overline{\nu}_e + p \rightarrow e^+ + n$, which is also used by Double Chooz and it is discussed in detail in section 2.2. The survival probability of ν_e must be equal to that of $\overline{\nu}_e$ by CPT symmetry, so the solar neutrino deficit can be reproduced using antineutrinos from nuclear reactors. The mean energy of reactor $\overline{\nu}_e$ is $\sim 4\,\text{MeV}$, which made the experiment sensitive to oscillations driven by $\Delta m^2_{21} \sim \mathcal{O}(10^{-5}\,\text{eV}^2)$, which is the squared-mass difference suggested by the combination of the solar experiments.

For the relevant length traveled by the antineutrinos at KamLAND, the oscillation due to Δm^2_{31} is averaged out, and the three-flavor survival probability can be approximated by

$$P_{ee}(L,E) = \cos^4(\theta_{13}) P^{2\nu}_{ee,m}(L,E) + \sin^4(\theta_{13}), \tag{1.42}$$

where $P^{2\nu}_{ee,m}$ is the two-flavor survival probability in matter and is given by

$$P^{2\nu}_{ee,m}(L,E) = 1 - \sin^2(2\theta_{12,m}) \sin^2\left(\frac{\Delta m^2_{21,m} L}{4E}\right), \tag{1.43}$$

where $\theta_{12,m}$ and $\Delta m^2_{21,m}$ are the θ_{12} mixing angle and the Δm^2_{21} squared-mass difference in matter, respectively, which can be obtained from equations 1.23 and 1.24 and replacing N_e by $N_e \cos^2(2\theta_{13})$.

KamLAND succeeded in observing an $\overline{\nu}_e$ deficit. Indeed, it has been able to study the $\overline{\nu}_e$ deficit as a function of the $\overline{\nu}_e$ energy, demonstrating the oscillatory behavior which is characteristic of neutrino oscillations, as figure 1.7 shows.

The combination of the solar experiments, which constrain best θ_{12}, and the KamLAND experiment, which constrains best Δm^2_{21}, produces the most precise measurements of θ_{12} and Δm^2_{21} (see figure 1.5) [51, 10]:

$$\sin^2(2\theta_{12}) = 0.846 \pm 0.021 \tag{1.44}$$

$$\Delta m^2_{21} = (7.53 \pm 0.18) \cdot 10^{-5}\,\mathrm{eV}^2 \tag{1.45}$$

Leaving θ_{13} as a free parameter or constraining it by the measurements from the experiments introduced in section 1.2.3 does not affect these values.

1.2.2 Measurement of θ_{23} and $|\Delta m^2_{32}|$

The sequence of experiments that led to the measurement of the θ_{23} mixing angle and the Δm^2_{32} squared-mass difference absolute value (since the sign is still unknown) was initiated by the *atmospheric neutrino anomaly*. The atmospheric neutrinos are the neutrinos and antineutrinos which are created as a result of the interaction of primary cosmic rays with atoms in the atmosphere. These energetic collisions generate hadrons, mostly pions (and kaons at higher energies) which decay dominantly into muons and neutrinos:

$$\pi^+ \to \mu^+ + \nu_\mu, \qquad \pi^- \to \mu^- + \overline{\nu}_\mu. \tag{1.46}$$

The muons themselves decay into electrons (or positrons), producing more neutrinos:

$$\mu^+ \to \overline{\nu}_\mu + \nu_e + e^+, \qquad \mu^- \to \nu_\mu + \overline{\nu}_e + e^-. \tag{1.47}$$

Therefore, at low energies ($\lesssim 1\,\mathrm{GeV}$) in order to let the muons decay before reaching the detectors, the expected flavor composition of the atmospheric neutrino flux is $\Phi(\nu_\mu + \overline{\nu}_\mu) : \Phi(\nu_e + \overline{\nu}_e) \approx 2 : 1$.

An anomalous deficit of the muon-flavor in the atmospheric neutrino flux was reported by the Kamiokande and IMB experiments [53, 54]. Both experiments were searching for the nucleon decay with water Cherenkov detectors, and the atmospheric neutrinos represented an undesired background which had to be modeled. For that reason, efforts were made to produce predictions of the flux. These neutrinos were detected through charged-current interactions with the nuclei in the detector:

$$\nu_l + N \to l^- + X, \qquad \overline{\nu}_l + N \to l^+ + X, \tag{1.48}$$

in which the lepton is revealed by the Cherenkov radiation emitted. Since the detectors were not magnetized, no charge discrimination was possible. In addition, tau leptons were very difficult to detect, since they decay almost immediately. In order to cancel theoretical and experimental uncertainties, the measurements can be reported as double ratios between the observed data and the Monte Carlo prediction, finding for both experiments:

$$\frac{\Phi(\nu_\mu + \overline{\nu}_\mu)/\Phi(\nu_e + \overline{\nu}_e)\Big|_{\mathrm{DATA}}}{\Phi(\nu_\mu + \overline{\nu}_\mu)/\Phi(\nu_e + \overline{\nu}_e)\Big|_{\mathrm{MC}}} \approx 0.6, \tag{1.49}$$

which shows a clear discrepancy with respect to the prediction.

The atmospheric neutrino flux was precisely measured by the Super-Kamiokande experiment (already introduced in the context of solar neutrinos) as a function of the distance L traveled (inferred from the zenith angle Θ) and the neutrino energy (see figure 1.8). The Cherenkov detection technique provides directional information which allows to distinguish if the neutrinos come directly from above (downward charged leptons), $\cos\Theta \approx 1$ and $L \sim 20\,\text{km}$, or have crossed the Earth on their way to the detector (upward charged leptons), $\cos\Theta \approx -1$ and $L \sim 13000\,\text{km}$. Super-Kamiokande found a deficit in the muon neutrinos with respect to the no-oscillation prediction which increased with the distance traveled by the neutrino, while the electron neutrinos agreed with the prediction. Therefore, the result was interpreted as $\nu_\mu \to \nu_\tau$ oscillation. Since ν_μ and ν_τ interact equally with the Earth, no matter effects apply. Neglecting the small Δm^2_{21}, the ν_μ survival probability can be approximated by:

$$P_{\mu\mu}(L,E) \approx 1 - 4\cos^2(\theta_{13})\sin^2(\theta_{23})\left[1-\cos^2(\theta_{13})\sin^2(\theta_{23})\right] \times \\ \times \sin^2\left(1.27\frac{\Delta m^2_{\text{atm}}[\,\text{eV}^2]L[\,\text{km}]}{E[\,\text{GeV}]}\right), \tag{1.50}$$

where $\Delta m^2_{\text{atm}} = \Delta m^2_{32}$ for normal mass hierarchy or $\Delta m^2_{\text{atm}} = \Delta m^2_{13}$ for inverted mass hierarchy, and the oscillatory phase is written in the usual units used by these experiments. The analysis of the Super-Kamiokande data [55] found the best fit at $\sin^2(\theta_{23}) = 0.5$, $\Delta m^2_{\text{atm}} = 2.1 \cdot 10^{-3}\,\text{eV}^2$ and $\sin^2(\theta_{13}) = 0.0$. Similar results were found by the Soudan 2 experiment [56], a 960 ton iron tracking calorimeter placed at the Soudan mine (Minnesota, USA); and by the MACRO experiment [57], using scintillator counters and streamer tubes, which was located at the Gran Sasso laboratory.

As in the case of the solar neutrinos, in order to reproduce the muon-flavor deficit observed in the atmospheric neutrinos, human-made neutrinos were used. Muon neutrino beams can be produced colliding protons with an energy of several GeV against a fixed target, creating secondary hadron showers. A magnetic device called horn focuses the resulting pions and kaons, which are allowed to decay in flight for $\sim 100\,\text{m}$ until they are stopped by an absorber material. The resulting beam is composed mostly of muon neutrinos with a small component of electron neutrinos. The beam energy is typically of a few GeV, so in order to explore squared-mass differences such as the atmospheric one of $\mathcal{O}(10^{-3}\,\text{eV}^2)$, baselines of several hundreds of kilometers are required.

The K2K experiment was the first long-baseline experiment which used this kind of beam. A $\langle E_\nu \rangle \sim 1.3\,\text{GeV}$ beam was produced at the KEK proton synchrotron and sent towards the Super-Kamiokande detector. The neutrino flux, energy spectrum and profile of the beam were measured by a near multi-detector system at $\sim 300\,\text{m}$ from the production target. Even though the baseline was not optimal, $L = 250\,\text{km}$, determined by the pre-existing locations of KEK and Super-Kamiokande, the K2K experiment confirmed the oscillation parameters obtained from the atmospheric neutrinos [58].

The second long-baseline experiment built is the MINOS experiment. The neutrino beam, peaked at $\sim 3\,\text{GeV}$, is produced using the protons extracted from the Main Injector at the Fermi National Accelerator Laboratory and sent towards a far detector located 735 km away from the target, in the Soudan mine. A near detector 1 km downstream the target characterizes the neutrino beam before the oscillation is developed, in order to predict the unoscillated flux and

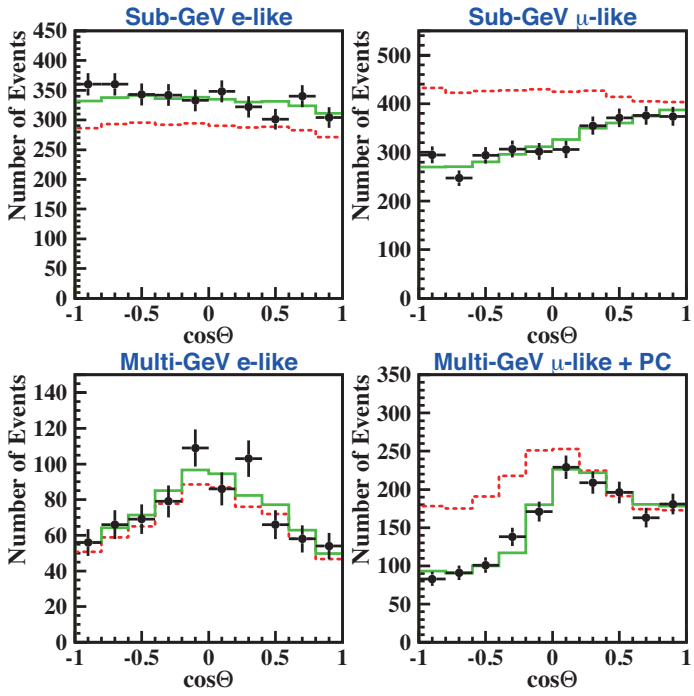


Figure 1.8: Zenith angle distributions of the atmospheric neutrino candidates observed by Super-Kamiokande. Electron-like events are shown on the left, and muon-like ones on the right. The upper plots correspond to visible energies $< 1.33\,\text{GeV}$ (Sub-GeV) and the lower plots to energies $> 1.33\,\text{GeV}$ (Multi-GeV). The dotted histograms show the no-oscillation Monte Carlo prediction, and the solid histograms show the best-fit oscillation prediction. From [10].

spectrum at the far detector. Both near (0.98 kton) and far (5.4 kton) detectors are steel-scintillator magnetized tracking calorimeters. MINOS has produced the most precise measurement of Δm^2_{32} combining its data from accelerator and atmospheric neutrinos and analyzing them in a three-flavor formalism [59]:

$$|\Delta m^2_{32}| = (2.37 \pm 0.09) \cdot 10^{-3}\,\mathrm{eV}^2 \quad \text{(NH)}, \quad (2.41^{+0.12}_{-0.09}) \cdot 10^{-3}\,\mathrm{eV}^2 \quad \text{(IH)}, \tag{1.51}$$

where $\Delta m^2_{32} > 0$ for normal hierarchy (NH) and $\Delta m^2_{32} < 0$ for inverted hierarchy (IH). The data includes the leading channel ${}^{(-)}\nu_\mu \to {}^{(-)}\nu_\mu$, and the subdominant channel ${}^{(-)}\nu_\mu \to {}^{(-)}\nu_e$, which is driven by θ_{13} and experiences matter effects (see equation 1.68 in the next section). The fit takes θ_{12} and Δm^2_{21} as fixed inputs, and θ_{13} is constrained by the measurements from the current reactor experiments introduced in the following section. In addition, the fit finds at 90% C.L. $\sin^2(2\theta_{23}) = 0.97^{+0.03}_{-0.06}$ (NH) or $\sin^2(2\theta_{23}) = 0.97^{+0.03}_{-0.09}$ (IH).

An evolution of the K2K experiment, called T2K, has been built. Super-Kamiokande is still used as far detector, but the muon neutrino beam is produced at J-PARC (Tokai, Japan), located 295 km away. The neutrino beam axis is aimed 2.5° off from Super-Kamiokande, resulting in a more monochromatic beam peaked at ~ 0.6 GeV, which enhances the neutrino conversion for the squared-mass difference studied, and a lower background contamination. A near detector complex at 280 m from the fixed target characterizes the unoscillated neutrino beam with multiple detectors placed both on-axis and off-axis. T2K has provided the most precise value of θ_{23} [60, 10]:

$$\sin^2(2\theta_{23}) = 0.999^{+0.001}_{-0.018} \quad \text{(NH)}, \quad 1.000^{+0.000}_{-0.017} \quad \text{(IH)}, \tag{1.52}$$

which depends slightly on the mass hierarchy. Both cases are compatible with maximal mixing, that is, $\theta_{23} = \pi/4$. In addition, T2K measures $\Delta m^2_{32} = (2.51 \pm 0.10) \cdot 10^{-3}\,\mathrm{eV}^2$ (NH) and $\Delta m^2_{13} = (2.48 \pm 0.10) \cdot 10^{-3}\,\mathrm{eV}^2$, which are consistent with the measurements from MINOS.

The dominant transformation in the atmospheric sector, $\nu_\mu \to \nu_\tau$, has been explicitly tested by observing the appearance of 5 τ leptons at the OPERA experiment using the CNGS muon neutrino beam [61]. The beam is produced colliding 400 GeV protons from the Super Proton Synchrotron (SPS) against a fixed target at CERN, resulting in a muon neutrino beam with an average energy of ~ 17 GeV. In order to detect the short-lived τ resulting from a ν_τ charged-current interaction, the OPERA detector uses nuclear emulsion films interleaved with large lead plates.

1.2.3 Measurement of θ_{13}

The neutrino mixing angle θ_{13} was the last angle to be measured. Unlike the other two angles θ_{12} and θ_{23}, it is small. In fact, when the research that has led to this thesis was started, only upper limits had been established.

The ideal neutrino source to measure θ_{13} are nuclear reactors. The daughter nuclei resulting from the fission of the nuclei making up the nuclear fuel have a neutron-excess, and decay into stable nuclei after several β^- decays, in which electron antineutrinos are produced. The result is an intense, isotropic, pure-$\overline{\nu}_e$ flux at the origin, which can be used to study the neutrino flavor oscillations. Since these neutrinos have energies of a few MeV, they are not energetic enough

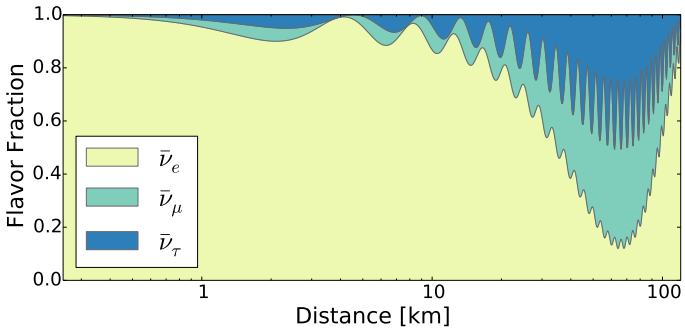


Figure 1.9: The flavor composition of a 4 MeV reactor antineutrino flux as a function of the distance to the reactor cores using the mixing angles and squared-mass differences from [10]. The "fast" oscillation due to Δm^2_{31}, Δm^2_{32} and the "slow" due to Δm^2_{21} are visible. The current experiments designed to measure θ_{13} have a near detector close to the reactors, where the oscillation is barely developed, and a far detector sitting close to the first local minimum at ~ 2 km. The fraction of non-electron antineutrinos at this distance is equal to $\sin^2(2\theta_{13})$. From [62].

to produce heavy leptons as the muon or tau upon charged-current interaction. The only experimentally accessible channel is $\overline{\nu}_e \to \overline{\nu}_e$, in which the electron neutrino is typically detected in the inverse β-decay (IBD) reaction, $\overline{\nu}_e + p \to e^+ + n$ (which is examined in detail in section 2.2).

From equation 1.16, the survival probability for electron antineutrino in vacuum can be written as

$$\begin{aligned} P_{ee}(L,E) = 1 &- \cos^4(\theta_{13})\sin^2(2\theta_{12})\sin^2\left(\frac{\Delta m^2_{21}L}{4E}\right) \\ &- \cos^2(\theta_{12})\sin^2(2\theta_{13})\sin^2\left(\frac{\Delta m^2_{31}L}{4E}\right) \\ &- \sin^2(\theta_{12})\sin^2(2\theta_{13})\sin^2\left(\frac{\Delta m^2_{32}L}{4E}\right). \end{aligned} \tag{1.53}$$

The experiments searching for θ_{13} have baselines of ~ 1 km. For these distances matter effects are negligible; furthermore, the oscillation driven by Δm^2_{21} has not developed yet and the second term can be neglected (see figure 1.9). This implies effectively $\Delta m^2_{21} \approx 0$, so $\Delta m^2_{31} \simeq \Delta m^2_{32}$. Then, the survival probability for $\overline{\nu}_e$ is simplified to

$$\begin{aligned} P_{ee}(L,E) &\simeq 1 - \sin^2(2\theta_{13})\sin^2\left(\frac{\Delta m^2_{31}L}{4E}\right) \\ &\approx 1 - \sin^2(2\theta_{13})\sin^2\left(1.27\frac{\Delta m^2_{31}[\,\mathrm{eV}^2]L[\,\mathrm{m}]}{E[\,\mathrm{MeV}]}\right), \end{aligned} \tag{1.54}$$

which has the form of the two-flavor approximation (equation 1.20). In the second line, the usual units used in reactor experiments have been explicitly written for convenience.

The tightest upper limit was set by the CHOOZ [5] experiment, a precursor of the Double Chooz experiment. The CHOOZ detector (see figure 1.10) was $\approx$ 1 km away from two powerful nuclear reactors producing 4.25 GW_{th} each at full power. The detector started taking data before the nuclear reactors started their operation, and $\sim$ 143 live-days of data with both reactors-off were acquired which provided a direct measurement of the backgrounds on-site. The ratio of the background-subtracted observed $\overline{\nu}_e$ candidates to the no-oscillation prediction was found to be

$$R_{\text{CHOOZ}} = 1.01 \pm 2.8\%(\text{stat}) \pm 2.7\%(\text{syst}), \tag{1.55}$$

which showed no indication of a $\overline{\nu}_e$ deficit induced by oscillation into other flavors. This allowed to exclude a large portion on the $\sin^2(2\theta_{13}) - \Delta m^2_{31}$ plane (see figure 1.11). In particular, for $\Delta m^2_{31} = 2.5 \cdot 10^{-3}\,\text{eV}^2$, $\sin^2(2\theta_{13}) < 0.14$ at 90% C.L. [63].

A similar experiment was carried out at the Palo Verde nuclear power plant in Arizona (USA) [65], finding a ratio between the observed candidate rate and the expected one in case of no-oscillation of

$$R_{\text{Palo Verde}} = 1.01 \pm 0.024(\text{stat}) \pm 0.053(\text{syst}), \tag{1.56}$$

which agrees with the CHOOZ conclusion but with bigger uncertainty, resulting in a less stringent exclusion.

In order to measure θ_{13} or improve the limit by CHOOZ, a new generation of reactor antineutrino experiments was proposed. These experiments are Double Chooz [66], Daya Bay [67] and RENO [68] (see table 1.3 for the main features). All of them use a similar detector design based on that from CHOOZ, with improvements (the Double Chooz detector is presented in detail in section 2.3). Since the CHOOZ result was limited not only by statistics due to the untimely conclusion (because of scintillator instabilities) but also from systematic uncertainties, in which the leading contributions were (in order) the reactor-flux prediction and the neutrino detection uncertainties; these experiments were conceived as two-detector experiments, in which a near detector measures the almost-unoscillated flux from the nuclear reactors and a far detector measures the flux after the oscillation has developed, in the fashion of the accelerator-based experiments in section 1.2.2. By building the two detectors as identical as possible, the systematic uncertainties due to the reactor prediction and the neutrino detection are minimized (see table 1.4).

The three experiments have reported non-null measurements of θ_{13}. The first one was by Double Chooz using only the far detector (the Double Chooz experiment is reviewed in chapter 2). Double Chooz measured an observed-to-expected ratio of

$$R_{\text{Double Chooz}} = 0.944 \pm 0.016(\text{stat}) \pm 0.040(\text{syst}). \tag{1.57}$$

An analysis based on the normalization and the shape of the positron spectrum (known as *Rate+Shape analysis*, described in chapter 6.2) found the best fit at

$$\sin^2(2\theta_{13}) = 0.086 \pm 0.041(\text{stat}) \pm 0.030(\text{syst}), \tag{1.58}$$

[5] The CHOOZ experiment, usually written capitalized, is named after the nuclear power plant Chooz-B, which is used as the reactor antineutrino source. The power plant itself is named after the close village of Chooz, in the French Ardennes.

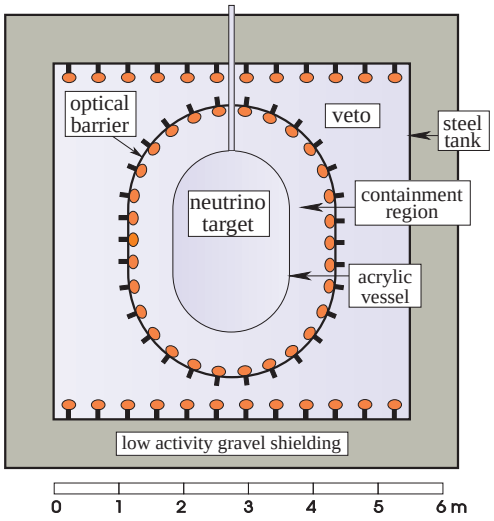


Figure 1.10: The CHOOZ detector. The "neutrino target" is a transparent acrylic vessel filled with 5 ton of liquid scintillator loaded with Gd (at 0.09%) to enhance the capture of the neutrons created by the $\overline{\nu}_e$. The "containment region" is a 70 cm thick liquid scintillator (undoped) layer enclosing the target to contain the γ-rays from neutron captures and shield the target from the radioactivity of 192 eight-inch PMTs used to detect the scintillation light. Surrounding this volume and optically separated from it, there is a 80 cm thick cosmic muon veto volume filled with undoped liquid scintillator and observed by two rings of 24 eight-inch PMTs, one at the top and one at the bottom. The vessel is made of stainless steel with the inner walls painted with a high-reflectivity white paint. The full detector is surrounded by 75 cm of low radioactivity sand and covered by 14 cm of cast iron. From [64].

Experiment	Reactor power (GW_{th})	Distance (m) Near / Far	Depth (mwe) Near / Far	Target mass (ton) $\times$ detectors
Double Chooz	8.5	400 / 1050	120 / 300	8×2
Daya Bay	17.4	470, 576 / 1648	260 / 860	20×8
RENO	16.5	294 / 1383	120 / 450	16×2

Table 1.3: Comparison between current reactor antineutrino experiments. The reactor power corresponds to the total nominal thermal power (summed over all cores). The distance is the average detector-reactor distance (in Daya Bay is flux-weighted). The target mass corresponds to the Gd-loaded liquid scintillator mass per detector.

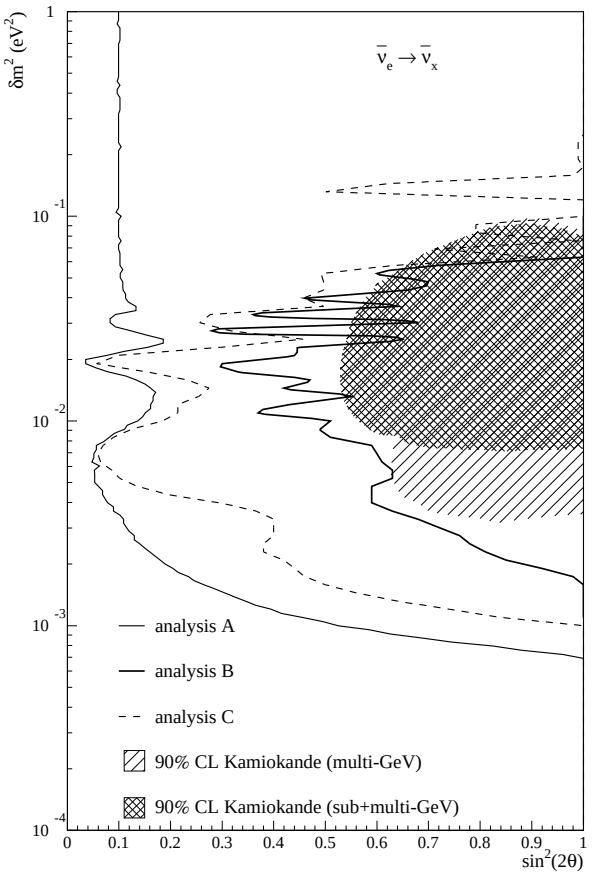


Figure 1.11: Exclusion contours at 90% C.L. in the $\sin^2(2\theta_{13}) - \Delta m^2_{31}$ plane from the final analysis of the CHOOZ data. Analysis A compared the normalization and the shape of the background-subtracted positron spectrum observed to the predicted one; analysis B used the ratio of the measured spectra from each reactor; analysis C corresponds to leaving the normalization free in analysis A (shape-only analysis). The allowed region for $\nu_\mu \leftrightarrow \nu_e$ from the Kamiokande atmospheric neutrino data is shown for comparison. From [63].

Systematic uncertainty	CHOOZ	Double Chooz
Reactor prediction	2.1%	$< 0.1\%$
Antineutrino detection	1.5%	$< 0.5\%$
Number of protons	0.8%	$< 0.2\%$
Total	2.7%	$< 0.5\%$

Table 1.4: Leading systematic uncertainties affecting the CHOOZ experiment and their expectation in Double Chooz (with two-detectors).

excluding the no-oscillation hypothesis at 94.6% C.L. [69].

This result was followed by the Daya Bay and RENO experiments, with both near and far detectors, which found the ratios:

$$R_{\text{Daya Bay}} = 0.940 \pm 0.011(\text{stat}) \pm 0.004(\text{syst}), \tag{1.59}$$
$$R_{\text{RENO}} = 0.920 \pm 0.009(\text{stat}) \pm 0.014(\text{syst}). \tag{1.60}$$

Analyses based exclusively on the total number of $\overline{\nu}_e$ (*rate-only analysis*) measured

$$\sin^2(2\theta_{13}) = 0.092 \pm 0.016(\text{stat}) \pm 0.005(\text{syst}), \tag{1.61}$$
$$\sin^2(2\theta_{13}) = 0.113 \pm 0.013(\text{stat}) \pm 0.019(\text{syst}), \tag{1.62}$$

for Daya Bay and RENO, respectively; which showed $\theta_{13} \neq 0$ with significances of 5.2 and 4.9 standard deviations [70, 71].

Double Chooz continued accumulating data with the one-detector configuration, releasing a second result [1] with increased statistics and an improved analysis shortly after:

$$\sin^2(2\theta_{13}) = 0.109 \pm 0.030(\text{stat}) \pm 0.025(\text{syst}). \tag{1.63}$$

In addition, Double Chooz produced the first θ_{13} measurement [72] using an independent data sample provided by the capture of the IBD neutrons on hydrogen nuclei instead of the intended gadolinium:

$$\sin^2(2\theta_{13}) = 0.097 \pm 0.034(\text{stat}) \pm 0.034(\text{syst}). \tag{1.64}$$

These two results were combined to find $\sin^2(2\theta_{13}) = 0.109 \pm 0.035(\text{total})$ [73]. Moreover, as it will be explained in section 6.1, Double Chooz is the unique experiment with a background-model independent analysis [74]. When it is applied to the same data set as the previous results, the best fit is found at:

$$\sin^2(2\theta_{13}) = 0.107 \pm 0.049 \quad (\text{Gd}), \tag{1.65}$$
$$\sin^2(2\theta_{13}) = 0.091 \pm 0.078 \quad (\text{H}), \tag{1.66}$$
$$\sin^2(2\theta_{13}) = 0.102 \pm 0.043 \quad (\text{Gd+H}). \tag{1.67}$$

A compilation of the abovesaid θ_{13} measurements is presented in figure 1.12 .

The latest θ_{13} results from Double Chooz are covered in this thesis. Naturally, the Daya Bay and RENO experiments have been updating their analyses too. Their most recent values can be found in the conclusions (chapter 7).

The neutrino mixing angle θ_{13} can also be measured using the subdominant channel $\nu_\mu \rightarrow \nu_e$, looking for the appearance of ν_e in the ν_μ beam in accelerator-based experiments such as those designed to measure θ_{23} and $|\Delta m^2_{32}|$ (section 1.2.2). In fact, this transition was observed for the first time by the T2K [75] and MINOS [76] experiments, providing the first hints of a non-null θ_{13}. For a small θ_{13}, as it happens to be the case, the $\nu_\mu \rightarrow \nu_e$ transition probability is

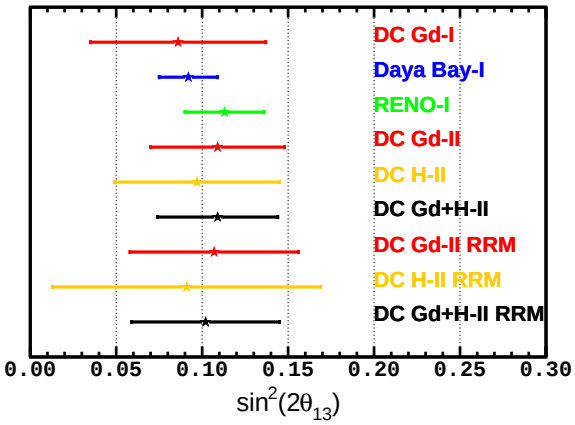


Figure 1.12: Compilation of the first measurements of $\sin^2(2\theta_{13})$ by the reactor antineutrino experiments (see text for details). Roman numbers correspond to the data release. "RRM" indicates the background-model independent analysis.

written up to second order in $\alpha \equiv \Delta m^2_{21}/\Delta m^2_{32}$ as [77]:

$$\begin{aligned} P_{\mu e}(L,E) = & \frac{1}{(A-1)^2}\sin^2(2\theta_{13})\sin^2(\theta_{23})\sin^2\left[(A-1)\Delta\right] \\ & \mp \frac{\alpha}{A(1-A)}\cos(\theta_{13})\sin(2\theta_{12})\sin(2\theta_{23})\sin(2\theta_{13})\times \\ & \times \sin(\delta)\sin(\Delta)\sin(A\Delta)\sin\left[(1-A)\Delta\right] \\ & + \frac{\alpha}{A(1-A)}\cos(\theta_{13})\sin(2\theta_{12})\sin(2\theta_{23})\sin(2\theta_{13})\times \\ & \times \cos(\delta)\cos(\Delta)\sin(A\Delta)\sin\left[(1-A)\Delta\right] \\ & + \frac{\alpha^2}{A^2}\cos^2(\theta_{23})\sin^2(2\theta_{12})\sin^2(A\Delta), \end{aligned} \tag{1.68}$$

where

$$\Delta \equiv \frac{\Delta m^2_{32} L}{4E},$$

the sign of the second line is negative for neutrinos and positive for antineutrinos, and the matter effects are encoded in the variable

$$A \equiv 2\sqrt{2} G_{\mathrm{F}} N_e \frac{E}{\Delta m^2_{32}}.$$

For accelerator-based experiments, N_e is the electron density of the Earth crust. The sign of A depends on the sign of Δm^2_{32}, which is not known yet. Moreover, the transition probability depends also on the CP-violating phase δ, which is not known either. Therefore, it is obvious that in order to produce a measurement of θ_{13}, some assumptions must be made, unlike in the simple case of reactor experiments (equation 1.54). For $\sin^2(\theta_{12}) = 0.306$, $\Delta m^2_{21} = 7.6 \cdot 10^{-5}\,\mathrm{eV}^2$, $\sin^2(\theta_{23}) = 0.5$, $|\Delta m^2_{32}| = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$ and $\delta = 0$, T2K finds the best-fit value

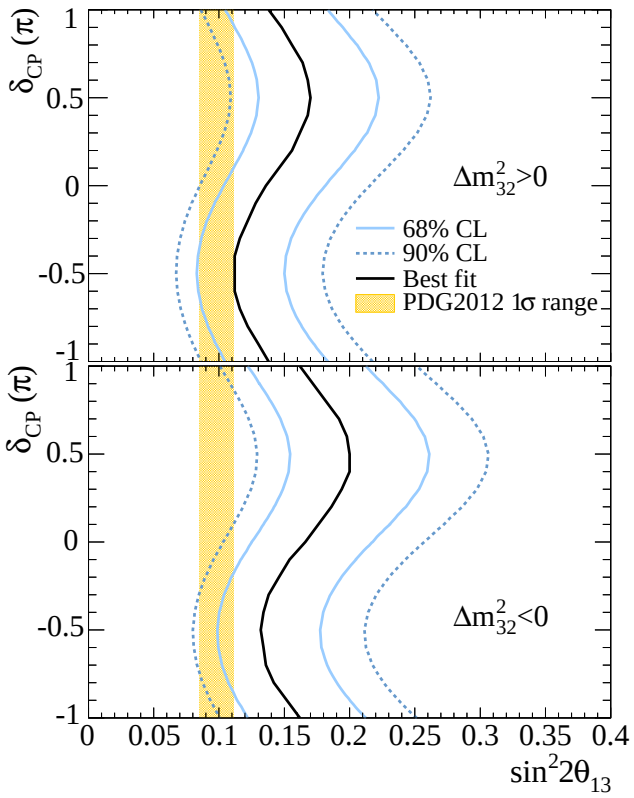


Figure 1.13: 68% and 90% allowed regions on the $\sin^2(2\theta_{13}) - \delta$ plane for normal hierarchy (top) and inverted hierarchy (bottom). The best fit for $\sin^2(2\theta_{13})$ for the given value of δ is shown as a solid line. θ_{23} and $|\Delta m^2_{32}|$ are varied in the fit constrained by the T2K own measurements. The vertical shaded region shows the average $\sin^2(2\theta_{13})$ from the PDG2012 $\sin^2(2\theta_{13}) = 0.098 \pm 0.013$. From [78] and references therein.

at [78]

$$\sin^2(2\theta_{13}) = 0.140^{+0.038}_{-0.032} \quad \text{(NH)}, \tag{1.69}$$

$$\sin^2(2\theta_{13}) = 0.170^{+0.045}_{-0.037} \quad \text{(IH)}, \tag{1.70}$$

where, as usual, NH stands for normal mass hierarchy ($\Delta m^2_{32} > 0$) and IH for inverted mass hierarchy ($\Delta m^2_{32} < 0$).

Nevertheless, although the dependence on the CP-violating phase δ is counterproductive for a clean measurement of θ_{13} by accelerator-based experiments, it has provided the first glimpse of the value of δ when the input from the reactors experiment on θ_{13} is used, as figure 1.13 shows. The preferred value is $\delta = \pi/2$, i.e. maximal CP violation, but the uncertainties are still large. However, δ values in the interval $(0.19\pi, 0.80\pi)$ in the normal hierarchy, and in $(\pi, 0.97\pi)$ and $(-0.04\pi, \pi)$ in the inverted hierarchy, are already excluded at 90% C.L.

1.3 Search for neutrino mass

Neutrino oscillation experiments tell us about the differences between the squared masses of the neutrino mass eigenstates, but only set a lower limit on the absolute mass scale: if the lightest neutrino mass is taken to be zero and using the squared-mass differences from [10], then for normal hierarchy:

$$m_2 \geq \sqrt{\Delta m^2_{21}} = 0.0087 \pm 0.0001\,\text{eV} \tag{1.71}$$

and

$$m_3 \geq \sqrt{\Delta m^2_{32} + \Delta m^2_{21}} = 0.0502 \pm 0.0006\,\text{eV}; \tag{1.72}$$

and for inverted hierarchy:

$$m_1 \geq \sqrt{|\Delta m^2_{32}| - \Delta m^2_{21}} = 0.0494 \pm 0.0007\,\text{eV} \tag{1.73}$$

and

$$m_2 \geq \sqrt{|\Delta m^2_{32}|} = 0.0502 \pm 0.0007\,\text{eV}; \tag{1.74}$$

In order to try to measure the absolute neutrino masses, other kinds of experiments must be carried out.

1.3.1 Kinematic limits

The mass of the neutrinos can be deduced from the conservation of energy and momentum in weak interactions involving neutrinos, providing a direct measurement.

In the β^- decay, the endpoint of the electron spectrum, which corresponds to the maximum kinetic energy carried away by the electron, $T_{\max}$, is sensitive to the electron antineutrino mass, $m_{\overline{\nu}_e}$:

$$T_{\max} = Q_\beta - m_{\overline{\nu}_e}, \tag{1.75}$$

where $Q_\beta = M_i - M_f - m_e$ is the Q-value of the reaction, with M_i and M_f the masses of the initial and final nuclei, and m_e the mass of the electron. In order to study this effect, nuclei with low Q-values are preferable to maximize the number of events at the endpoint. A good candidate is tritium, with a Q-value of 18.6 keV and a half-life of 12.3 years. Nevertheless, as a result of the neutrino mixing, $\overline{\nu}_e$ has not a definite mass. Then, the β^- decay of tritium is an incoherent (because of the different final states) sum of decay modes:

$$^3\text{H} \rightarrow {}^3\text{He} + e^- + \overline{\nu}_j, \tag{1.76}$$

where $\overline{\nu}_j$ are the antineutrino mass eigenstates. Therefore, the endpoint of the electron spectrum will exhibit a shift from the massless neutrinos case $T_{\max} = Q_\beta$ to $T_{\max} = Q_\beta - \min(m_j)$, where $\min(m_j)$ is the lightest mass, with additional kinks at $T = Q_\beta - m_k$, with $m_k \neq \min(m_j)$.

However, measuring the shift and the kinks is very difficult since the rate of events close to the endpoint energy tends to zero. Instead, it is more practical to measure deviations in the Kurie function defined as

$$K(T) \equiv \left[(Q_\beta - T) \sum_{j=1}^{3} |U_{ej}|^2 \sqrt{(Q_\beta - T)^2 - m_j^2} \right]^{\frac{1}{2}}, \tag{1.77}$$

where T is the kinetic energy of the emitted electron and U_{ej} is the j-th matrix element from the electron-neutrino row of the PMNS matrix (equation 1.5). If the resolution of the experimental apparatus is such that it cannot resolve the individual masses, the Kurie plot can be approximated considering $Q_\beta - T >> m_j$ to

$$K(T) \simeq \left[(Q_\beta - T) \sqrt{(Q_\beta - T)^2 - m_\beta^2} \right]^{\frac{1}{2}}, \tag{1.78}$$

where m_β

$$\begin{aligned} m_\beta &= \left[\sum_{j=1}^{3} |U_{ej}|^2 m_j^2 \right]^{\frac{1}{2}} \\ &= \left[c_{12}^2 c_{13}^2 m_1^2 + s_{12}^2 c_{13}^2 m_2^2 + s_{13}^2 m_3^2 \right]^{\frac{1}{2}} \end{aligned} \tag{1.79}$$

is the effective electron antineutrino mass in β^- decay (which is equal to the electron neutrino mass if CPT symmetry holds).

No β^- decay experiment has measured a neutrino mass different from zero yet. The best upper limits come from the Troitsk [79] and Mainz [80] β^- spectrometers:

$$m_\beta < 2.05\,\text{eV} \quad \text{(Troitsk)} \tag{1.80}$$

$$m_\beta < 2.3\,\text{eV} \quad \text{(Mainz)} \tag{1.81}$$

at 95% C.L. Both collaborations have merged to perform an improved experiment in the near future, KATRIN, which will reach a sensitivity of 0.2 eV (90% C.L.) [81].

Less stringent limits have been set to the effective muon and tau neutrino masses, $(m_{\nu_\mu}^{\text{eff}})^2 = \sum_{j=1}^{3} |U_{\mu j}|^2 m_j^2$ and $(m_{\nu_\tau}^{\text{eff}})^2 = \sum_{j=1}^{3} |U_{\tau j}|^2 m_j^2$. In the case of the muon neutrino, the analysis of the decay $\pi^+ \to \mu^+ + \nu_\mu$ [82] results in

$$m_{\nu_\mu}^{\text{eff}} < 0.17\,\text{MeV (90\% C.L.)}. \tag{1.82}$$

For the tau neutrino, the analysis of the channels $\tau^- \to 2\pi^-\pi^+\nu_\tau$ and $\tau^- \to 3\pi^- 2\pi^+(\pi^0)\nu_\tau$ [83] gives

$$m_{\nu_\tau}^{\text{eff}} < 18.2\,\text{MeV (95\% C.L.)}. \tag{1.83}$$

1.3.2 Neutrinoless double beta decay

In section 1.1, the mechanism of Higgs-induced neutrino masses was put forward. However, this is not the only possibility. If right-handed neutrinos exist, since they are not charged under any of the Standard Model interactions,

another possibility to generate neutrino masses is given by (considering only one neutrino for simplicity, again):

$$\mathcal{L}_{\text{mass}}^{\text{Majorana}} = -\frac{1}{2}\overline{\nu_{\text{R}}} M_{\text{M}}^{\nu} \nu_{\text{R}}^{C} + \text{H.c.}, \tag{1.84}$$

where the charge-conjugated field $\nu_{\text{R}}^{C} = \mathcal{C}\overline{\nu_{\text{R}}}^{T}$, with $\mathcal{C}$ the charge conjugation operator, is left-handed (i.e. the right-handed projector P_{R} gives $P_{\text{R}}\nu_{\text{R}}^{C} = 0$); and M_{M}^{ν} is its Majorana mass [6].

This second mechanism of mass generation is attractive because it offers a plausible explanation to the smallness of the neutrino mass. If both Dirac and Majorana terms are present, the most general neutrino mass term is written

$$\begin{aligned} \mathcal{L}_{\text{mass}}^{\text{Dirac+Majorana}} &= -\overline{\nu_{\text{R}}} M_{\text{D}}^{\nu} \nu_{\text{L}} - \frac{1}{2}\overline{\nu_{\text{R}}} M_{\text{M}}^{\nu} \nu_{\text{R}}^{C} + \text{H.c.} \\ &= -\frac{1}{2} N_{\text{L}}^{T} \mathcal{C}^{\dagger} M N_{\text{L}}, \end{aligned} \tag{1.85}$$

where

$$N_{\text{L}} = \begin{pmatrix} \nu_{\text{L}} \\ \nu_{\text{R}}^{C} \end{pmatrix}, \tag{1.86}$$

and the mass matrix

$$M = \begin{pmatrix} 0 & M_{\text{D}}^{\nu} \\ M_{\text{D}}^{\nu} & M_{\text{M}}^{\nu} \end{pmatrix}. \tag{1.87}$$

When the mass matrix is diagonalized, the two eigenvalues result:

$$m_{\text{light}} \simeq \frac{\left(M_{\text{D}}^{\nu}\right)^2}{M_{\text{M}}^{\nu}} \tag{1.88}$$

$$m_{\text{heavy}} \simeq M_{\text{M}}^{\nu} \tag{1.89}$$

Since the right-handed neutrino fields are singlets under the Standard Model gauge symmetries, their Majorana mass is not linked to the electroweak symmetry breaking scale and can take any value, in particular $M_{\text{M}}^{\nu} >> M_{\text{D}}^{\nu}$. Hence, even if the left-handed neutrinos fields have a Yukawa coupling which is similar to that of the charged fermions; two very different neutrinos would arise: a light one (mostly left-handed, the one we are familiar with) with mass $m_{\text{light}} << M_{\text{D}}^{f}$, and a heavy one (mostly right-handed, and thereby sterile) with mass m_{heavy}. These neutrinos are Majorana fermions, that is, they are their own antiparticle. This mechanism is known as the seesaw mechanism [84, 85, 86, 87, 88]: the heavier the heavy neutrinos, the lighter the light ones.

A Majorana mass term violates the total lepton number symmetry of the Standard Model by two units. This allows to test experimentally the Majorana neutrino hypothesis, since if such term is present, neutrinoless double β decay ($\beta\beta0\nu$) should occur:

$${}^{\text{A}}_{\text{Z}}\text{X} \rightarrow {}^{\text{A}}_{\text{Z}\pm2}\text{Y} + 2\text{e}^{\mp}. \tag{1.90}$$

[6] No renormalizable Majorana term can be built with ν_{L} in the Standard Model because it would require a weak isospin Higgs triplet with $Y = 2$ which is absent. However, if the Standard Model is regarded as a low-energy effective theory and renormalizability is given up, a Majorana mass term using the Standard Model Higgs doublet would be possible, resulting in a mass $M_{\text{M}}^{\nu} = \kappa v^2/\Lambda$, where κ is a dimensionless coupling constant and Λ is a heavy mass characteristic of the symmetry-breaking scale of the new physics.

The simplest realization of this process consists in the exchange of a virtual light Majorana neutrino between the two virtual $W^{\mp}$ bosons emitted by the decaying nucleons [7]. In this case, the half-life of the process for the nucleus X is written:

$$T^{0\nu}_{1/2}(X) = \left[G^{0\nu}_X \left| \mathcal{M}^{0\nu}_X \right|^2 \left(\frac{m_{\beta\beta}}{m_e} \right)^2 \right]^{-1}, \tag{1.91}$$

where $G^{0\nu}_X$ and $\mathcal{M}^{0\nu}_X$ are a phase space factor and a nuclear matrix element which depend on the nucleus, respectively. While $G^{0\nu}_X$ can be calculated precisely, the calculation of $\mathcal{M}^{0\nu}_X$ requires to rely on a nuclear model, introducing considerable theoretical uncertainty. $m_{\beta\beta}$ is the effective Majorana mass, which is the coherent sum of the masses of the neutrino mass eigenstates, m_j:

$$\begin{aligned} m_{\beta\beta} &= \left| \sum_{j=1}^{3} U^2_{ej} m_j \right| \\ &= \left| c^2_{12} c^2_{13} m_1 + \mathrm{e}^{2\mathrm{i}\alpha_1} s^2_{12} c^2_{13} m_2 + \mathrm{e}^{2\mathrm{i}(\alpha_2 - \delta)} s^2_{13} m_3 \right|, \end{aligned} \tag{1.92}$$

where the U_{ej} matrix elements of the electron-neutrino row of the PMNS matrix including the diagonal matrix with the Majorana phases (equations 1.5 and 1.6) have been written explicitly in the second line.

Therefore, measuring the half-life of the $\beta\beta0\nu$ process gives information about the neutrino mass. However, the connection to the individual masses m_i is not direct, since the presence of complex phases in equation 1.92 may result in cancellations.

Equation 1.92 can be rewritten as a function of the unknown lightest neutrino mass and the measured squared-mass differences. In the normal hierarchy, m_1 is the lightest mass, $m_2 \simeq \sqrt{\Delta m^2_{21}}$ and $m_3 \simeq \sqrt{\Delta m^2_{\mathrm{atm}}}$, so $m_{\beta\beta}$ can be approximated by

$$m_{\beta\beta} \simeq \left| \left(c^2_{12} m_1 + \mathrm{e}^{2\mathrm{i}\alpha_1} s^2_{12} \sqrt{\Delta m^2_{21}} \right) c^2_{13} + \mathrm{e}^{2\mathrm{i}(\alpha_2 - \delta)} s^2_{13} \sqrt{\Delta m^2_{\mathrm{atm}}} \right|. \tag{1.93}$$

In the inverted hierarchy, m_3 is the lightest mass and $m_2 \simeq m_1 \simeq \sqrt{\Delta m^2_{\mathrm{atm}}} >> s^2_{13} m_3$, so $m_{\beta\beta}$ can be approximated by

$$m_{\beta\beta} \simeq \sqrt{\Delta m^2_{\mathrm{atm}}} c^2_{13} \left| c^2_{12} + \mathrm{e}^{2\mathrm{i}\alpha_1} s^2_{12} \right|. \tag{1.94}$$

If the neutrino masses do not exhibit a strong hierarchy, $m_1 \simeq m_2 \simeq m_3 \equiv m_\nu >> \sqrt{\Delta m^2_{\mathrm{atm}}}$, a situation known as *quasi-degenerated*, then

$$m_{\beta\beta} \simeq m_\nu \left| \left(c^2_{12} + \mathrm{e}^{2\mathrm{i}\alpha_1} s^2_{12} \right) c^2_{13} + \mathrm{e}^{2\mathrm{i}(\alpha_2 - \delta)} s^2_{13} \right|. \tag{1.95}$$

These three possibilities are depicted in figure 1.14. It is clear that a positive observation of $\beta\beta0\nu$ corresponding to a $m_{\beta\beta} < 0.01\,\mathrm{eV}$ would reveal that the neutrino mass hierarchy is normal. Conversely, the light Majorana neutrino exchange can be ruled out if oscillation experiments determine that the hierarchy is inverted and no $\beta\beta0\nu$ is observed for $m_{\beta\beta} > 0.01\,\mathrm{eV}$.

[7] Neutrinoless double β decay can proceed through the exchange of other virtual particles. However, regardless of the actual mechanism, the observation of the process implies a Majorana mass term for the neutrinos [89].

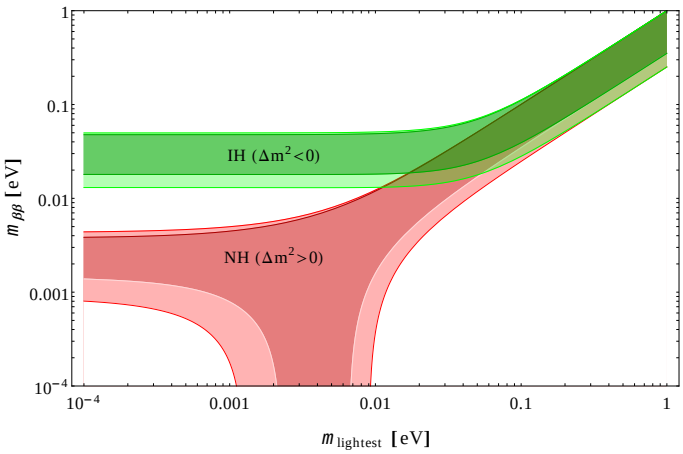


Figure 1.14: The effective Majorana mass in neutrinoless double β decay ($m_{\beta\beta}$) as a function of the lightest neutrino mass: m_1 for the normal hierarchy (NH); or m_3 for the inverted hierarchy (IH). The quasi-degenerate (QD) case corresponds to the upper right region. The width of the bands corresponds to a 3σ uncertainty on the oscillation parameters and the unknown CP-violating phases. From [90].

From the experimental point of view, the evidence of $\beta\beta0\nu$ would be a peak in the electron spectrum at the energy $T_{\max} = Q_{\beta\beta}$, where $Q_{\beta\beta} = M_i - M_f - 2m_e$, with M_i and M_f the masses of the initial and final nuclei, and m_e the mass of the electron. Since the recoil energy of the final nucleus is negligible, this decay is essentially a two-body reaction in which the two electrons carry away all the energy available. Being a weak interaction of second order, if existing, this reaction will be eclipsed by the single β decay, which is easier to materialize. Hence, the candidate nucleus to be used should have the single β decay either energetically forbidden or heavily suppressed because of a large change of spin. Even in this case, the $\beta\beta0\nu$ process is disfavored, since the Standard-Model-allowed double β decay ($\beta\beta2\nu$)

$$ {}^{\mathrm{A}}_{\mathrm{Z}}\mathrm{X} \rightarrow {}^{\mathrm{A}}_{\mathrm{Z+2}}\mathrm{Y} + 2\mathrm{e}^- + 2\overline{\nu}_\mathrm{e}, \tag{1.96} $$

already observed in some nuclei, will occur preferentially. This is a four-body reaction, resulting in a continuous spectrum up to $Q_{\beta\beta}$. Consequently, the observation of $\beta\beta0\nu$ requires a large number of the chosen nuclei, in a low-background and high-efficiency detector, with good energy resolution to separate the $\beta\beta0\nu$ peak from the $\beta\beta2\nu$ continuum (and the remaining background).

With the exception of a claim [91] of $\beta\beta0\nu$ in the $\beta\beta2\nu$-decaying isotope ^{76}Ge by part of the Heidelberg-Moscow collaboration, corresponding to $m_{\beta\beta} = 0.32 \pm 0.03\,\mathrm{eV}^2$; the $\beta\beta0\nu$ decay has not been observed yet and only upper limits have been set. The GERDA experiment [92], using an array of high purity germanium detectors enriched with ^{76}Ge has set a lower limit on the $\beta\beta0\nu$ half-life of $T^{0\nu}_{1/2}(^{76}\mathrm{Ge}) > 2.1 \cdot 10^{25}$ y (90% C.L.), which implies

$$ m_{\beta\beta} < 0.2 - 0.4\,\mathrm{eV}\ (90\%\ \mathrm{C.L.}) \tag{1.97} $$

depending on the nuclear matrix element calculation.

The EXO-200 experiment using a time projection chamber filled with liquid xenon enriched with the $\beta\beta2\nu$-decaying ^{136}Xe has set a lower limit on $T^{0\nu}_{1/2}(^{136}\text{Xe}) > 1.6 \cdot 10^{25}$ y (90% C.L.) [93] [8]. Another ^{136}Xe-based experiment, KamLAND-Zen, in which the KamLAND detector introduced in section 1.2.1 is used to look for the $\beta\beta0\nu$, has established the lower limit $T^{0\nu}_{1/2}(^{136}\text{Xe}) > 1.9 \cdot 10^{25}$ y (90% C.L.) [95]. When the EXO-200 and KamLAND-Zen limits are combined, the lower limit increases to $T^{0\nu}_{1/2}(^{136}\text{Xe}) > 3.4 \cdot 10^{25}$ y (90% C.L.), which corresponds to

$$m_{\beta\beta} < 0.12 - 0.25\,\text{eV (90\% C.L.)}. \tag{1.98}$$

More upper limits on $m_{\beta\beta}$ for other $\beta\beta2\nu$-decaying nuclei exist, but they are less stringent.

The current generation of $\beta\beta0\nu$ experiments has the potential to reach $m_{\beta\beta} \sim 0.1\,\text{eV}$. However, in order to explore the inverted hierarchy regime (cf. figure 1.14), larger experiments will be needed in the future.

1.3.3 Cosmology bounds

As the rest of the elementary particles, neutrinos were created in the first stages of the Universe after the Big Bang. Therefore, the evolution of the Universe is sensitive to the neutrino properties, which leave an imprint on the cosmic microwave background (CMB), the large-scale structure of the Universe or the element abundances from the primordial nucleosynthesis. A detailed discussion of this measurements is out of the scope of this thesis, so only the conclusions regarding the neutrinos are highlighted. The cosmological data is very well described by a minimal model (the Λ Cold Dark Matter model, abbreviated ΛCDM) which assumes 3 active neutrinos with normal hierarchy and a total mass $\sum m_i = 0.06\,\text{eV}$, where m_i are the masses of the neutrino mass eigenstates, with $m_1 \approx m_2 \approx 0$ and $m_3 \approx 0.06\,\text{eV}$.

However, the cosmological data can be used to test extensions of the model. From the latest CMB maps from the Planck satellite [96] and references therein, assuming 3 degenerate massive neutrinos, bounds ranging from $\sum m_i < 0.17 - 0.72\,\text{eV}$ (95% C.L.) are obtained depending on the data considered, with a conservative $\sum_i m_i < 0.23\,\text{eV}$ (95% C.L.). This result disfavors the neutrinoless double β decay in the quasi-degenerate region with $m_{\text{lightest}} \gtrsim 0.2\,\text{eV}$ in figure 1.14; and is comparable to the expected limit achievable by KATRIN on m_β. Nevertheless, the cosmology bounds are model dependent; hence, the independent information brought by a direct experiment is always valuable. In any case, the data supports the idea of light neutrinos, which are too light to be the primary component of the dark matter.

The cosmological data is also sensitive to the effective number of relativistic degrees of freedom beyond the photons, N_{eff}, during the radiation-dominated era. In the ΛCDM model, $N_{\text{eff}} = 3.046$ because of the 3 neutrinos of the Standard Model (and the corrections due to their partial decoupling during the electron-positron annihilation phase) [96, 97]. The fit to the data results in 3.04 ± 0.18, in good agreement with the prediction, and disfavors any additional contribution from light particles (e.g. light sterile neutrinos). If both N_{eff} and

[8] The latest analysis with increased statistics sets the lower limit on $T^{0\nu}_{1/2}(^{136}\text{Xe}) > 1.1 \cdot 10^{25}$ y (90% C.L.), consistent with the statistical fluctuation in the data [94].

$\sum_i m_i$ are fitted simultaneously, consistent results are obtained: $N_{\text{eff}} = 3.2 \pm 0.5$ and $\sum_i m_i < 0.32\,\text{eV}$ (both at 95% C.L.).

1.4 Open questions in neutrino oscillations

In addition to the underway experimental effort to measure the mass of the neutrinos and determine their Dirac or Majorana nature, the field of neutrino oscillations is still active and poses questions that remain unanswered.

Table 1.5 and figure 1.15 summarize the current knowledge obtained from a global fit to the data from the experiments presented in section 1.2.

The two most prominent challenges are determining if there is CP violation in the neutrino sector, and if positive, what is the value of the CP-violating phase δ (cf. right middle panel in figure 1.15); and resolve the neutrino mass hierarchy, that is, what is the lightest neutrino mass eigenstate: ν_1 or ν_3. Both questions can be answered measuring precisely the channel $\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_e$ with accelerator-based experiments, since terms sensitive to the sign of Δm^2_{32}, due to the matter effects, and the value of δ appear in the transition probability (equation 1.68). As it was discussed in section 1.2.3, the recently measured θ_{13} enters also the transition probability. Therefore, the precision in the knowledge of the θ_{13} mixing angle is crucial to address these questions; and reactor antineutrino experiments like Double Chooz provide a degeneracy-free measurement to be used as input.

If the baseline is long enough, the matter effects result in an enhancement of the conversion (see the resonance in equation 1.23) of neutrinos if the hierarchy is normal ($\Delta m^2_{32} > 0$), or for antineutrinos if the hierarchy is inverted $\Delta m^2_{32} < 0$. This a consequence of Earth not being CP symmetric (it is made of matter, not antimatter), so matter effects induce a CP violation background that must be handled with care when measuring δ. Therefore, the determination of the neutrino mass hierarchy affects the sensitivity of the long-baseline oscillation experiments searching for CP violation because of the degeneracy between δ and $\text{sign}(\Delta m^2_{32})$.

The new NOνA experiment [100], with a baseline of $\sim 810\,\text{km}$ might be able to shed some light on both δ and the hierarchy. The current T2K experiment has not a baseline long enough to contribute to the mass hierarchy research, but it has already probed some of the possible δ values. An updgrade of the Super-Kamiokande detector to a 25-times bigger detector (Hyper-Kamiokande [101]) would increase its potential to measure δ. However, in order to study the mass hierarchy precisely, an experiment with a longer baseline (to increase matter effects) will be needed. The DUNE experiment (formerly LBNE [102]), currently in the design phase, with a long baseline ($\sim 1300\,\text{km}$) and an intense beam with a broad energy spectrum (peaked at $\sim 2.5\,\text{GeV}$) might answer both δ and $\text{sign}(\Delta m^2_{32})$ questions.

The mass hierarchy can also be resolved using non-accelerator experiments. The disappearance of atmospheric muon neutrinos is also sensitive to the matter effects, and huge baselines are available from the study of the neutrinos crossing the Earth. Future experiments using large masses of water, either frozen as the Antarctic ice (IceCUBE-PINGU [103]) or liquid as in the Mediterranean sea (KM3Net-ORCA [104]) or confined in a volume (Hyper-Kamiokande), as water Cherenkov detectors will be able to compete in the search. Other large detectors

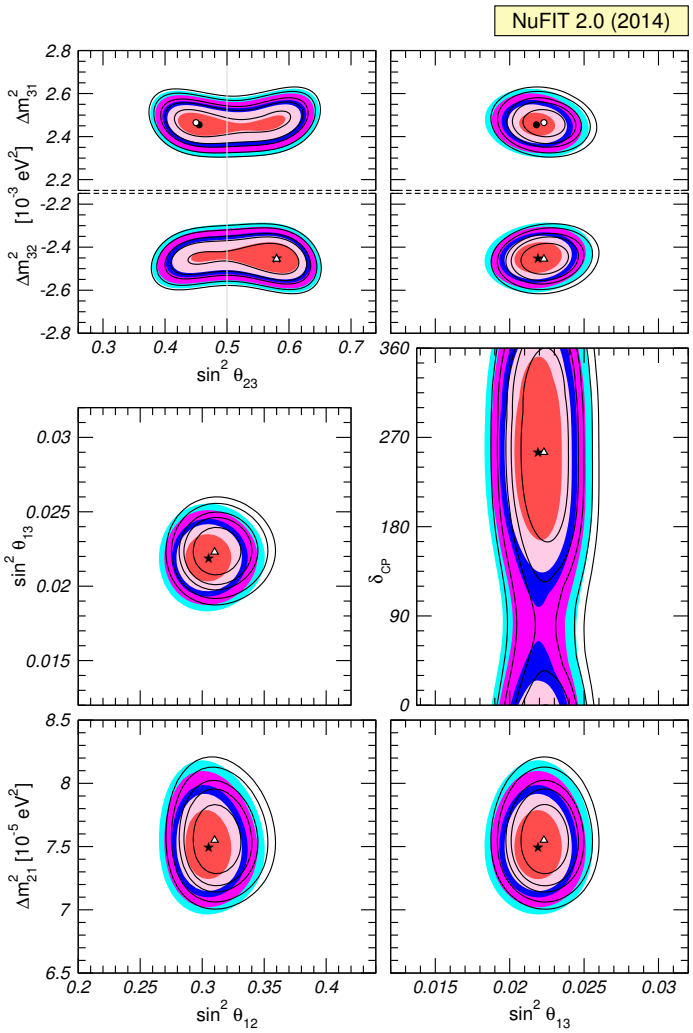


Figure 1.15: Two-dimensional allowed regions from `NuFit 2.0` to global data in 2014. In each panel, a minimization with respect to the parameters not displayed is performed. The different contours correspond to 1σ, 90%, 2σ, 99% and 3σ C.L. (for 2 degrees of freedom). Shaded regions show the best fit results leaving the normalization of the reactor fluxes free and including data from short baseline reactor experiments ($L < 100\,\mathrm{m}$). Empty contours use the reactor flux prediction from [98] and do not use the short baseline reactor data. The Double Chooz results covered in this thesis are not included. From [99].

	NH ($\Delta\chi^2 = 0.97$)		IH (best fit)		Any hierarchy
	Best fit $\pm 1\sigma$	3σ range	Best fit $\pm 1\sigma$	3σ range	3σ range
$\sin^2\theta_{12}$	$0.304^{+0.013}_{-0.012}$	$0.270 \to 0.344$	$0.304^{+0.013}_{-0.012}$	$0.270 \to 0.344$	$0.270 \to 0.344$
$\theta_{12}/^\circ$	$33.48^{+0.78}_{-0.75}$	$31.29 \to 35.91$	$33.48^{+0.78}_{-0.75}$	$31.29 \to 35.91$	$31.29 \to 35.91$
$\sin^2\theta_{23}$	$0.452^{+0.052}_{-0.028}$	$0.382 \to 0.643$	$0.579^{+0.025}_{-0.037}$	$0.389 \to 0.644$	$0.385 \to 0.644$
$\theta_{23}/^\circ$	$42.3^{+3.0}_{-1.6}$	$38.2 \to 53.3$	$49.5^{+1.5}_{-2.2}$	$38.6 \to 53.3$	$38.3 \to 53.3$
$\sin^2\theta_{13}$	$0.0218^{+0.0010}_{-0.0010}$	$0.0186 \to 0.0250$	$0.0219^{+0.0011}_{-0.0010}$	$0.0188 \to 0.0251$	$0.0188 \to 0.0251$
$\theta_{13}/^\circ$	$8.50^{+0.20}_{-0.21}$	$7.85 \to 9.10$	$8.51^{+0.20}_{-0.21}$	$7.87 \to 9.11$	$7.87 \to 9.11$
$\delta/^\circ$	306^{+39}_{-70}	$0 \to 360$	254^{+63}_{-62}	$0 \to 360$	$0 \to 360$
$\frac{\Delta m^2_{21}}{10^{-5}\,\mathrm{eV}^2}$	$7.50^{+0.19}_{-0.17}$	$7.02 \to 8.09$	$7.50^{+0.19}_{-0.17}$	$7.02 \to 8.09$	$7.02 \to 8.09$
$\frac{\Delta m^2_{3k}}{10^{-3}\,\mathrm{eV}^2}$	$+2.457^{+0.047}_{-0.047}$	$+2.317 \to +2.607$	$-2.449^{+0.048}_{-0.047}$	$-2.590 \to -2.307$	$\begin{bmatrix} +2.325 \to +2.599 \\ -2.590 \to -2.307 \end{bmatrix}$

Table 1.5: Three-flavor neutrino oscillation parameters from `NuFit 2.0` to global data in 2014 (from [99] and references therein). The Double Chooz results covered in this thesis are not included. To account for the reactor anomaly (see text), the normalization of the reactor fluxes is left free in the fit and data from short baseline reactor experiments ($L < 100\,\mathrm{m}$) are included. The numbers in the first and second column are obtained minimizing with respect to the normal (NH) or inverted (IH) mass hierarchy local minimum. In the third column the minimization is also done with respect to the hierarchy. In the last row, Δm^2_{3k} corresponds to the largest absolute squared-mass difference, i.e. $\Delta m^2_{31} > 0$ for NH and $\Delta m^2_{32} < 0$ for IH.

such as magnetized iron calorimeters (ICAL [105]), or the liquid argon time projection chamber acting as the far detector in DUNE might profit also from the atmospheric neutrinos channel. The disappearance of electron antineutrinos from reactors at the first minimum due to $(\Delta m^2_{21}, \theta_{12})$, equal to a distance of $\sim 50\,\text{km}$, is also sensitive to the mass hierarchy from the interplay between the terms containing Δm^2_{32} and Δm^2_{31} in equation 1.53. In order to observe this effect, a detailed knowledge of the reactor antineutrino spectrum is necessary and an excellent energy resolution is required. The JUNO experiment [106] will pursue this measurement.

Currently, it is known that θ_{23} is close to $\pi/4$ (or 45°, see table 1.5), but it is not known whether it lies in the first octant ($\theta_{23} < \pi/4$), in the second ($\theta_{23} > \pi/4$) or it is just the maximal mixing (between the ν_2 and ν_3 mass eigenstates) value $\theta_{23} = \pi/4$ (cf. upper left panel in figure 1.15). The precise study of the channels $\overset{(-)}{\nu}_\mu \rightarrow \overset{(-)}{\nu}_e$ and $\overset{(-)}{\nu}_\mu \rightarrow \overset{(-)}{\nu}_\mu$ with the aforementioned detectors has the potential to discriminate between the different options. The determination of the actual value of θ_{23} is important, not only because it impacts the sensitivity of the future experiments that will measure the mass hierarchy and the CP violation, but also because it can point to flavor symmetries to explain the structure of the PMNS matrix,

$$|U_{\text{PMNS}}| \sim \begin{pmatrix} 0.8 & 0.5 & 0.2 \\ 0.5 & 0.6 & 0.6 \\ 0.2 & 0.6 & 0.8 \end{pmatrix}, \tag{1.99}$$

and why it is so different from the almost-diagonal CKM matrix in the quark sector,

$$|V_{\text{CKM}}| \approx \begin{pmatrix} 1 & 0.2 & 0.004 \\ 0.2 & 1 & 0.04 \\ 0.009 & 0.04 & 1 \end{pmatrix}. \tag{1.100}$$

As it has been described in this chapter, the three-flavor neutrino framework describes successfully almost all the experimental data available. However, some experiments have obtained results which cannot be accommodated in this framework.

A long-standing anomaly is that initiated by the excess of $\overline{\nu}_e$ events in the LSND experiment. The LSND experiment ran between 1993 and 1998 at Los Alamos National Laboratory (New Mexico, USA). LSND used a $\overline{\nu}_\mu$ beam generated from decay-at-rest pions directed towards a liquid scintillator detector 30 m downstream, and looked for the appearance of $\overline{\nu}_e$ via inverse-β decay on carbon nuclei. A final excess of $87.9 \pm 22.4 \pm 6$ of $\overline{\nu}_e$ events above backgrounds in the positron energy range between 20 MeV and 60 MeV was reported [107]. If interpreted as a result of oscillations, it points to a squared-mass difference in the range $\Delta m^2 = 0.2 - 10\,\text{eV}^2$. This squared-mass difference is much larger than Δm^2_{21} or Δm^2_{32}, suggesting the existence of a new mass eigenstate heavier than the known ones. This mass eigenstate would require a new state in the flavor basis, which must be sterile in order to escape the constrain imposed by the Z boson width explained in section 1.1. However, the KARMEN experiment, with similar beam and detector distance did not observe any excess [108], although it could not exclude completely the LSND result (see figure 1.16).

The LSND anomaly was further investigated by the MiniBooNE experiment at the Fermi National Accelerator Laboratory using both $\overline{\nu}_\mu$ and ν_μ beams

created from the decay in flight of pions. The beams had a wide energy distribution, peaked at 400 MeV for $\overline{\nu}_\mu$ and 600 MeV for ν_μ, and a low contamination of electron-flavor (anti)neutrinos ($\leq 0.5\%$). The detector was placed at a distance of ~ 540 m from the production point, so it could explore a similar L/E region as LSND. The detector was a mineral-oil Cherenkov detector, in which the flavor of the neutrino was revealed from the Cherenkov light emitted by the charged lepton produced upon charged-current interaction. MiniBooNE observed excesses in both antineutrino and neutrino modes in the neutrino energy interval $(200, 1250)$ MeV [109]. The antineutrino excess, 78.4 ± 28.5 events (2.8σ), when interpreted as oscillation-induced, gives an allowed parameter region which overlaps partially with the LSND one (cf. figure 1.16). On the other hand, the neutrino excess, 162.0 ± 47.8 events (3.4σ), is concentrated at low energy, where is barely compatible with an oscillation interpretation.

It is important to remark that the MiniBooNE detector cannot distinguish between electrons and γ-rays. For that reason, MicroBooNE [110], a new experiment in the same beamline as MiniBooNE has been built, in which a liquid argon time projection chamber (LArTPC) will be used as detector at ~ 470 m from the production point. The excellent tracking and calorimetric capabilities of the LArTPC allow to tell the electrons from γ-rays, addressing the nature of the excess. Furthermore, adding SBND (formerly LAr1-ND [111]), a near detector at ~ 100 m from the beam production point, will reduce the systematic uncertainties affecting the oscillation search, enhancing the sensitivity in the electron-flavor appearance and muon-flavor disappearance charged-current detection channels, and opening the possibility to look for disappearance of active neutrinos into sterile neutrinos using neutral-current interactions. Finally, the LSND allowed region will be completely explored with the addition of the large ICARUS T600 LArTPC detector at ~ 600 m, which is moving from Gran Sasso to the Fermi National Accelerator Laboratory with an overhauling stop at CERN [112].

Another anomaly, the *gallium anomaly* [113], has arisen from the reanalysis of the calibration data in the GALLEX/GNO and SAGE solar neutrino radiochemical detectors (introduced in section 1.2.1). These detectors used ^{51}Cr and ^{37}Ar (only SAGE) sources, which decay via electron capture emitting monochromatic ν_e with energies below 1 MeV, to measure their efficiency to detect solar neutrinos using the reaction in equation 1.29. The combination of four calibration runs, two at each detector, gives an observed-to-expected ratio of

$$R_{\mathrm{Ga}} = 0.76^{+0.09}_{-0.08}, \tag{1.101}$$

which is 3σ away from 1. If the result is interpreted as a consequence of ν_e disappearance due to oscillation into an undetectable flavor in a two-neutrino approximation, it is found $\sin^2(2\theta) \gtrsim 0.07$ and $\Delta m^2 \gtrsim 0.35\,\mathrm{eV}^2$ at 99% C.L.

Moreover, the reevaluation of the reactor antineutrino fluxes to be used by the current θ_{13} experiments resulted in a normalization shift of +3% [114, 98] (the Double Chooz prediction is explained in section 2.4.1). When the data from past reactor experiments performed at short baselines (< 100 m) is reanalyzed with the new flux prediction, an overall deficit appears in the ratio of the observed to predicted $\overline{\nu}_e$ events:

$$R_{\mathrm{SBL}} = 0.943 \pm 0.023. \tag{1.102}$$

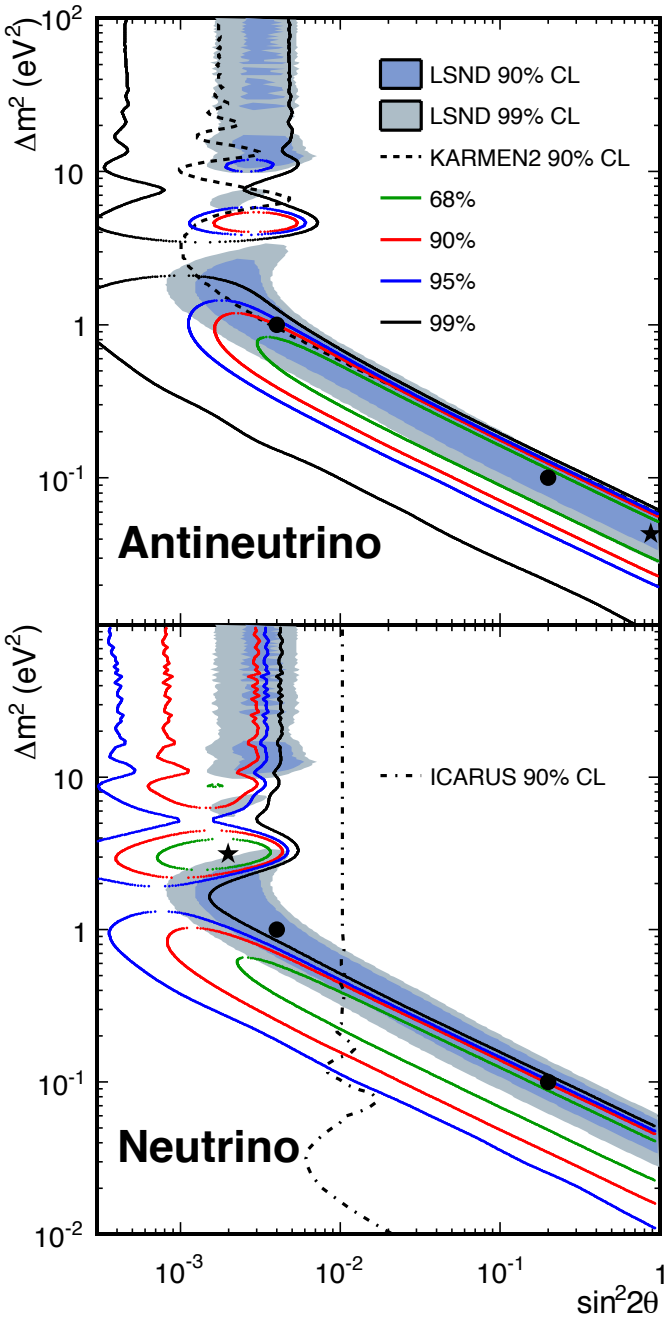


Figure 1.16: Allowed regions (empty contours) in the $\sin^2(2\theta) - \Delta m^2$ plane from a two-flavor oscillation analysis of the MiniBooNE antineutrino (top) and neutrino (bottom) data. The black star shows the best-fit value. The allowed regions from LSND are shaded. The upper limits from KARMEN on antineutrinos and ICARUS on neutrinos are also shown. From [109] and references therein.

This has been called the *reactor antineutrino anomaly* [115]. The significance of this anomaly depends critically on the precision of the reactor flux knowledge. The interpretation of the deficit as a consequence of neutrino oscillation due to a fourth nonstandard neutrino (the so-called 3+1 scheme), gives a $\sin^2(2\theta_{\rm new})$ in the range $0.02 - 0.21$ and $|\Delta m^2_{\rm new}| > 0.23\,{\rm eV}^2$ (95.3% C.L.).

This anomaly has prompted a new series of experiments to search for an oscillation signal at very short baselines from nuclear reactors. These are challenging experiments since the reactor-induced background at such short distances ($\sim 10\,{\rm m}$) is high. Another possibility is to use an intense radioactive source such as those used to calibrate the gallium experiments to measure the spatial dependence of the number of neutrino interactions, looking for a modulation characteristic from oscillations. Suitable detectors for this approach are the Borexino [116] and KamLAND detectors [117].

Finally, it must be noted that additional sterile neutrinos appear naturally if both Dirac and Majorana mass terms exist, as it was pointed out at the beginning of section 1.3.2. If neutrinos are proved to be Majorana from the neutrinoless double β decay experiments and CP violation is measured in the neutrino sector, the predominance of matter over antimatter could be originated because of the neutrinos; a possible realization of *leptogenesis*. In this model [118], the heavy right-handed Majorana neutrinos created at the beginning of the Universe could undergo CP-violating decays into Higgs bosons and lighter leptons, allowing for a lepton number asymmetry. As the Universe expanded, the process would go out of equilibrium, leaving a net excess of leptons. The nonperturbative sphaleron process of the Standard Model can convert this excess of leptons into baryons, satisfying all the Sakharov conditions [119] to generate dynamically a matter-antimatter asymmetry. The large energy scales of these processes preclude a direct study, but neutrinos would be the heirs of this new physics.

Chapter 2

The Double Chooz experiment

Double Chooz [1] is a two-detector experiment to measure the neutrino mixing angle θ_{13} from the oscillation-induced deficit of electron antineutrinos in the flux emitted by the nuclear reactors of the Chooz power plant. The experiment concept and site are described in section 2.1. The detection of the electron antineutrinos is discussed in section 2.2 and the design of the detector is explained in section 2.3. The Monte Carlo simulation which provides the expected number of electron antineutrinos in the detector is described in section 2.4.

2.1 Experiment concept

The idea of a two-detector reactor experiment to search for the disappearance of $\overline{\nu}_e$ due to the oscillation driven by θ_{13} (see equation 1.54) was first proposed in [120]. The use of two detectors, one close to the reactor at a distance of $\mathcal{O}(100\,\mathrm{m})$ (the Near Detector) and one distant at $\mathcal{O}(1\,\mathrm{km})$ (the Far Detector), to measure simultaneously the $\overline{\nu}_e$ flux emitted by the reactor would allow to cancel the uncertainty in the reactor flux in the relative comparison of both measurements [2]. In addition, by building the two detectors as identical as possible, the systematic uncertainty in the $\overline{\nu}_e$ detection could be canceled as well, except for the uncorrelated part accounting for the remaining differences between the detectors, which could be minimized with thorough calibration and monitoring. Since these two sources (reactor flux and $\overline{\nu}_e$ detection) were the dominant systematic uncertainties affecting the best experimental limit on θ_{13} at the time, set by the CHOOZ experiment [63] (figure 1.11), a two-detector experiment was guaranteed to make a significant improvement in the knowledge of this mixing angle.

Double Chooz [66] was originally proposed around the year 2004 as a fast and moderate-cost experiment to measure $\sin^2(2\theta_{13}) \geq 0.05$ at 3σ or set a limit $\sin^2(2\theta_{13}) < 0.03$ at 90% C.L. after 3 years of data taking with two detectors,

[1] Chooz is pronounced as "sho".

[2] The magnitude of the cancellation of the uncertainty in the $\overline{\nu}_e$ flux when two or more reactor cores are present depends on the location of the detectors with respect to the reactors. This is further discussed at the end of the section.

assuming a $\Delta m^2_{31} = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$. This covered about 85% of the allowed region at the time of the proposal, and it was due to happen before the second generation experiments (Daya Bay and RENO), with more antineutrino luminosity owing to the use of more reactor cores and larger detectors, started their operation and potentially improved the result. Double Chooz relied on the advantage granted by the existence of the laboratory where the CHOOZ experiment had been located, which was available. Therefore, a phased construction was planned. The Far Detector would be installed first in the CHOOZ site and take data alone for 1.5 years, while the Near Detector and its laboratory were being built. This one-detector phase would allow to surpass the CHOOZ sensitivity in a few months. Then the two-detector phase of the experiment would begin, which would last for 3 years. However, the schedule was disrupted by several delays which affected both detectors. In the end, the Far Detector started data taking in 2011. This made the Double Chooz one-detector phase concurrent with the second generation experiments, which started their data taking directly in a multi-detector configuration. Nevertheless, Double Chooz managed to be the first reactor experiment in providing an indication of a non-null θ_{13} [69], opening the way for the Daya Bay and RENO experiments, which measured θ_{13} with higher precision in 2012 [70, 121]. Concerning the Near Detector, its commissioning began at the end of 2014, while this thesis was being written. For that reason, no Near Detector data will be used in this thesis.

The source of electron antineutrinos for the Double Chooz experiment are the two reactor cores of the Chooz nuclear power plant, located close to the Chooz village in the Ardennes department in the northeast of France. These two cores are N4 type pressurized water reactors (PWR) with a nominal thermal power of $4.25\,\mathrm{GW_{th}}$ each, which makes them two of the most powerful reactors worldwide. The $\overline{\nu}_e$ are produced in the β-decay of the fission products. More details about the $\overline{\nu}_e$ generation are given in the section 2.4.1 covering the reactor flux simulation.

In an experiment with only one reactor, the cancellation of the flux uncertainty is always total regardless where the two detectors are located. When there are two or more reactors, as is the case for Double Chooz, the total cancellation only occurs if the detectors receive the flux from the reactors in the same proportion (e.g. in an experiment with two reactors, named R1 and R2, all detectors receive x% of the flux from R1 and $(100-x)$% from R2.). This is known as the isoflux condition [122]. In Double Chooz, the location of the Far Detector was predetermined since it occupies the same pit that hosted the CHOOZ detector, at $997.839 \pm 0.015\,\mathrm{m}$ from the East reactor and $1114.656 \pm 0.015\,\mathrm{m}$ from the West reactor (see figure 2.1). The laboratory is excavated in the rock under a hill which provides 300 m.w.e. of overburden against the cosmic rays. This defines a line of isoflux positions for the Near Detector. Ultimately, the Near Detector was placed at $\sim 400\,\mathrm{m}$ of the reactors [3], slightly off the isoflux line in order to have the protection conferred by a small nearby hill which gives 120 m.w.e. of overburden.

Even though Double Chooz does not fulfill the isoflux condition exactly, it is close enough to greatly benefit from the suppression of the reactor flux systematic uncertainty, leaving a conservatively estimated 0.1% uncertainty as remnant [122]. This represents an advantage over the other reactor experiments, with

[3] Results from the geodetic survey have not been disclosed yet.

Figure 2.1: Aerial view of the Double Chooz experimental site. The locations of the two reactors and the two detectors are indicated on the photograph.

more complicated setups which do not meet the isoflux condition (see figure 2.2). In addition, because Double Chooz has only two reactors, it becomes a single reactor experiment, with full cancellation of the flux uncertainty, every time one of the reactors must be stopped (e.g. for refuelling). Moreover, although it is economically very unfavorable for the power plant company, there is the possibility of having the two reactors stopped at the same time, allowing a direct measurement of the backgrounds (see section 4.1.3.4). These circumstances are highly unlikely in Daya Bay or RENO, with six reactors each.

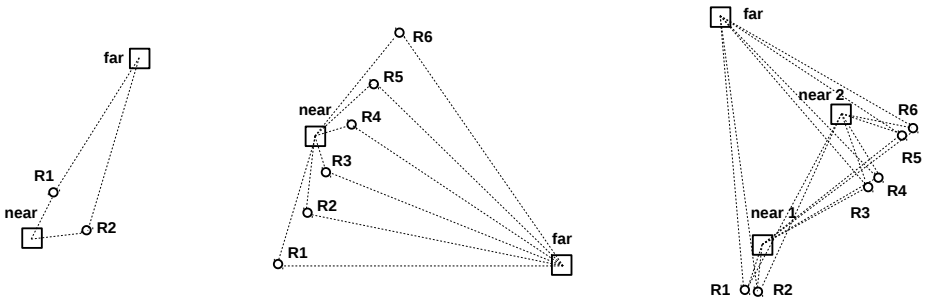


Figure 2.2: Top view layout of the experimental setups of the Double Chooz (left), RENO (middle) and Daya Bay (right) experiments. The circles depict the position of the reactor cores (labeled as R) and the squares depict the position of the near and far detector halls. From [122].

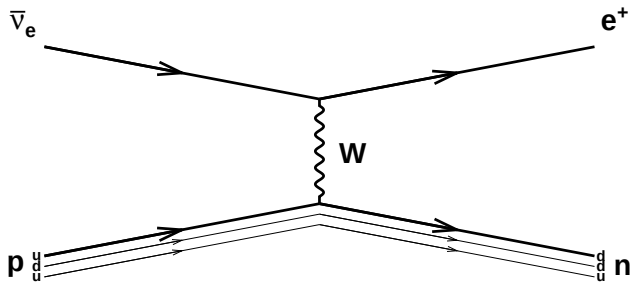


Figure 2.3: Feynman tree-level diagram for the inverse β-decay reaction, in which the electron antineutrino is transformed into a positron through the exchange of a virtual W boson with one of the up quarks in the proton, which is converted into a down quark. As a result, the proton is transformed into a neutron.

2.2 Electron antineutrino detection

The electron antineutrinos emitted by the reactor cores are detected through the inverse β-decay (IBD) reaction,

$$\overline{\nu}_e + p \rightarrow e^+ + n, \tag{2.1}$$

in which an electron antineutrino undergoes a weak interaction with a proton, becoming a positron and a neutron, respectively (see figure 2.3).

The $\overline{\nu}_e$ threshold energy for the IBD interaction in the laboratory frame (where the proton is at rest), neglecting the antineutrino mass, is

$$E^{\mathrm{LAB}}_{\overline{\nu}_e,\,\mathrm{th}} = \frac{(m_e + m_n)^2 - m_p^2}{2m_p} = 1.806\,\mathrm{MeV}, \tag{2.2}$$

so the experiment is not sensitive to the lowest energy region of the β spectra of the fission products (see figure 2.4). In addition, since the reactors produce antineutrinos with a kinetic energy of a few MeV, there is not enough energy to create the charged leptons when the antineutrino oscillate into the μ or τ flavors. Consequently, the unique evidence of neutrino oscillation is a deficit of $\overline{\nu}_e$ in the reactor flux, which makes Double Chooz a *disappearance* experiment.

At zeroth order in $1/M$, where M is the nucleon mass, and invoking the relation with the decay of the free neutron, the IBD total cross-section in the laboratory frame can be written as [123]

$$\sigma_{\mathrm{IBD}} = K(E_{\overline{\nu}_e} - \Delta)\sqrt{(E_{\overline{\nu}_e} - \Delta)^2 - m_e^2}(1 + \delta), \tag{2.3a}$$

with $\Delta = m_n - m_p$ and

$$K = \frac{2\pi^2}{m_e^5 f^{\mathrm{R}} \tau_n}, \tag{2.3b}$$

where τ_n is the measured neutron lifetime, which we take from the MAMBO-II experiment [124]; and $f^{\mathrm{R}} = 1.7152$ is the phase space factor of the free neutron

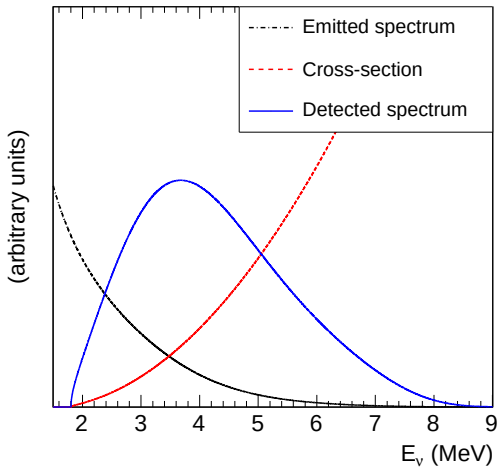


Figure 2.4: Polynomial parametrization of the ^{235}U antineutrino spectrum (black dash-dotted line), inverse β-decay cross-section (red dashed line) and the detectable antineutrino spectrum (blue solid line, oscillation not included) resulting from multiplying the first two. From [114].

decay calculated in [125]. We find $K = 9.61 \cdot 10^{-44}\,\mathrm{cm}^2\,\mathrm{MeV}^{-2}$. Moreover, the δ corrections to the cross-section (eq. 2.3a) due to neutron recoil, weak magnetism and outer radiative corrections of order α [126] are incorporated numerically. This cross-section is well-known, with an uncertainty driven by the neutron lifetime uncertainty (0.2%).

The IBD reaction is chosen because it has the highest cross-section in the energy range of the reactor antineutrinos and leaves a very distinctive imprint on the detector: firstly, the positron deposits its kinetic energy, T_e, in the detector and annihilates with an electron. This gives a *prompt* signal following the $\overline{\nu}_e$ interaction. The visible energy of this prompt signal is closely related to the $\overline{\nu}_e$ energy:

$$\begin{aligned} E_{\mathrm{vis}} = T_e + 2m_e &= E_{\overline{\nu}_e} - \Delta + m_e - T_n \\ &\simeq E_{\overline{\nu}_e} - 0.782\,\mathrm{MeV}, \end{aligned} \tag{2.4}$$

where the neutron kinetic energy, T_n, is small compared to $E_{\overline{\nu}_e}$. Secondly, after the neutron has thermalized (which takes a few microseconds), it is captured on a detector nucleus, releasing γ rays which constitute a *delayed* signal. The coincidence of these two signals grants the experiment an intrinsic background suppression mechanism.

Furthermore, the detector design (described in the next section 2.3) is optimized for the detection of IBD interactions. The use of organic liquid scintillator provides numerous free protons for the reaction. Part of this liquid scintillator is loaded with gadolinium, which has the highest thermal neutron capture cross-section among all stable nuclides (see table 2.1). This increases the neutron detection efficiency and shortens the mean capture time to $\sim 31.1\,\mu$s. Moreover, the radiative capture of the neutron on Gd,

$$n + {}^{\mathrm{A}}\mathrm{Gd} \rightarrow {}^{\mathrm{A+1}}\mathrm{Gd}^* \rightarrow {}^{\mathrm{A+1}}\mathrm{Gd} + \sum_{\mathrm{i}} \gamma_{\mathrm{i}}, \tag{2.5}$$

Gd isotope	$\sum_i E_{\gamma_i}$ (keV)	Abundance (%)	Cross-section (barns)	Relative intensity
152	6247	0.20	735	$3.0 \cdot 10^{-5}$
154	6438	2.18	85	$3.8 \cdot 10^{-5}$
155	8536	14.80	60900	0.1848
156	6360	20.47	1.50	$6.3 \cdot 10^{-6}$
157	7937	15.65	254000	0.8151
158	5942	24.84	2.20	$1.1 \cdot 10^{-5}$
160	5635	21.86	0.77	$3.5 \cdot 10^{-6}$

Table 2.1: Naturally occurring gadolinium isotopes (except the radioisotope ^{152}Gd with a half-life of $1.08 \cdot 10^{14}$ years, the rest are observationally stable). Starting from the second column, the columns show the total energy of the γ rays emitted upon a radiative neutron capture, the natural abundance, the thermal neutron capture cross-section, and the relative intensity resulting from multiplying the abundance by the cross-section. From [63].

releases a total energy of $\approx 8\,\mathrm{MeV}$, which is far above the natural radioactivity energies. This differentiates the IBD signature from the random coincidences of that background. Although not optimized for it, the experiment is also capable of using the neutrons which are captured on hydrogen forming deuterium,

$$n + {}^1\mathrm{H} \rightarrow {}^2\mathrm{H} + \gamma(2.224\,\mathrm{MeV}). \tag{2.6}$$

This is a process with a lower cross section (0.33 b), resulting in a longer mean capture time $\sim 200\,\mu$s, but it profits from the substantial presence of H nuclei in the hydrocarbon compounds.

By requiring the coincidence of the two characteristic signals of the IBD reaction, most of the background events will be rejected at the selection level. Only those which are able to mimic such coincidence will remain in the sample. From those, the one which consists of *accidental* coincidences of two uncorrelated signals has already been mentioned. Since its dominant component is the natural radioactivity of the detector materials, it can be further decreased by building the detector as radiopure as possible. The other backgrounds are byproducts of the muon flux that bathes the experiment and which is generated in the decay of the charged pions produced by the cosmic rays interacting in the Earth's atmosphere. One of these backgrounds are the cosmogenic isotopes, that is, unstable isotopes which are produced by the passage of muons through the detector. Some of them are specially dangerous, since they have decay modes in which an electron is emitted along with a neutron, reproducing the final state of the IBD interaction, and are long-lived enough to make a veto time not feasible. Another of these backgrounds are the high-energy neutrons released when the muons break up the nuclei of the matter surrounding the detector. These *fast neutrons* are highly penetrating, and can reach the inside of the detector, where they can make proton recoils or be captured, causing sequential triggers which resemble those of the IBD reaction. Finally, the muons which stop in the detector can also look like IBD interactions when the Michel electron/positron from the decay is mistaken as the delayed signal. In order to reduce the impact of these backgrounds, and added to the underground location of both detec-

tors, the detector includes several active subdetectors which are described in the next section. These backgrounds and how the active subdetectors are used to reject most of their contribution are explained in detail in sections 4.1.3 (for the Gd-based $\overline{\nu}_e$ selection) and 4.2.3 (for the H-based $\overline{\nu}_e$ selection).

2.3 Detector design

The Double Chooz detector consists of a cylindrical main detector, enclosed by shielding and covered by an outer veto. The main detector is formed by four concentric cylindrical volumes (called Target, Gamma Catcher, Buffer and Inner Veto) which are nested in the fashion of a matryoshka doll (see figure 2.5). In order to avoid mechanical stress on the vessels, the liquids which fill each volume have their densities matched to 0.804 ± 0.001 g/cm^3 at 15 °C, which is close to the detector mean temperature.

The detector can be conceptually divided into two big subdetectors by considering the kind of events they are designed for: an Inner Detector for measuring the signal events, and an Outer Detector to act as a background veto.

2.3.1 Inner Detector

The Inner Detector is the part of the detector aimed to measure the signals produced by the IBD interaction of the $\overline{\nu}_e$. It is a liquid scintillator calorimeter which is divided into three concentrical cylindrical volumes: the Target, the Gamma Catcher and the Buffer. A central chimney, connected to the three volumes, was used to fill them. It is observed by 390 low background 10-inch photomultiplier tubes (PMT) (Hamamatsu R7081) installed on the inner walls of the Buffer.

Target

The innermost volume, the Target, is an 8 mm thick acrylic vessel (radius 1150 mm, height 2458 mm), transparent to UV to visible light, filled with 10.3 m^3 of liquid scintillator loaded with Gd to enhance the capture of the IBD neutrons. The scintillator solvent consists of o-PXE (ortho-phenylxylylethane) and n-dodecane, in a volume concentration of 20% and 80%, respectively. Two additional fluors are added, PPO (2,5-diphenyloxazole) as the primary fluor (7 g/l), and bis-MSB (1,4-bis-(2-methylstyryl)benzene) as second wavelength shifter (20 mg/l), in order to shift the scintillation light to a region in which the mixture is more transparent and the PMTs have the highest quantum efficiency.

The gadolinium metal does not dissolve well in the organic liquids which make up the scintillator. This can cause problems in the optical stability of the scintillator, as it happened in the CHOOZ experiment [63]. In order to achieve a higher solubility and a durable chemical stability, a metalorganic complex is used, Gd(thd)$_3$ (Gd(III)-tris-(2,2,6,6-tetramethyl-heptane-3,5-dionate)), a metal-β-diketone [127]. These complexes have a high vapor pressure, which makes possible to purify them by sublimation, thus reducing the radioactive contamination of K, U and Th. The Gd concentration (0.123% by weight, corresponding approximately to 1 g/l) is chosen as a compromise between a high

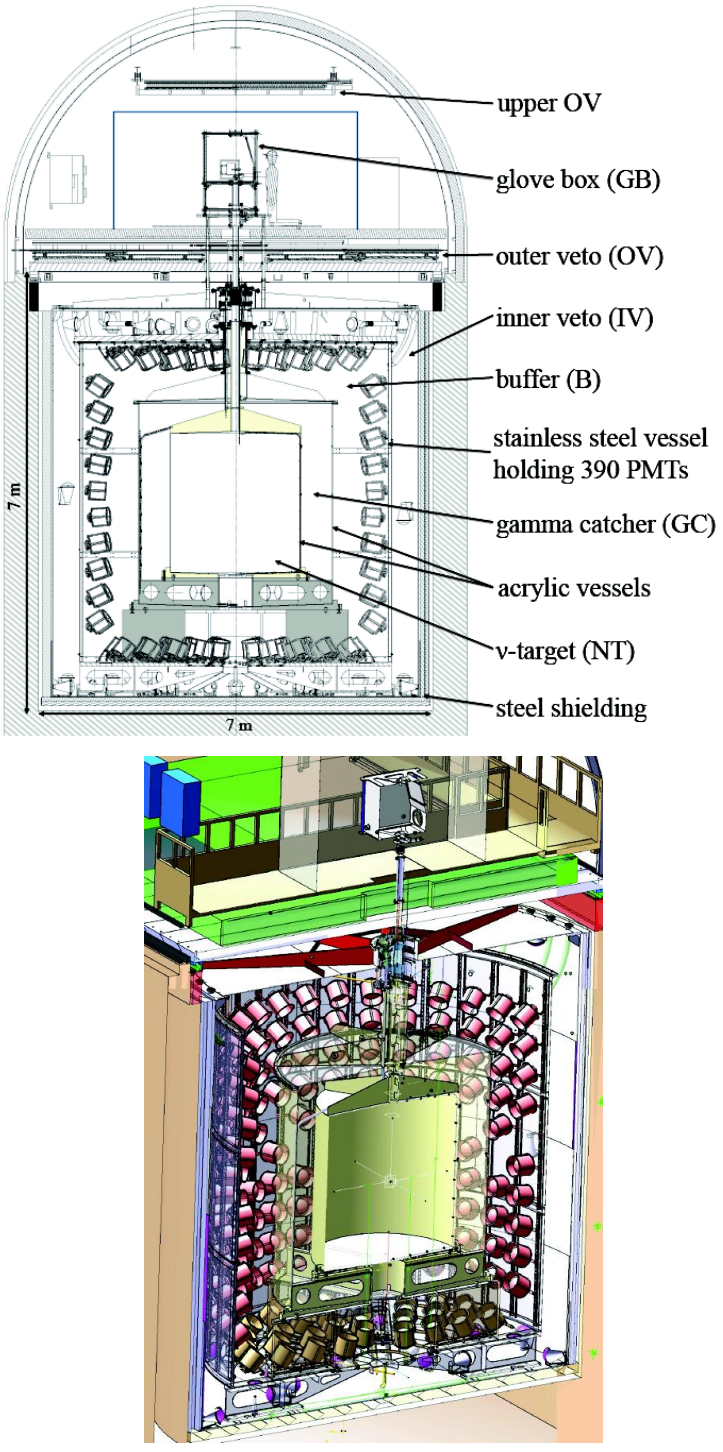


Figure 2.5: Top: Double Chooz Far Detector cross-sectional view. Bottom: Axonometric projection of the sectioned detector. Courtesy of CEA IRFU.

neutron capture probability ($\sim$ 85%) and high light yield (since the Gd compound decreases it). A small amount of oxolane (0.5% by weight) is added to liquefy the Gd complex during the production process.

Gamma Catcher

The Gamma Catcher is a 55 cm thick layer of liquid scintillator (Gd-free) which surrounds the Target to maximize the energy containment of the γ-rays escaping from it, especially for the events happening at the Target edges. It has a volume of 22.5 m^3 enclosed by a 12 mm thick acrylic vessel. In order to have the same photoelectron yield and density as the Target, the scintillator solvent volume concentration is changed to 4% of o-PXE, 30% of n-dodecane and 66% of a non-aromatic mineral oil (Shell Ondina 909) [128]. As in the Target, PPO (2 g/l) and bis-MSB (20 mg/l) are used as fluors.

Buffer

The Buffer is a 105 cm thick layer of a non-scintillating mixture (by volume, 53.5% medicinal white oil (Shell Ondina 917) and 46.5% n-alkanes (Cobersol C70)) which surrounds the Gamma Catcher. It is an innovation with respect to the CHOOZ detector (cf. figure 1.10), and it serves as a shield against incoming radioactivity from the PMTs or the surrounding rock. It amounts to 110 m^3 and is contained in a 3 mm thick stainless steel tank.

The PMTs are attached to the Buffer tank inner surface by angle-adjustable mounting jigs, and are oriented to provide a uniform light collection. Each PMT is shielded by a mu-metal cylinder from external magnetic fields [129], such as the ones from the detector shielding and the Earth.

All the liquids used are ensured to be compatible with the detector materials, especially with the delicate acrylics, and comply with the safety regulations. Concerning transparency, the attenuation lengths in the region of the scintillator emission ($\sim$ 430 nm) are measured to be $>$ 5 m, longer than the detector dimensions.

2.3.2 Outer Detector

The Outer Detector is the part of the detector oriented toward the attenuation and detection of backgrounds. Unlike the Inner Detector, it does not need to be identical between the two Double Chooz detectors. Since this thesis only deals with Far Detector data, the description is limited to the Far Detector case. It is formed by two subdetectors, the Inner Veto and the Outer Veto, which are used as background vetoes; and a passive shielding given by 15 cm of demagnetized steel which encloses the Inner Veto tank, protecting the main detector from external γ-rays.

Inner Veto

The Inner Veto is the fourth and outermost of the volumes which comprise the cylindrical main detector, and it is optically separated from the other three by the Buffer tank. It is a 10 mm thick stainless steel vessel (radius 3.3 m and

height 6.8 m) filled with 90 m^3 of liquid scintillator, in which the Inner Detector is immersed. The scintillator solvent is a mixture of 50% LAB (linear alkyl benzene) and 50% n-alkanes (Cobersol C70) by volume; with 2 g/l of PPO as fluor and 20 mg/l of bis-MSB as secondary wavelength-shifter. It provides the Inner Detector with a 50 cm thick layer of active shield against external radioactivity and fast neutrons which also serves to detect cosmic muons passing through.

The Inner Veto is optimized to detect a high number of photoelectrons per deposited energy and achieve a high-efficiency rejection of muons and the correlated background associated to them [130]. In order to do so, the inner surface of the Inner Veto is painted with a highly reflective white coating and the Buffer outer side wall is covered with a reflective VM2000 film. 78 8-inch PMTs (Hamamatsu R1408, previously used in the IMB and Super-Kamiokande experiments) are positioned around the Inner Veto: 24 on the top, 12 on the side at mid height and 42 on the bottom. Each PMT and its base are encapsulated in a stainless steel cone with a mu-metal shield fitted in, and closed by a transparent PET window. This capsule is filled with mineral oil to match the optical properties of the scintillator, and it is used to protect the electronic components from the scintillator chemicals.

Outer Veto

The Outer Veto is a plastic scintillator tracker which is used as muon veto. It consists of two parts: the largest one, known as the Lower Outer Veto, is placed above the main detector, directly over the 15 cm shielding, and it covers a 13 m × 7 m area centered on the chimney, leaving an aperture of 110 cm × 30 cm around it. The smallest one, known as the Upper Outer Veto, is hanged from the laboratory ceiling and covers the chimney area.

The Outer Veto is formed by 44 modules (36 for the Lower part and 8 for the Upper part). Each module has 64 scintillator strips arranged in two equal layers which are separated by 2.5 cm. The strips are made of polystyrene with 1% of PPO and 0.03% of POPOP (1,4-bis(5-phenyloxazol-2-yl)) as fluors, have a cross-section of 5 cm × 1 cm and are 320 or 360 cm long, with a 1.5 mm diameter wavelength-shifting fiber (Kuraray Y-11 (175 ppm), cladded) running through them. The 64 fibers have one end mirrored and another connected to a multi-anode PMT (Hamamatsu H8804). To improve the light containment, the strips are coated with a reflective layer of titanium dioxide. The modules are laid over the main detector in two layers with the strips orthogonally aligned, allowing a two-dimensional reconstruction of the impinging muons.

The Outer Veto installation began when the main detector was already operating. This results in a fraction of the data having no Outer Veto information, another fraction having only the Lower Outer Veto active, and a final one with the full Outer Veto configuration.

2.3.3 Acquisition systems

The data acquisition system used in the main detector is outlined in figure 2.6. The light produced by an energy deposition in the Inner Detector or the Inner Veto reaches the PMTs, where it is converted into an amplified electric current (the PMT gain is 10^7). The signal is extracted using the same cable

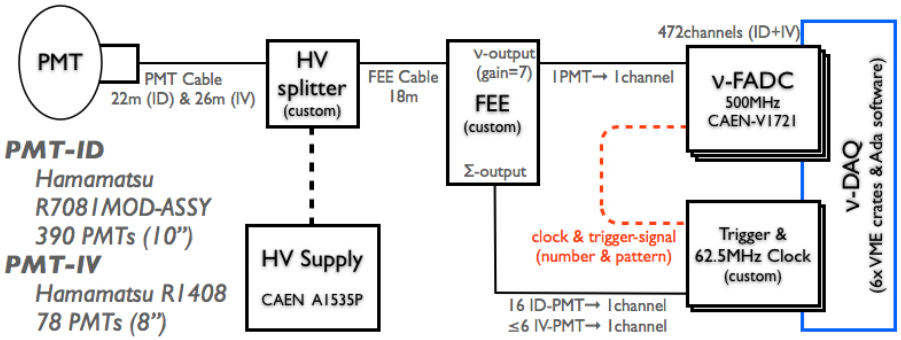


Figure 2.6: Block diagram of the readout and data acquisition systems in the main detector. From [1].

which carries the high voltage (HV) from the HV power supply (CAEN-A153P) to the PMT. An intermediate custom HV-splitter circuit decouples the signal ($\sim 5\,\text{mV}$ per photoelectron) from the HV component ($\sim 1.3\,\text{kV}$). Afterwards, the signal is sent to the Front End Electronics (FEE), where it is amplified and clipped, the baseline is corrected and the coherent noise is filtered. This stage prepares the signal for digitization, which happens at the ν-FADC system.

The FEE also produces sums of the Inner Detector (Inner Veto) PMT signals in groups up to 16 (6) which are sent to a custom Trigger System [131]. There, a trigger decision is made based on the energy and multiplicity of the group inputs, resulting in a trigger threshold at $\sim 350\,\text{keV}$ for the Inner Detector, well below the minimum visible energy of an IBD positron ($\approx 1.02\,\text{MeV}$, from equations 2.2 and 2.4); and $\sim 10\,\text{MeV}$ for the Inner Veto, corresponding to a 8 cm track of a minimum ionizing muon.

The ν-FADC system are 64 modules (CAEN VX1721) which perform flash analog-to-digital conversion of the signal waveforms. Each module has 8 channels with 8-bit resolution which are sampled at 500 MHz. Each channel is connected to a single PMT and has a memory buffer capable of storing 1024 waveforms of $4\,\mu\text{s}$ each, allowing the detector to acquire data in a dead-time free way within the expected trigger rates. Upon trigger, 256 ns of waveform from the 478 PMTs are recorded.

The Outer Veto acquisition system is independent of the main detector one. Each multianode PMT is connected to a custom front-end board with a readout chip (MAROC2), which allows to adjust the gain of the individual channels and compare them to a common analog theshold. When two overlapping scintillator strips are hit, a FPGA (Altera EP1C6Q240) triggers the readout and the signals are sent to a multiplexed 12-bit ADC and recorded. Afterwards, the Outer Veto data are aligned in time and merged with the ν-FADC data during the offline processing.

2.3.4 Calibration systems

Several calibration systems are used to survey the detector response.

A multi-wavelength LED-fiber system allows to inject light in the detector.

Source	Main total γ energy (MeV)	Origin
^{137}Cs	0.662	Deexcitation upon β^- decay.
^{68}Ge	1.022*	Annihilation of the e^+ of the ^{68}Ga β^+ decay.
^{60}Co	2.506*	Deexcitation upon β^- decay.
^{252}Cf	2.224	Capture of the fission neutron on ^{1}H.
^{252}Cf	7.937*	Capture of the fission neutron on ^{157}Gd.
^{252}Cf	8.536*	Capture of the fission neutron on ^{155}Gd.

Table 2.2: γ-ray sources used for detector calibration. The total energy of the most intense emission is quoted. The asterisk indicates that the energy is divided into two or more γ-rays (e.g. the ^{68}Ge entry corresponds to two 0.511 MeV γ rays; the ^{60}Co entry corresponds to two sequential γ rays with energies 1.173 and 1.333 MeV). Energies and decay information from [38].

The light is generated by UV LEDs with wavelengths of 385, 425 and 470 nm for the Inner Detector, and 365 and 475 nm for the Inner Veto. The rate, intensity and width of the light pulses are controlled remotely. The light is guided into the detector through optical fibers with their ends attached to some of the PMT support structures. Some of the fibers produce pencil beams of light and some have a diffuser to illuminate wide areas of the detector. Calibration runs are taken regularly to monitor the PMT and the readout electronic gains, and the time offsets.

The Inner Detector chimney can be opened to insert calibration devices. A glove box on top of the detector ensures the cleanliness of the operation. A diffuser ball connected through an optical fiber to an external laser system can be lowered into the detector, allowing the measurement of the time offset of each readout channel with 0.15 ns precision. Encapsulated radioactive sources, with ~ 50 Bq activities, are used to produce γ-rays with known energies within the detector (see table 2.2). A motorized pulley-weight system is used to deploy the sources along the Target symmetry axis (also known as z axis) with 1 mm precision. A rigid hermetic tube, the Guide Tube (see figure 2.7), allows the circulation of the sources inside the Gamma Catcher volume.

In addition, several natural sources can be used to calibrate the detector response in the full volume and extended in time. One example are the spallation neutrons released by the cosmic muons, which are mostly captured on H and Gd, but also on C, yielding ≈ 4.95 MeV [38]. Another example is the decay of the ^{210}Po from the ^{232}Th chain, which is part of the natural radioactive contamination, and produces a 8.95 MeV α [38] which is detected with a visible energy ~ 1 MeV due to the scintillator quenching.

2.4 Signal simulation

As explained in section 2.1, due to the phased construction of the Double Chooz experiment, a one-detector period was foreseen. In this period, the measurement of θ_{13} comes from the comparison of the observed DATA with the prediction of a MC simulation of the experiment, searching for the deficit of electron antineutrinos caused by the oscillation. Since the precision on θ_{13} is

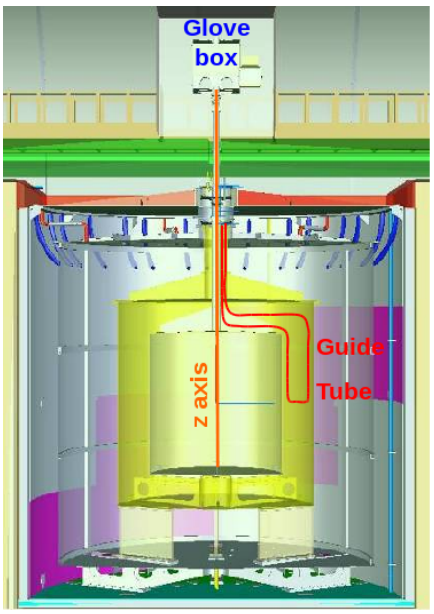


Figure 2.7: Sketch of the available deployment positions within the detector using the z axis and Guide Tube systems.

limited by the knowledge of the reactor flux and the $\overline{\nu}_e$ detection, efforts were made to produce accurate reproductions of the two.

2.4.1 Antineutrino flux prediction

In the null oscillation hypothesis, the expected rate of $\overline{\nu}_e$ from a reactor R detected in one detector through the IBD reaction at time t is given by:

$$\frac{\mathrm{d}N_R(t)}{\mathrm{d}t} = \frac{\epsilon N_p}{4\pi L_R^2} \frac{P_{\mathrm{th},R}(t)}{\langle E_\mathrm{f}\rangle_R(t)} \langle\sigma_\mathrm{f}\rangle_R(t), \tag{2.7}$$

where ϵ is the signal detection efficiency and N_p is the number of protons in the detector considered. L_R^2 represents the distance between the reactor R and the detector. The remaining variables are reactor-related and time dependent: $P_{\mathrm{th},R}(t)$ denotes the thermal power, $\langle E_\mathrm{f}\rangle_R(t)$ denotes the mean energy released per fission, and $\langle\sigma_\mathrm{f}\rangle_R(t)$ denotes the mean cross-section per fission.

The value of the detection efficiency depends on the selection cuts used in the analysis of the data; so its calculation is deferred to the chapters 4, which covers the two $\overline{\nu}_e$ selections developed by Double Chooz, and 5, which is devoted to the detection efficiency of the neutrons created in the IBD reaction. The number of protons is included in the detector simulation; the estimation can be found later in section 2.4.2.1. The baselines of the two cores and the Far Detector were already given in section 2.1. Henceforth, the discussion is restricted to the reactor-related variables.

Thermal power

The instantaneous thermal power of each reactor core, $P_{\mathrm{th},R}(t)$, is provided by the power plant company (Électricité de France) in time steps of less than 1 minute. It is computed using the core instrumentation; and it is dominated by the measurement of the water temperature in the primary loop [4].

The precision of the thermal power measurement determines the maximum power at which the core is allowed to operate, so the power plant company has a detailed knowledge about it. The calibration of the core instrumentation is tested weekly measuring the heat balance in the secondary loop. If a deviation beyond the uncertainty on the heat balance is found, the core instrumentation is recalibrated. The test is carried out with the reactor at full power, resulting in a 0.5% relative uncertainty on the thermal power. This uncertainty increases slightly when the reactor is not operating at full power, but this condition is rarely met; so the variation can be neglected to a very good approximation when integrating over all the thermal powers.

Mean energy released per fission

The mean energy released per fission, $\langle E_\mathrm{f}\rangle_R(t)$, is computed as

$$\langle E_\mathrm{f}\rangle_R(t) = \sum_k \alpha_{k,R}(t)\langle E_\mathrm{f}\rangle_k, \tag{2.8}$$

where $\alpha_{k,R}(t)$ is the fractional fission rate of the nuclide k, that is, the number of fissions of the nuclide k divided by the total number of fissions occurring in the reactor core R per unit of time. $\langle E_\mathrm{f}\rangle_k$ is the mean energy released in a fission of the nuclide k. The index k runs over the nuclides being fissioned. It turns out that 99.7% of the fissions are due to four nuclides: ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu; so it is enough to consider just these in order to predict the $\overline{\nu}_e$ flux. The uranium isotopes are initially present in the fresh nuclear fuel (uranium dioxide, with an initial ^{235}U enrichment ranging between 1.8% and 4%). The plutonium isotopes are created as the fuel is burned up in the reactions:

$$^{238}\mathrm{U} \xrightarrow{(n,\gamma)} {}^{239}\mathrm{U} \xrightarrow[23.45\,\mathrm{min}]{\beta^-} {}^{239}\mathrm{Np} \xrightarrow[2.356\,\mathrm{d}]{\beta^-} {}^{239}\mathrm{Pu}$$

$$^{239}\mathrm{Pu} \xrightarrow{(n,\gamma)} {}^{240}\mathrm{Pu} \xrightarrow{(n,\gamma)} {}^{241}\mathrm{Pu}$$

(half-lives from [38]).

The fractional fission rate of each nuclide will change with time as the fuel is consumed. This evolution is simulated with the `MURE` code [132, 133] (see figure 2.8). `MURE` (`MCNP` Utility for Reactor Evolution) uses the Monte Carlo particle transport code `MCNP` [134] to model the neutron flux inside a 3D reactor core, and solves numerically the differential equations (Bateman equations) which describe the time-evolution of the fuel composition. The validity of the `MURE` results has been demonstrated by performing crosschecks with the `DRAGON` code [135], with a two-dimensional deterministic simulation of the neutron transport,

[4] In a pressurized water reactor, the primary loop contains water which is heated directly by the nuclear fuel assemblies. This water is kept liquid at a high pressure, so it does not boil. The heat is transferred to a secondary loop, which contains water which is transformed into steam and directed to a turbine coupled to an electric generator.

Nuclide	$\langle E_\mathrm{f} \rangle_k$ (MeV)
^{235}U	201.92 ± 0.46
^{239}Pu	209.99 ± 0.60
^{238}U	205.52 ± 0.96
^{241}Pu	213.60 ± 0.65

Table 2.3: Mean energy released per fission of nuclide. From [137].

and comparisons with the experimental data obtained from destructive assay of fuel rods of the Takahama-3 reactor [136].

In order to model the reactor core accurately, the geometry and the materials within it must be defined in the `MURE` simulation. At the beginning of a reactor cycle, only one third of the fuel assemblies is fresh, the rest corresponds to fuel partially used. The power plant company makes available the location and the initial burnup of each assembly. By running the `MURE` and `DRAGON` simulations for the assemblies, it is possible to determine their initial composition; which is also compared to the result of the power plant company simulation (`APOLLO2-F`). Then, `MURE` is used to simulate the time-evolution of the core in steps of 6–48 hours, depending on the stability of the reactor operation.

The systematic uncertainties on the fractional fission rates are estimated from the discrepancies found when the simulation inputs are varied. The systematic uncertainty is dominated by the uncertainties on the neutron moderator density and the position of the control rods which are used to regulate the neutron flux inside the core. Other contributions are the thermal power, the boron concentration in the primary loop (boron is a neutron absorber and is used to control the fission rate), the neutron moderator temperature, the initial burnup error, the nuclear databases and mean energy released per fission data considered, and the statistical error of the Monte Carlo. Ultimately, the uncertainty on the fractional fission rates implies a 0.8% uncertainty on the normalization of the $\overline{\nu}_e$ flux.

Returning to equation 2.8, the mean energy released per fission of a nuclide is found by computing the global Q value corresponding to the fission reaction and the subsequent β^- decays of the fission products, which is known with great precision. To this global Q value, the energy deposited within the core by the capture of fission neutrons not contributing to the chain reaction must be added; and the energy taken away by the escaping $\overline{\nu}_e$ and the undecayed fission products must be subtracted [137]. The resulting values of the mean energies released per fission of the four nuclides are given in table 2.3. The uncertainty on them is translated into a 0.2% uncertainty on the $\overline{\nu}_e$ rate.

Mean cross-section per fission

The mean cross-section per fission corresponds to the IBD reaction cross-section averaged over the full energy spectra of the $\overline{\nu}_e$ emitted following a fission. It is written as

$$\langle \sigma_\mathrm{f} \rangle_R(t) = \sum_k \alpha_{k,R}(t) \langle \sigma_\mathrm{f} \rangle_k, \tag{2.9}$$

where $\alpha_{k,R}(t)$ is the aforedescribed fractional fission rate of the nuclide k, and $\langle \sigma_\mathrm{f} \rangle_k$ is the mean cross-section per fission of the nuclide k, which is computed

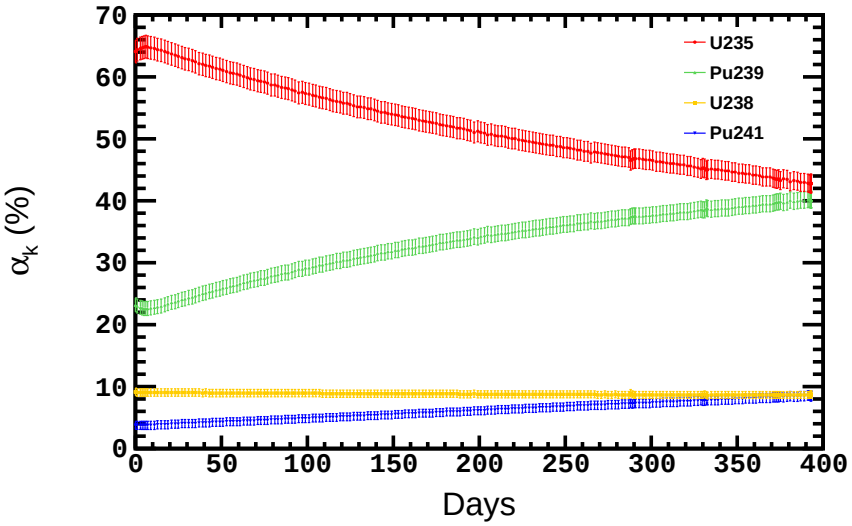


Figure 2.8: Evolution of the fractional fission rates as a function of time calculated with the MURE code for a typical fuel cycle of a Chooz reactor. The error bars represent the uncertainty on each fractional fission rate. The contribution of ^{235}U (red) to the fission decreases as it is progressively "burned"; while the contributions of ^{239}Pu (green) and ^{241}Pu (blue) increase as they are bred. The fission rate of ^{238}U (orange) remains approximately constant. Edited from [138].

as

$$\langle \sigma_{\mathrm{f}} \rangle_k = \int_0^\infty S_k(E_{\overline{\nu}_e}) \sigma_{IBD}(E_{\overline{\nu}_e}) \mathrm{d}E_{\overline{\nu}_e}. \tag{2.10}$$

$\sigma_{IBD}(E_{\overline{\nu}_e})$ is the IBD cross-section (equation 2.3) and $S_k(E_{\overline{\nu}_e})$ is the cumulated energy spectrum of all the $\overline{\nu}_e$ emitted upon a fission of the nuclide k.

When a nuclide fissions stimulated by a neutron capture, it usually splits into two lighter nuclei, known as the fission products. These are unstable neutron-rich nuclei which undergo several β^- decays, in which the $\overline{\nu}_e$ are produced, before reaching stability (typically, each fission product decays thrice, resulting in 6 $\overline{\nu}_e$ per fission).

Since there are not measurements of the fission product $\overline{\nu}_e$ spectra with the required precision, measurements of the electron spectra are used instead, which are then converted into $\overline{\nu}_e$ spectra. For the fissile nuclides ^{235}U, ^{239}Pu and ^{241}Pu, the cumulated electron spectra from references [139, 140, 141] are used. These spectra are obtained by exposing thin films enriched with one of the fissile nuclides to the thermal neutron flux of the research reactor at the Institut Laue-Langevin (ILL), and analyzing the electron spectrum with a magnetic β spectrometer. The conversion to $\overline{\nu}_e$ spectrum is carried out according to reference [98] by fitting virtual β branches—functions with the shape of an hypothetical allowed[5] branch with an effective nuclear charge whose parameters (normalization, endpoint energy) are determined by the fit—to the electron spectrum, which are then converted into $\overline{\nu}_e$ spectra invoking energy conservation. A correction accounting for the off-equilibrium effects arising due to the limited exposure of the films to the neutron flux at ILL (up to $\sim$ 2 days) in contrast to the much longer irradiation time inside a reactor like the ones at Chooz ($\sim$ 1 year) is also included [114].

The fissionable nuclide ^{238}U has a negligible fission cross-section for thermal neutrons, which prevented a measurement at the ILL reactor. A recent measurement of the cumulated electron spectrum of ^{238}U using a fast neutron beam at the Forschungsreaktor München II (FRM II) neutron source is used [142]. The low statistics of the electron spectrum at high energy preclude the use of the method based on virtual β branches, so the conversion into an $\overline{\nu}_e$ spectrum is made with a transfer function. This transfer function is computed using the measured cumulated electron spectrum of ^{235}U [140] and a corresponding $\overline{\nu}_e$ spectrum calculated by an *ab initio* approach—summation of the individual $\overline{\nu}_e$ spectra produced by the numerous branches of β^- decays of the ^{235}U fission products, which are obtained from electron spectroscopy experimental data [114]. The ^{238}U $\overline{\nu}_e$ spectral data from reference [142] is restricted to the $2.875 - 7.625\,\mathrm{MeV}$ energy range, so an extrapolation to lower and higher energies is made by combination of an exponential-polynomial fit to the spectral data and the *ab initio* spectrum for ^{238}U [143], whose uncertainty is reduced by expressing it as a relative spectrum to other *ab initio* spectra (^{239}Pu and ^{241}Pu), which is then made absolute using the spectra from [98].

As it can be concluded from the previous explanation, the obtention of the $\overline{\nu}_e$ spectra is rather complicated, what makes the mean cross-section per fission

[5]The shape is further corrected by considering the finite size of the nucleus, the screening of the nuclear charge due to the atomic electrons, radiative corrections and the effect of weak magnetism.

the dominant source of systematic uncertainty on the $\overline{\nu}_e$ flux. This uncertainty varies with the antineutrino energy, but it is at the level of 3% of the $\overline{\nu}_e$ normalization. In order to improve the precision of the θ_{13} measurement during the one-detector phase of the experiment, the measurement of the mean cross-section per fission at the Bugey4 experiment [144], $\langle\sigma_\mathrm{f}\rangle^{\mathrm{B4}}$ is used. In this case, the mean cross-section per fission (equation 2.9) is rewritten as

$$\langle\sigma_\mathrm{f}\rangle_R(t) = \langle\sigma_\mathrm{f}\rangle^{\mathrm{B4}} + \sum_k (\alpha_{k,R}(t) - \alpha_k^{\mathrm{B4}})\langle\sigma_\mathrm{f}\rangle_k, \tag{2.11}$$

where α_k^{B4} is the published fractional fission rate of the nuclide k in the Bugey reactor.

Bugey4 was a short-baseline (15 m) reactor experiment which made a precise measurement of the mean cross-section per fission, $\langle\sigma_\mathrm{f}\rangle^{\mathrm{B4}} = (5.752 \pm 0.081) \cdot 10^{-43}\,\mathrm{cm}^2/\mathrm{fission}$ (1.4% relative uncertainty). Because the composition of the Bugey reactor was very similar to the ones of the Chooz reactors, the term between brackets of equation 2.11 is very small, thus suppressing the contribution of $\langle\sigma_\mathrm{f}\rangle_k$ to the normalization uncertainty, which changes to 0.2%. In addition, an hypothetical deficit due to the oscillation of the $\overline{\nu}_e$ into a fourth sterile antineutrino with a squared mass difference $\Delta m^2_{41} \gtrsim 1\,\mathrm{eV}^2$ (as suggested by the anomalies discussed in section 1.4) would be already included in the Bugey4 measurement, rendering the θ_{13} oscillation fit unaffected by such sterile antineutrino. In a certain sense, the Bugey4 experiment serves as an effective near detector until the actual Double Chooz Near Detector is ready.

In the preceding discussion, only the expected rate resulting from integration over the complete $\overline{\nu}_e$ energy spectrum has been considered. However, the expected shape is also relevant since the oscillation probability is energy-dependent. In order to have a prediction of the detected $\overline{\nu}_e$ flux binned in $\overline{\nu}_e$ energy, equation 2.7 must be modified to

$$\begin{aligned}\frac{\mathrm{d}N_{i,R}(t)}{\mathrm{d}t} =& \frac{\epsilon N_p}{4\pi L_R^2}\frac{P_{\mathrm{th},R}(t)}{\langle E_\mathrm{f}\rangle_R(t)}\times \\ &\times\left(\frac{\langle\sigma_\mathrm{f}\rangle^{\mathrm{B4}} + \sum_k(\alpha_{k,R}(t) - \alpha_k^{\mathrm{B4}})\langle\sigma_\mathrm{f}\rangle_k}{\sum_k \alpha_{k,R}(t)\langle\sigma_\mathrm{f}\rangle_k}\sum_k \alpha_{k,R}(t)\langle\sigma_\mathrm{f}\rangle_{i,k}\right),\end{aligned} \tag{2.12}$$

where the index i represents the i-th bin, corresponding to the $\overline{\nu}_e$ energy interval $[E_i, E_i + \Delta E_i)$. This requires to define the mean cross-section per fission of the nuclide k in the energy bin i as:

$$\langle\sigma_\mathrm{f}\rangle_{i,k} = \int_{E_i}^{E_i+\Delta E_i} S_k(E_{\overline{\nu}_e})\sigma_{IBD}(E_{\overline{\nu}_e})\mathrm{d}E_{\overline{\nu}_e}. \tag{2.13}$$

Double Chooz custom Monte Carlo event generator, `DCRxtrTools` [145], generates $\overline{\nu}_e$ according to equation 2.12 for each of the detector runs used in the oscillation analysis, reading the time-dependent reactor variables (thermal power and fractional fission rates) from a database. It also assigns an $\overline{\nu}_e$ interaction vertex with a proton of the simulated detector, and produces the positron and the neutron resulting from the IBD reaction with the proper kinematics [123].

Source	Uncertainty (%)
Bugey4 mean cross-section per fission	1.4
Fractional fission rates	0.8
Thermal power	0.5
Mean cross-section per fission	0.2
Mean energy released per fission	0.2
Distance between cores and detector	< 0.1
Total	1.7

Table 2.4: Reactor $\overline{\nu}_e$ flux normalization uncertainties. Shape uncertainties not included here are also accounted in the covariance matrix. The uncertainties due to the proton number and signal detection efficiency are considered in sections 2.4.2.1 and chapter 5, respectively.

Additionally, it produces a covariance matrix binned in $\overline{\nu}_e$ energy which describes the uncertainty on the $\overline{\nu}_e$ spectrum due to the variables used in equation 2.12 (see table 2.4).

2.4.2 Detector simulation

Double Chooz has developed a complete detector simulation which reproduces step by step all the processes happening in the real detector, starting with the interaction of the secondary particles (positron and neutron) with the detector materials and finishing with the charge output of the readout system. The detector simulation, `DCGLG4sim`, is based on the `Geant4` toolkit (`Geant4.9.2.p02`) [146, 147], with specific features for liquid scintillator antineutrino detectors. The *physics list* (the Physics processes which can experience the simulated particles) is similar to `Geant4`'s `QGSP_BERT_HP` list, which includes the high precision neutron package `NeutronHP` to describe the neutron physics from 20 MeV to thermal energies. In order to enhance the accuracy of the simulation, custom models for neutron scattering and radiative capture are used for neutrons at epithermal energies and below (further details are given in section 5.2 of the chapter 5 devoted to the neutron detection efficiency estimation).

The detector geometry and materials are implemented in the simulation to a great level of detail, including the proton number of the detector (its measurement is described in the following section 2.4.2.1). The dimensions of the detector vessels and supporting structures were measured during assembly and installation, and the placement was validated by photographic survey. The position and orientation of the PMTs were determined also by photographic survey with sub-millimeter accuracy. The PMT's photocathode is custom-modelized as a thin semitransparent surface with absorption and refractive indices from [148], and includes the variation in the collection efficiency dependent on where the photoelectron was emitted.

The liquid scintillators are modeled based on measurements of their optical parameters: the relative light yield of the Target liquid with respect to the Gamma Catcher one [128] (relative and absolute light yields in the simulation were further tuned from calibration data taken with the detector), scintilla-

tor emission spectrum, photon emission time probabilities, attenuation and re-emission probabilities [127], the ionization quenching (also known as Birk's law quenching) for electrons [149] and α particles [150] of different energies. The optical model also includes the measured refraction indices and reflectivities of the detector materials.

The readout system simulation, `RoSS`, reproduces the response of the full chain (PMT, front-end electronics, flash-ADC, trigger and the data acquisition systems) to a photoelectron. Given the time when the PMT was hit by a photon, `RoSS` generates the corresponding digitized waveform. The model is based on an empirical probability density function which is built from measurements of the response of the components. The readout characteristics (gain, baseline, noise, single photoelectron width...) are implemented channel-wise, so the output of the simulation exhibits a similar dispersion as the DATA.

2.4.2.1 Proton number

The number of free protons (the nuclei of hydrogen-1 atoms) determines the number of IBD interactions occurring in the detector. During the one-detector phase, it is essential to know it accurately to have a reliable prediction of the number of $\overline{\nu}_e$ expected (in the two-detector phase, the uncertainty in this value is irrelevant since both detectors are identical).

The proton number in the Target liquid governs almost completely the number of IBD interactions which end up with the neutron captured on Gd [6]. From the precise knowledge of the chemical composition and weight of the ingredients used to produce the scintillator, the hydrogen fraction in the Target is calculated to be $13.60 \pm 0.04\%$ by weight [127]. After measuring the total mass of the Target scintillator, $(6.747 \pm 0.020) \cdot 10^{29}$ protons are estimated [69].

The proton number in the Gamma Catcher liquid is estimated differently since the hydrogen content of the mineral oil used is not precisely known. A CHN (Carbon, Hydrogen, Nitrogen) analysis found a hydrogen fraction of $14.6 \pm 0.2\%$ by weight. The total volume of the Gamma Catcher scintillator is estimated from geometrical measurements and converted into mass using the well-known density, resulting in $(1.582 \pm 0.016) \cdot 10^{30}$ protons [72].

Although the acrylic vessels and supporting structures do not scintillate, turning the positron kinetic energy invisible, the IBD interactions within them can be detected when the γ-rays escape. The proton number for those bodies is computed from their mass, assuming a 8.05% a hydrogen fraction (from the monomer formula $C_5H_8O_2$). The results are $(2.08 \pm 0.48) \cdot 10^{28}$ protons in the Target and $(6.96 \pm 0.21) \cdot 10^{28}$ protons in the Gamma Catcher acrylics [151]; where the uncertainties have been estimated from the differences found with the respective numbers in the detector simulation.

The proton number in the Buffer is not relevant since it does not scintillate, so the IBD interactions within it are practically undetectable. Only the few $\overline{\nu}_e$ interacting close to the Gamma Catcher border have a chance to be detected (estimated in section 5.5.3). Since it is a skin effect, the hydrogen fraction rather than the total number of protons applies. A CHN analysis of the Buffer liquid found a hydrogen fraction of $14.80 \pm 0.15\%$ by weight. Hence, the proton number

[6] Except for a small number of neutrons created in the Gamma Catcher which migrate into the Target, estimated in section 5.4.3

Material	Proton number ($\times 10^{29}$)	MC correction factor	Uncertainty (%)
T liquid	6.747 ± 0.020	1.00	0.30
T acrylics	0.208 ± 0.048	1.39	23
GC liquid	15.82 ± 0.16	1.00	1.04
GC acrylics	0.696 ± 0.021	1.00	3
Buffer liquid	67.99 (MC)	1.00	1.00

Table 2.5: Proton number of the detector materials, with the associated correction factor and uncertainty to be applied to the normalization of the $\overline{\nu}_e$ interacting within them.

is taken from the detector simulation to be $6.799 \cdot 10^{30}$ with a 1% uncertainty from the CHN measurement.

All the previous materials are implemented in the detector simulation so the proton number is taken into account. Only the proton number of the Target acrylics needs to be corrected due to the differences found between the simulation and the estimation based on measurements, which might be caused by the approximate reproduction of the supporting structures in the MC simulation (the stiffener on top of the Target lid and the feet under the Target vessel in figure 2.5). A summary of the proton numbers, their correction and uncertainty is found in table 2.5. The magnitude of the final normalization correction and uncertainty which must be applied to the MC simulation due to the proton number depends on the number of the $\overline{\nu}_e$ detected in each material, which is estimated from the MC and depend on the selection cuts of chapter 4.

Chapter 3

Event reconstruction

This chapter introduces the basic variables which are computed from the reconstruction of the event and which are later used in the selection of the IBD interactions, explained in the next chapter.

A special emphasis is made on the energy. The oscillation probability is energy-dependent (e.g. equation 1.54), so the antineutrino oscillations will leave a characteristic imprint on the electron antineutrino spectrum (and on the closely related positron visible energy spectrum as equation 2.4 indicates) which should be used in order to exploit all the available empirical information to measure θ_{13} with the highest precision. This technique is known as *Rate+Shape analysis*, since both normalization and spectral shape are considered (this is discussed further in section 6.2). In fact, Double Chooz was the first experiment to measure a non-null θ_{13} using a Rate+Shape analysis [69]. Moreover, because the background sources have energy spectra different from the $\overline{\nu}_e$ one, their rates in the selected IBD candidate sample (which contains both $\overline{\nu}_e$ and background events) can be constrained from the total IBD candidates spectrum, increasing further the precision on θ_{13}. In addition, the $\overline{\nu}_e$ selection relies on energy cuts to select the positron and neutron signals from the IBD reaction, so the energy accuracy has an explicit impact on the $\overline{\nu}_e$ normalization. All these capabilities require to dominate the energy response of the detector.

The ambitious Monte Carlo simulation of the Double Chooz detector (introduced in section 2.4.2) produces signal waveforms equivalent to those of the real detector; so the energy reconstruction is done in parallel for both DATA and MC. The visible energy per trigger, E_{vis}, is obtained as

$$E_{\text{vis}} = N_{\text{pe}}^{m} \cdot f_{\text{u}}^{m}(\rho, z) \cdot f_{\text{MeV}}^{m} \cdot f_{\text{s}}^{m}(E_{\text{vis},\odot}^{m}, t) \cdot f_{\text{qnl}}^{m}(E_{\text{vis},\odot}^{m}), \tag{3.1}$$

where the superscript m refers to DATA or MC. The first step is to obtain N_{pe}^{m}, the total number of photoelectrons of the event, which is calculated from the reconstructed charge in each channel as explained in section 3.1 and then converted into photoelectrons as described in section 3.2.

The second step is the uniformity calibration, where $f_{\text{u}}^{m}(\rho, z)$ is a function of the reconstructed vertex coordinates $\rho = \sqrt{x^2 + y^2}$ and z (introduced in section 3.3) and corrects for the spatial dependence of the photoelectron yield as described in section 3.4.

The third step is the conversion to energy units, carried out by f_{MeV}^{m}, a conversion factor from photoelectrons to MeV (introduced in section 3.5).

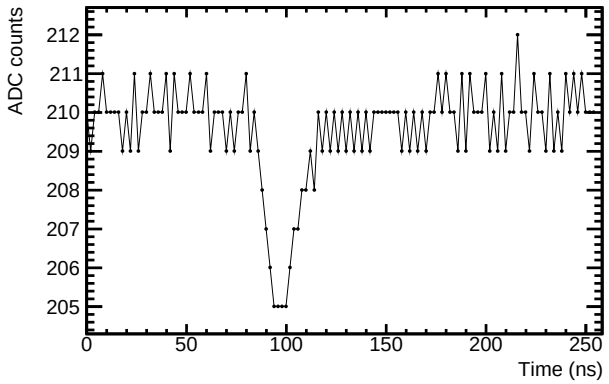


Figure 3.1: Example of a recorded digitized waveform from one readout channel. The dots indicate the samples taken every 2 ns.

The fourth step is the stability calibration, where $f_{\rm s}^{m}(E_{{\rm vis},\odot}^{m},t)$ is a function of the uniform energy, $E_{{\rm vis},\odot}^{m}=N_{\rm pe}^{m}\cdot f_{\rm u}^{m}(\rho,z)\cdot f_{\rm MeV}^{m}$, and the trigger time t, and corrects for the temporal dependence of the energy scale, which only affects the DATA (ergo for the simulation $f_{\rm s}^{\rm MC}(E_{{\rm vis},\odot}^{\rm MC},t)=1$) as explained in section 3.6.

The fifth and last step is the linearity calibration, where $f_{\rm qnl}^{m}(E_{{\rm vis},\odot}^{m})$ is a function of $E_{{\rm vis},\odot}^{m}$ and corrects the non-linearity of the MC with respect to the DATA (so $f_{\rm qnl}^{\rm DATA}(E_{{\rm vis},\odot}^{\rm DATA})=1$ by definition) as discussed in section 3.7.

The calibration steps just described render a visible energy variable which is uniform within the detector, stable in time, and scales linearly with the charge in the PMTs; which is used to perform the $\overline{\nu}_e$ selection in both DATA and MC samples. However, to account for the systematic uncertainties due to the calibration process, a further implementation of the visible energy to be used in the Rate+Shape analysis is presented in section 3.8.

3.1 Charge reconstruction

The energy deposited by a particle in the scintillators is partially transformed into photons which are caught by the PMTs. These photons can cause the emission of electrons (photoelectrons) from the photocathode of the PMTs (photoelectric effect). These photoelectrons are focused by electrodes and accelerated using a voltage until they collide with the first dynode of the PMT. The collision frees more electrons (secondary emission), which are accelerated again and made to collide with the second dynode. This process is repeated for every dynode inside the PMT, resulting in a multiplication of the number of electrons. In the end, an electric current pulse is collected in the anode. This pulse is fed into the acquisition system (described in section 2.3.3) which converts it into a digitized waveform (figure 3.1 shows an example). If the event fulfills the trigger conditions, the 256 ns waveform in each readout channel are recorded for posterior analysis.

The first step to reconstruct the energy of the event is to integrate the electric current waveforms to compute the charge in each PMT. In order to

know how much of the charge is due to the event, the reference baseline of each readout channel must be known (i.e. the signal level in the digitizer modules in absence of an actual PMT pulse). This is achieved using a 1 Hz periodic trigger, in which the baseline mean (B_{mean}) and the standard deviation (σ_B) are measured using the full window (256 ns) for each of the 468 channels of the main detector (corresponding to the 390 Inner Detector PMTs and the 78 Inner Veto PMTs).

A sliding window algorithm is used to integrate the waveform. The length of the window is fixed at 112 ns, the optimal value in terms of charge resolution and reconstruction efficiency. The window position is set so as the resulting integral is maximized. The charge q is then computed from the integral value, once the pedestal (the result of integrating B_{mean} over 112 ns) has been subtracted. In case of not having an actual photoelectron signal, this algorithm would integrate over the largest baseline fluctuation, biasing the charge reconstruction. In order to ensure that the pulse comes from a photoelectron, two conditions must be satisfied to accept the reconstruction:

$$I_{\text{max}} \geq 2\,\text{ADC counts} \tag{3.2a}$$

$$q_{\text{min}} > N \cdot \sigma_B \cdot \sqrt{W}. \tag{3.2b}$$

The first condition, eq. 3.2a, states that the maximum amplitude of the waveform must reach 2 ADC counts at least with respect to the baseline (the typical amplitude of a single photoelectron is $\sim$ 6 ADC counts, cf. figure 3.1). The second condition, eq. 3.2b, establishes a minimum charge, where N is a real number, W is the integration window size in number of samples (since a sample corresponds to 2 ns, $W = 56$ for a 112 ns window), and σ_B is the standard deviation of the reference baseline in the channel considered (typically, $\sigma_B \approx 0.4$ ADC counts). The choice of N is the subject of the next section.

The reconstruction algorithm computes also the pulse start time in each PMT, defined as the time when the pulse reaches 20% of its maximum amplitude. The time offset of each readout channel is corrected using the light injection calibration systems.

3.1.1 Improvement of single photoelectron efficiency

The value of N in equation 3.2b was originally set to 5.0 for the first Double Chooz θ_{13} measurement [69] using the MC simulation. However, a non-linearity of up to 3% in the MC energy scale with respect to the DATA was observed at low energies, which was partly due to the rejection of actual signal pulses because of the too stringent minimum charge condition. This prompted a review of the N value to increase the charge reconstruction efficiency of low-energy signals.

In the low-energy range (up to a few MeV), each PMT produces one photoelectron at most. Therefore, the number of PMTs giving a reconstructed signal (PMT multiplicity) is proportional to the single photoelectron (SPE) reconstruction efficiency, but also to the noise. The capture of neutrons on hydrogen yields a 2.224 MeV γ-ray which serves as the prototypical event for this study.

The ^{252}Cf calibration source is used as the neutron source. The ^{252}Cf is an unstable nucleus (half-life of 2.645 years) which undergoes α decay or spontaneous fission with branching ratios 96.91% and 3.09%, respectively [38]. The source is doubly encapsulated in stainless steel [152], so the α particles do not

escape. In average, 3.77 neutrons per fission are released [153]. Along with the neutrons, γ radiation with a total energy of $\sim 7\,\mathrm{MeV}$ is emitted from the fission fragments. This allows to select neutron capture events following a fission using a prompt-delayed coincidence similar to the one used to select $\overline{\nu}_e$, profiting from the intrinsic background reduction granted by the coincidence requirement. In this case, the prompt trigger is given by the γ emission (and the proton recoils induced by the fast neutrons just released); and the delayed trigger is given by the radiative neutron captures on hydrogen, which happen after the neutrons have thermalized.

In order to obtain the most homogeneous illumination of all the Inner Detector PMTs, a one-hour long calibration run taken with the ^{252}Cf source at the center of the detector is used.

The mean PMT multiplicity is determined from the mean of a Gaussian fitted to the PMT multiplicity distribution of the delayed triggers. Figure 3.2 shows how the mean PMT multiplicity (red line) grows asymptotically as the charge threshold is decreased (when the amplitude threshold, eq. 3.2a, is kept at the default value). The asymptotical behavior, saturating at 260 PMTs, proves that the original $N = 5.0$ charge threshold is rejecting signal: if it were noise, the PMT multiplicity would continue growing as the threshold is reduced. The reason why the charge threshold is dispensable is the existence of the other threshold, the amplitude threshold, which is rejecting the baseline noise. In fact, if the amplitude threshold is nullified by setting it to 0 ADC counts, the mean PMT multiplicity (green line) is seen to burst as the charge threshold is decreased. Since the amplitude threshold cannot be softened more (the minimum baseline fluctuation is of 1 ADC count because of the discretization), the only way to improve the single photoelectron efficiency is to set N to a lower value. A plateau is observed for $N \lesssim 1$, so the chosen value is $N = 1.0$; since it provides the same increase in PMT multiplicity (+4%) as $N = 0.0$ with respect to $N = 5.0$, but it offers an additional protection in case a sudden noise source affected the readout channels. This value was used for first time in the second θ_{13} measurement by Double Chooz [1], and it is the current value used for all Double Chooz analyses.

Another test to study the improvement brought by reducing the charge threshold in the reconstruction is the resolution of the peak corresponding to the total absorption of the 2.224 MeV γ-ray. In order to do so, the delayed trigger energy spectrum of the ^{252}Cf source, corresponding to neutron captures on H is fitted with the following function [154]:

$$F(E) = A\big(A_\mathrm{t} N_\mathrm{t} F_\mathrm{t}(E) + F_\mathrm{p}(E)\big), \tag{3.3a}$$

where

$$F_\mathrm{t}(E) = \mathrm{e}^{sE}\,\mathrm{erfc}\left(\frac{E-\kappa\mu}{\sigma}\right) \tag{3.3b}$$

$$F_\mathrm{p}(E) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{1}{2}\left(\frac{E-\mu}{\sigma}\right)^2}. \tag{3.3c}$$

The function has 5 free parameters to be determined by the fit: μ, σ, s, A_t and A, which are explained next. Ignoring for the moment the normalization parameters A, A_t and N_t; the function has two terms, $F_\mathrm{p}(E)$ and $F_\mathrm{t}(E)$, which model the photopeak and the Compton tail, respectively.

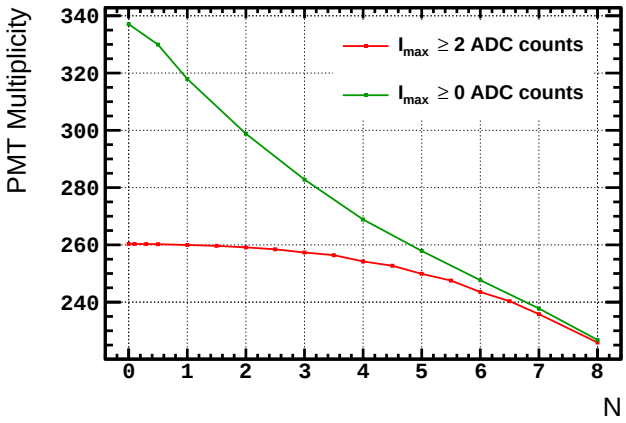


Figure 3.2: Mean PMT multiplicity as a function of the parameter N of the reconstruction charge threshold (eq. 3.2b). Data from neutron captures on H using the ^{252}Cf source. Red: maximum amplitude threshold (eq. 3.2a) at default value (2 ADC counts). Green: maximum amplitude threshold removed.

The function $F_{\mathrm{p}}(E)$ is a normalized Gaussian function with mean μ, which corresponds to the resulting energy of the photopeak after summing the reconstructed charges of the readout channels; and standard deviation σ, which accounts for the widening caused by the statistical and electronic noise. The amplitude of the photopeak is controlled by the normalization parameter A.

The function $F_{\mathrm{t}}(E)$ is the outcome of convoluting an exponential function to describe the shape of the Compton tail with a Gaussian function to model the noise, which is parametrized using the same σ as in the photopeak. This results in the exponential being multiplied by the complementary error function,

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{-t^2}\,\mathrm{d}t. \tag{3.4}$$

The Compton tail is modeled by e^{sE}, where s controls the steepness of the tail. The position of the Compton edge, $(\kappa \cdot \mu)$, is left as a function of the photopeak energy, μ. Using the expression for the maximum energy transfer in the Compton scattering, the value of κ can be calculated as

$$\kappa = 1 - \frac{1}{1 + \frac{\mu}{m_e}(1 - \cos\pi)}, \tag{3.5}$$

where m_e is the electron mass. The photopeak of the neutron capture on H has $\mu = 2.224\,\mathrm{MeV}$, so $\kappa = 0.897$. The amplitude of the Compton tail is expressed as a fraction of the photopeak amplitude, A, using the fractional tail amplitude A_{t}. In order to do so, $F_{\mathrm{t}}(E)$ is normalized to unity in the interval $(0, \mu + 2\sigma)$

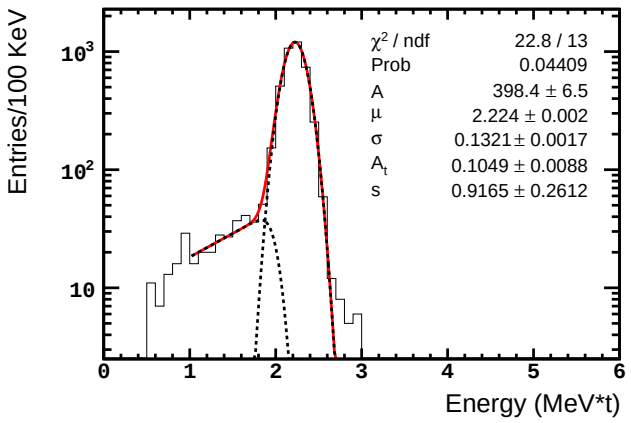


Figure 3.3: Fit of the function in eq. 3.3 to the energy spectrum of the ^{252}Cf neutrons captured on H (the parameter N of the reconstruction charge threshold is set to 5.0).

using the normalization constant $N_{\rm t}$, which is defined as

$$N_{\rm t} \equiv \left(\int_0^{\mu+2\sigma} F_{\rm t}(x)\mathrm{d}x \right)^{-1}$$
$$= s\left(\mathrm{e}^{sx}\,\mathrm{erfc}\left(\frac{x-\kappa\mu}{\sigma}\right) - \mathrm{e}^{\kappa\mu s+\sigma^2 s^2/4}\,\mathrm{erf}\left(\frac{\kappa\mu - x + \sigma^2 s/2}{\sigma}\right)\right)\Bigg|_0^{\mu+2\sigma}.$$

The resolution is estimated as the ratio between the width ($\sigma_{\rm fit}$) and the mean ($E_{\rm fit}$) energy of the H capture peak, which are taken from the best-fit values of the parameters σ and μ of the function in equation 3.3 (see figure 3.3 for an example of a fit).

Figure 3.4 shows the resolution of the peak as a function of the parameter N of the charge threshold. It is observed how the resolution worsens when the charge threshold is increased, as a result of considering a smaller number of channels for the calorimetric sum. On the contrary, the resolution improves slightly when the charge threshold is diminished (for $N = 1.0$, the improvement with respect to $N = 5.0$ is $(2.3 \pm 1.8)\%$). If reducing the charge threshold had led to the reconstruction of noise, the resolution would have deteriorated. This result confirms the conclusion of the PMT multiplicity study.

3.2 Photoelectron conversion

The total number of photoelectrons produced by one event in the detector, $N_{\rm pe}$, is obtained from the sum of the number of photoelectrons of each channel, pe_i, where the subindex i denotes the i-th channel. The number of photoelectrons in the i-th channel is calculated from the reconstructed charge q_i, using a

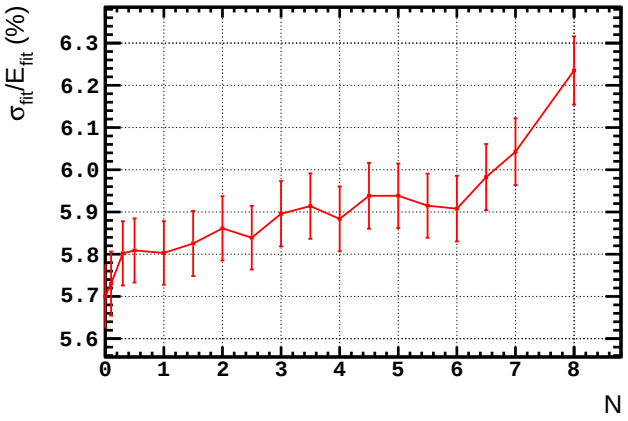


Figure 3.4: Resolution of the neutron capture on H peak (2.224 MeV) as a function of the parameter N of the reconstruction charge threshold (eq. 3.2b) using the ^{252}Cf source.

conversion function $g_i(q_i, t)$ known as gain. Therefore, it can be written

$$N_{\mathrm{pe}} = \sum_i pe_i = \sum_i \frac{q_i}{g_i(q_i, t)} \tag{3.6}$$

where the sum is over the well-behaved channels (mean charges, pedestals, pulse start times are required to be stable).

Leaving aside for the moment the dependence on the charge and time, the gain in each channel, g_i, is extracted using the light injection calibration system. Illuminating the PMTs with pulsed light from a LED with constant intensity, the mean charge observed in the i-th PMT, $\langle q_i \rangle$, is related to the mean number of photoelectrons, $\langle pe_i \rangle$, released by the LED photons as

$$\langle q_i \rangle = g_i \langle pe_i \rangle. \tag{3.7}$$

The standard deviation of the observed charge distribution, $\sigma_{q,i}$, is understood to be caused by the fluctuations in the number of photons per LED shot (which follows a Poisson distribution, hence $\sigma_{\mathrm{Poisson},i} = g_i \sqrt{\langle pe_i \rangle}$), the fluctuations of the charge obtained for each photoelectron ($\sigma_{\mathrm{SPE},i} = g_i \alpha \sqrt{\langle pe_i \rangle}$, where α is the standard deviation of the single photoelectron distribution, which we assume to be a characteristic of the PMT model so it is the same for all the PMTs and constant in time), and the electronic noise. Therefore, $\sigma_{q,i}$ can be written as

$$\sigma^2_{q,i} = \frac{1}{\alpha_0} g_i^2 \langle pe_i \rangle \tag{3.8}$$

where the parameter α_0 absorbs the dependence on the width of the single photoelectron distribution and the electronic noise. Then, from equations 3.7 and 3.8, the gain can be computed as

$$g_i = \alpha_0 \frac{\sigma^2_{q,i}}{\langle q_i \rangle}. \tag{3.9}$$

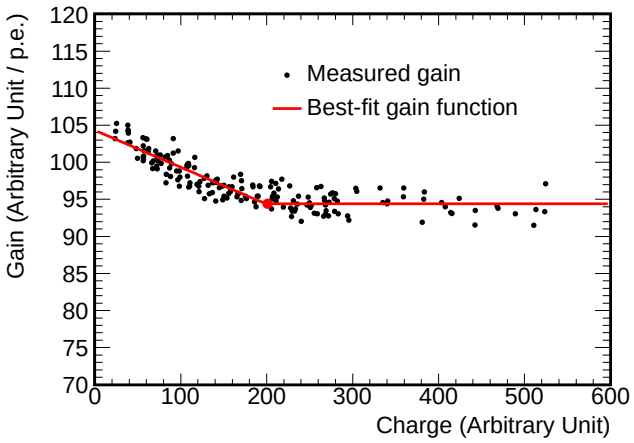


Figure 3.5: Gain as a function of the reconstructed charge for a typical readout channel. The black points show the measurements made with the light injection calibration system at different intensities. The red line shows the best fit of a function as the one in equation 3.10, which is used as the conversion function. From [2].

The parameter $\alpha_0 = 1.053$/photoelectron is determined by imposing that, at low energy (such as the 2.2 MeV released in the capture of spallation neutrons on H) the total number of photoelectrons observed must equal the expected number of PMT hit, $\pi_{\text{hit}} = \overline{\mu} \cdot N_{\text{PMT}}$; where N_{PMT} is the number of PMTs in the detector and $\overline{\mu}$ is the mean number of photoelectrons in one PMT, which is calculated from Poisson statistics as $\overline{\mu} = -\ln(1 - N_{\text{hit}}/N_{\text{PMT}})$, with N_{hit} being the observed number of hit PMTs.

When the gain is computed using different light intensities (see figure 3.5), a nonlinear dependence of the gain on the reconstructed charge is discovered. The origin of the non-linearity lies in the non-linearity of the flash-ADC conversion and in the discretization of the waveform shape [155], and it affects especially the pedestal, which can be biased within ±1 ADC count. This causes a charge-dependent gain non-linearity which is more noticeable at low charges. In order to obtain a conversion function to photoelectrons $g_i(q_i)$ which corrects for this effect, the curve of measured gains as a function of the charge is fitted with the function

$$g_i(q_i) = \begin{cases} g_{0,i} + l_i \cdot (q_i - c_i), & \text{if } q_i < c_i \\ g_{0,i}, & \text{if } q_i \geq c_i \end{cases} \tag{3.10}$$

where $g_{0,i}$, l_i and c_i are parameters to be determined by the fit.

The gain curve changes upon power cycling the readout electronics (which is occasionally needed due to power glitches in the laboratory or to fix readout channels), which causes its time dependence, $g_i(q_i, t)$.

3.3 Vertex reconstruction

The reconstruction of the event vertex (either the positron energy deposition or the radiative capture of the neutron) in the Inner Detector is used in the $\overline{\nu}_e$ selection (reviewed in next chapter) to correlate the prompt and delayed triggers, but it is also used to correct the event energy for the variations due to its position within the detector as explained in next section.

The Double Chooz vertex reconstruction algorithm is called `RecoBAMA`. For a point-like event defined by the set $\boldsymbol{X} = (x_0, y_0, z_0, t_0, \Phi)$, where (x_0, y_0, z_0) are the spatial coordinates within the detector, t_0 is the event time and Φ is the light intensity per solid angle (expressed in photons/steradian); the `RecoBAMA` vertex reconstruction is given by the $\boldsymbol{X}$ which maximizes the following likelihood:

$$\mathcal{L}(\boldsymbol{X}) = \prod_{\substack{i \\ (pe_i=0)}} f_{\mathrm{pe}}(0;\mu_i) \prod_{\substack{i \\ (pe_i>0)}} f_{\mathrm{pe}}(pe_i;\mu_i) f_{\mathrm{t}}(t_i;\tau_i,\mu_i), \tag{3.11}$$

where $f_{\mathrm{pe}}(pe_i;\mu_i)$ is the probability to measure the observed number of photoelectrons pe_i in the i-th PMT when the expected number is μ_i; and $f_{\mathrm{t}}(t_i;\tau_i,\mu_i)$ is the probability to measure the event at time t_i (corrected for the channel offset) when the expected time is τ_i for the expected μ_i photoelectrons. The expected number of photoelectrons and time in the i-th PMT are predicted as

$$\mu_i = \Phi\, \epsilon_i\, \Omega_i\, A_i \tag{3.12a}$$

$$\tau_i = t_0 + \frac{r_i}{c_n} \tag{3.12b}$$

where ϵ_i is the PMT quantum efficiency, Ω_i is the solid angle subtended by the PMT for a vertex located at a distance r_i, A_i is the light transmission amplitude (which includes the effects of the light attenuation as it traverses detector and of the dependence on the angle of incidence, tuned using the radioactive source data); and c_n is the effective speed of light in the Inner Detector. The likelihood takes into account the possibility that some PMTs do not receive a photon by dividing the product into the PMTs which have not been hit ($pe_i = 0$) and therefore do not have time information, and those which have been hit ($pe_i > 0$). The probability functions f_{pe} and f_{t}, and the effective speed of light c_n are derived from the calibration data taken with the laser source.

In practice, instead of finding the $\boldsymbol{X}$ which maximizes equation 3.11, the negative log-likelihood function

$$F_V(\boldsymbol{X}) = -\ln \mathcal{L}(\boldsymbol{X}) \tag{3.13}$$

is minimized, which is equivalent. The value of F_V at the minimum is a measure of the goodness of the reconstruction and can be used to discriminate signal from background (explained in section 4.1.3.1).

3.4 Uniformity calibration

The Inner Detector response is not completely homogeneous due to geometrical effects (e.g. PMT positions and coverage) and the different media inside, so the total number of photoelectrons depends on the position of the event

within the detector. In order to correct for this spatial dependence, a uniformity calibration map is applied, $f_{\mathrm{u}}(\rho, z)$, which is a function of the reconstructed coordinates $\rho = \sqrt{x^2 + y^2}$ and z. The map is defined such that the corrected photoelectron number, $N_{\mathrm{pe}}(\rho, z) \cdot f_{\mathrm{u}}(\rho, z)$, equals the one at the center of the detector, $N_{\mathrm{pe}}(\rho = 0, z = 0)$, for the same physics event:

$$f_{\mathrm{u}}(\rho, z) = \left. \frac{N_{\mathrm{pe}}(\rho = 0, z = 0)}{N_{\mathrm{pe}}(\rho, z)} \right|_{n-\mathrm{H}}, \tag{3.14}$$

where the subscript $n - \mathrm{H}$ indicates that the map is built using the peak of the photoelectron spectrum produced by the 2.2 MeV γ-ray released in the capture of a neutron on a hydrogen nucleus. The DATA map is made using the spallation neutrons produced after the passage of a muon through the detector, and the MC map is made using simulated IBD neutrons (see figure 3.6). Both are abundant sources of neutrons which fully explore the scintillator volumes (Target and Gamma Catcher).

The data used to evaluate the systematic uncertainty on the visible energy because of a residual non-uniformity are different in the two $\overline{\nu}_e$ selections (one using the IBD neutrons captured on Gd, another using the IBD neutrons captured on H).

In the case of the Gd selection, the remaining non-uniformity in the Target between DATA and MC is taken as the systematic uncertainty. It is found to be 0.36%, estimated as the standard deviation of the relative differences across the Target volume between the DATA and MC maps built using the ~ 8 MeV energy released in the capture of a neutron on a gadolinium nucleus, after the visible energy calibration (equation 3.1) is completed.

In the case of the H selection, the systematic uncertainty cannot be assessed using gadolinium-based maps, since these do not represent the Gamma Catcher volume. Therefore, a different strategy is followed resorting to hydrogen-capture maps. However, the uniformity calibration maps are already made using neutrons captured on H. For the MC this is not a problem, since an statistically independent MC can be generated. For the DATA, an auxiliary uniformity calibration map is made just for the uncertainty estimation using half of the spallation neutron sample (selected to be evenly distributed in time). This auxiliary map is verified to agree with the official one which uses the full sample. As in the previous case, the systematic uncertainty is estimated from the residual non-uniformity after the full visible energy calibration is completed; but this time evaluated in the whole detector (Target and Gamma Catcher) using the other half of the spallation neutron sample as DATA (which have been corrected using the auxiliary map) and the independent MC sample. The standard deviation of the relative differences between the two results in a 0.25% uncertainty.

3.5 Energy conversion

The photoelectron-to-megaelectronvolt conversion factor, f_{MeV}, is determined from the position of the photopeak in the energy spectrum of neutron captures on hydrogen, resulting from the deployment of the ^{252}Cf source at the center of the detector (see figure 3.7). The energy spectrum is fitted with

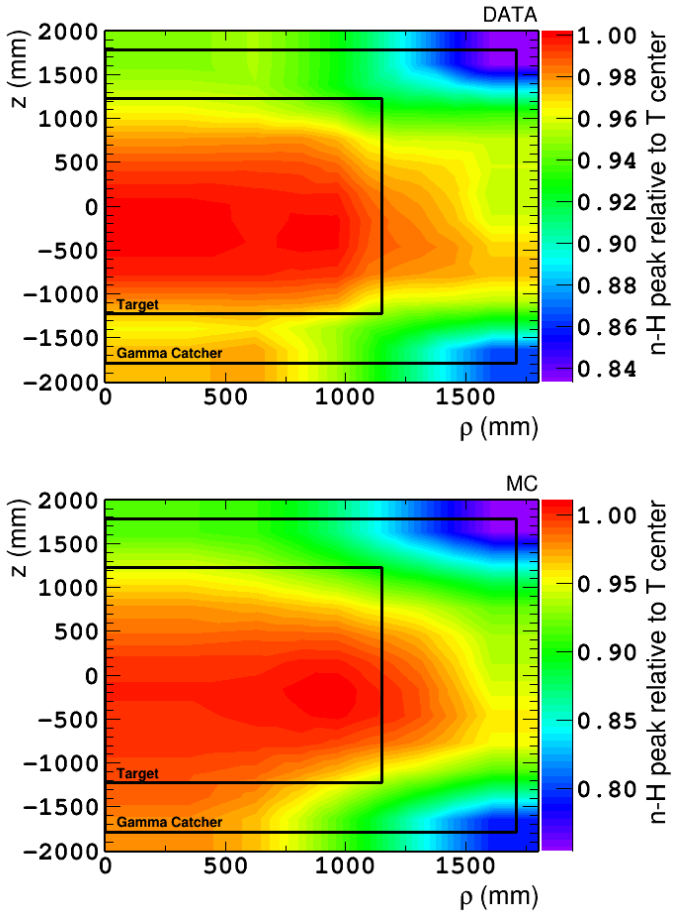


Figure 3.6: Uniformity calibration maps (corresponding to $f_{\mathrm{u}}^{-1}(\rho, z)$ of eq. 3.14) made using neutrons captured on hydrogen, for DATA (top) and MC (bottom), for the $\overline{\nu}_e$ Gd selection. A two-dimensional interpolation has been performed for a smooth application. From [156].

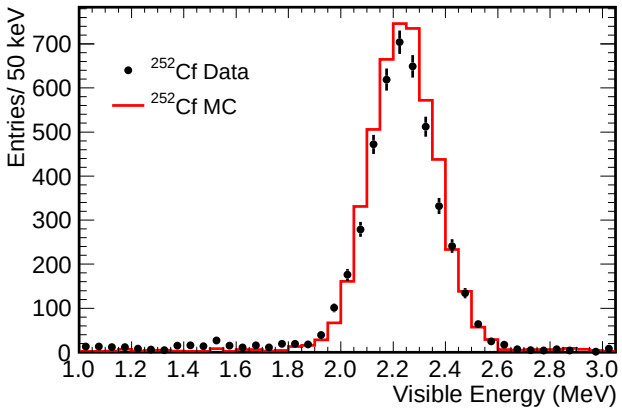


Figure 3.7: Energy spectrum of the capture of neutrons on hydrogen from a one-hour run taken with the ^{252}Cf source at the center of the detector, after application of the photoelectron-to-megaelectronvolt conversion factor. The DATA is shown as black circles and the MC simulation as a red line. From [2].

the function of equation 3.3, finding $f_{\mathrm{MeV}}^{\mathrm{DATA}} = 1\,\mathrm{MeV}/186.2$ photoelectrons and $f_{\mathrm{MeV}}^{\mathrm{MC}} = 1\,\mathrm{MeV}/186.6$ photoelectrons.

3.6 Stability calibration

The stability calibration is applied only to the DATA to correct their temporal dependence. The correction function is written

$$f_{\mathrm{s}}^{\mathrm{DATA}}(E_{\mathrm{vis},\odot}^{\mathrm{DATA}}, t) = \\ = 1 + \delta\alpha_0(t)\left(0.7866 - 0.07101 \cdot E_{\mathrm{vis},\odot}^{\mathrm{DATA}}\right) - 8.24 \cdot 10^{-6}(t - t_0), \qquad (3.15)$$

where $\delta\alpha_0(t)$ represents the measured variations in time of the parameter α_0 from equation 3.9, which is computed using the capture of spallation neutrons on hydrogen.

Examining the magnitude of the temporal dependence for different classes of events, an energy dependence is found: the low energy events are more affected due to the remaining gain non-linearity and the SPE inefficiency. Therefore, the correction function must incorporate an energy-dependent term based on the available energy definition so far, $E_{\mathrm{vis},\odot}^{\mathrm{DATA}} = N_{\mathrm{pe}}^{\mathrm{DATA}} \cdot f_{\mathrm{u}}^{\mathrm{DATA}}(\rho, z) \cdot f_{\mathrm{MeV}}^{\mathrm{DATA}}$. In order to cover the energy range of the $\overline{\nu}_e$ spectrum, three sources are used to devise this correction. The first source is ^{212}Po, which is a short-lived nucleus (half-life of $0.299\,\mu$s) which decays emitting an α with an energy of 8.95 MeV [38], but it is observed as a 1 MeV deposition due to the scintillator quenching, thus exploring the low-energy range. The ^{212}Po is one of the steps in the ^{232}Th radioactive chain which pervades the detector, and since it is created by the β^- decay of ^{212}Bi, it is easily selectable using a coincidence technique as the one for $\overline{\nu}_e$. The other two sources are the spallation neutron captures on H, releasing a 2.22 MeV γ-ray which provides a calibration point close to the $\overline{\nu}_e$

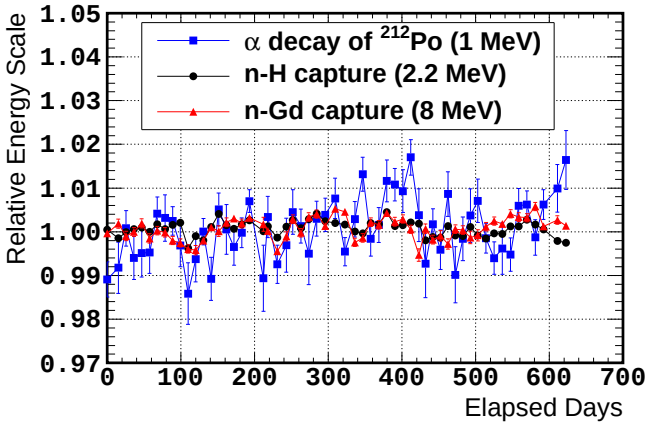


Figure 3.8: Ratio of the observed peak energy to the nominal energy as a function of time after stability calibration (eq. 3.15) for the three sources: ^{212}Po (blue squares) and the capture of spallation neutrons on H (black circles) and on Gd (red triangles). Error bars show the statistical uncertainty on the peak energy position. From [2].

mean energy, and on Gd, releasing several γ-rays with a total energy ~ 8 MeV which serve to study the high-energy behavior. The values of the parameters of the second term in equation 3.15 are determined to be those that minimize the time variation of the energies of the three sources.

The third term of the correction function of equation 3.15 corrects for the remaining temporal dependence present after the application of the second term. This persisting time variation of the detector response is attributed to the evolution of the scintillator light yield and the readout electronics. It is obtained by monitoring the position of the peak in the energy spectrum of spallation neutron captures on H in the Target and Gamma Catcher volumes, taking as reference the day $t_0 = 398$ corresponding to the deployment of the ^{252}Cf source which fixed the energy conversion factor (section 3.5).

The result of applying the stability calibration to the three sources described is shown in figure 3.8. The H capture exhibits the smaller dispersion (standard deviation of 0.17%) because the stability calibration is anchored to it. The ^{212}Po decay and Gd capture have standard deviations of 0.70% and 0.25%, respectively.

In the $\overline{\nu}_e$ selection based on Gd captures, the systematic uncertainty due to the residual temporal dependence is estimated using the dispersions of ^{212}Po decays and Gd captures of spallation neutrons (as it has been mentioned before, the H captures of spallation neutrons cannot be used since they are used to built the stability correction, and the IBD neutrons have a rate too low to be useful). A linear interpolation between the two sources gives the dispersion for all energies, which is further weighted by the $\overline{\nu}_e$ spectrum, resulting in a systematic uncertainty of 0.50%.

In the case of the $\overline{\nu}_e$ selection using H captures, in addition to the two sources used for the Gd selection (^{212}Po decay and Gd captures), an independent H

capture source is obtained from the IBD neutrons, so that a linear interpolation can be built using three points instead of two. Weighting by the $\overline{\nu}_e$ spectrum as before, this results in a systematic uncertainty of 0.33%.

3.7 Linearity calibration

The linearity calibration is applied to the energy of the MC simulation to correct for its non-linearity displayed relative to the DATA. The origin of this non-linearity is twofold: part is related to the reconstruction of the event charge, so it is referred to as charge non-linearity; and part is related to the generation of light in the scintillator by the event and is called light non-linearity correction. The charge non-linearity arises from biases in the modeling of the readout system and the charge reconstruction algorithm used for the simulation, thus affecting all Inner Detector triggers. The light non-linearity originates in the modeling of the scintillator light properties. Therefore, it depends on the particle being considered and the detector volume where it interacted, since Target and Gamma Catcher have different scintillator mixtures.

For the abovesaid reasons, the calibration of the two non-linearities is done differently: the charge non-linearity is corrected for all events using a function $f_{\rm qnl}^{\rm MC}(E_{\rm vis,\odot}^{\rm MC})$, which is a function of the visible energy prior to this correction, $E_{\rm vis,\odot}^{\rm MC} = N_{\rm pe}^{\rm MC} \cdot f_{\rm u}^{\rm MC}(\rho, z) \cdot f_{\rm MeV}^{\rm MC}$. Regarding the light non-linearity, instead of including a correction in the visible energy definition (equation 3.1), it is corrected in the Rate+Shape analysis (section 3.8). The explanation of how the two corrections are obtained is given in the rest of this section.

Charge non-linearity

The charge-non-linearity correction function $f_{\rm qnl}^{\rm MC}(E_{\rm vis,\odot}^{\rm MC})$ is extracted analyzing the DATA of the neutron captures, either on H or on Gd, obtained from the deployment of the ^{252}Cf source at the center of the detector and the corresponding MC simulation. Even though the total energy released upon neutron capture on H (one γ-ray of 2.22 MeV) and on Gd (typically three or four γ-rays totaling $\sim$ 8 MeV [157, 158]) are different, the average energy per γ-ray ($\sim$ 2.2 MeV) is very similar. Therefore, the scintillator response for the two cases is almost the same and the DATA-MC discrepancy found between the two can be attributed to a non-linearity in the total charge obtention. The following function is used to correct for such discrepancy:

$$f_{\rm qnl}^{\rm MC}(E_{\rm vis,\odot}^{\rm MC}) = 0.0023 \cdot E_{\rm vis,\odot}^{\rm MC} + 0.9949. \tag{3.16}$$

The systematic uncertainty on this correction, $\delta f_{\rm qnl}^{\rm MC}(E_{\rm vis}^{\rm MC})$, is estimated from the DATA-MC residual non-linearity observed in the ^{252}Cf data deployed along the detector z axis, after the $f_{\rm qnl}^{\rm MC}(E_{\rm vis,\odot}^{\rm MC})$ correction has been applied. It can be parametrized as a a linear function,

$$\delta f_{\rm qnl}^{\rm MC}(E_{\rm vis}^{\rm MC}) = b_{\rm qnl} + c_{\rm qnl} \cdot E_{\rm vis}^{\rm MC}, \tag{3.17}$$

with $b_{\rm qnl} = (1.004 \pm 0.004)$ and $c_{\rm qnl} = (-0.0001 \pm 0.0006)\,{\rm MeV}^{-1}$, where the parameter uncertainties are given by their standard deviation found when all the z axis positions are used. The correlation between the $b_{\rm qnl}$ and $c_{\rm qnl}$ parameters

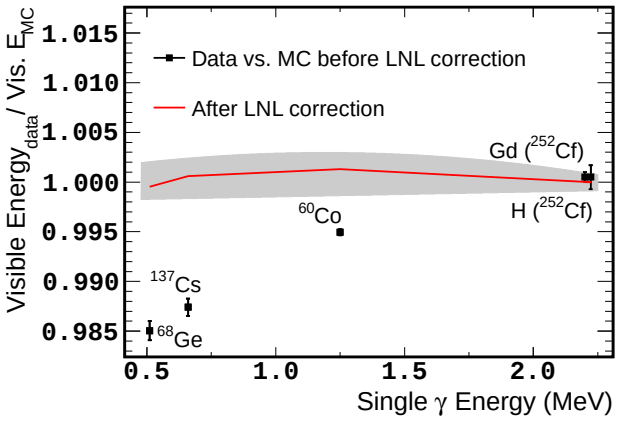


Figure 3.9: Ratio of the peak visible energy in the DATA to that of the MC simulation after the application of the charge-non-linearity correction (eq. 3.16), plotted as a function of the average γ-ray energy. The results correspond to deployments of the calibration sources (symbols indicated on the plot) at the center of the detector. The black squares show the results before applying the light-non-linearity correction, with the error bars showing the statistical uncertainty on the peak position; and the red line shows the results after the correction found for the Gd-based $\overline{\nu}_e$ selection, with the gray band showing the systematic uncertainty. From [2].

is -0.6 and is taken into account in the treatment of the energy scale uncertainty in the Rate+Shape analysis (section 3.8).

Light non-linearity

The calibration has only been implemented for the positron events (prompt triggers), as these are the ones used to build the energy spectrum from which θ_{13} is measured (explained in section 6.2). Moreover, since the correction for the light non-linearity depends on the scintillator properties, it will be different for the Gd-based $\overline{\nu}_e$ selection, in which practically all IBD interactions occur within the Target, and the H-based $\overline{\nu}_e$ selection, in which the IBD interactions happen in both Target and Gamma Catcher scintillators.

Figure 3.9 shows the performance of the visible energy of the MC simulation with respect to the DATA after the charge non-linearity correction has been applied. An energy non-linearity is clearly observed. The fact that the magnitude of this non-linearity scales monotonically as a function of the average single γ-ray energy rather than the total γ energy produced by the calibration source (see table 2.2) suggests that this effect is caused by the scintillator local response. However, no correction can be derived directly from figure 3.9 since it displays the light non-linearity corresponding to γ interactions instead of the positron ones.

Concerning the Gd-based $\overline{\nu}_e$ selection, in order to obtain a light-non-linearity correction function for the positrons, the parameters of the scintillator simulation are varied within the uncertainties of their laboratory measurements

[149, 128]. Specifically, the parameters varied are the Birks' constant of the scintillator quenching and the scintillator light yield. The variation of the latter alters the ratio of scintillation to Cherenkov light, which is suspected as a cause of the DATA-MC discrepancy. The calibration source data of figure 3.9 are used to identify which combinations of the MC parameters result in a good linearity of the energy scale with respect to the DATA. For these DATA-like combinations of MC parameters, a positron MC simulation is generated and its spectrum is compared to the one obtained with the default MC parameters. The non-linearity observed in the DATA-like simulation with respect to the default MC is then fitted with a function like

$$\delta f_{\rm lnl}^{\rm MC}(E_{\rm vis}^{\rm MC}) = \frac{a_{\rm lnl}}{E_{\rm vis}^{\rm MC}} + b_{\rm lnl}, \tag{3.18}$$

where $a_{\rm lnl}$ and $b_{\rm lnl}$ are free parameters. Analyzing the fit results for the DATA-like combinations, the parameters are found to be $a_{\rm lnl} = (-0.027 \pm 0.006)\,{\rm MeV}$ and $b_{\rm lnl} = (1.008 \pm 0.003)$; where the central values and uncertainties correspond to the mean and the standard deviation of the best-fit parameter distributions, respectively. These values, together with their correlation (-0.81), determine the light-non-linearity correction function which is implemented in the energy scale model of the Rate+Shape analysis (discussed in the next section 3.8).

The obtention of a light-non-linearity correction for the H-based $\overline{\nu}_e$ selection is more complex than the Gd-based, since the IBD interactions can happen in the Target and Gamma Catcher scintillators. Therefore, an alternative strategy is followed. Instead of trying to find a correction function, a generic correction with the functional form of equation 3.18 is implemented in the Rate+Shape analysis, and the $a_{\rm lnl}$ and $b_{\rm lnl}$ parameters are left to be determined during the θ_{13} analysis. At the beginning, $a_{\rm lnl}$ and $b_{\rm lnl}$ are initialized to 0 and 1, respectively (i.e. no correction). The uncertainties on these parameters are estimated from an educated guess relying on the results found for the Target scintillator. Each uncertainty is defined so that it covers a shift in the value of the associated parameter of magnitude twice the central value plus one standard deviation found for the Gd-based $\overline{\nu}_e$ selection, in order to allow for possible deviations in the Gamma Catcher scintillator. Hence, the uncertainties on $a_{\rm lnl}$ and $b_{\rm lnl}$ amount to $2 \times (0.027 + 0.006)\,{\rm MeV}$ and $2 \times (0.008 + 0.003)$, respectively. The validity of this approach was tested using fake MC data and running the Rate+Shape analysis of the Gd-based selection in this way. In all cases the Rate+Shape analysis was able to recover the nominal $a_{\rm lnl}$ and $b_{\rm lnl}$ parameters without a significant degradation of the precision on θ_{13}.

After the completion of the visible energy calibration (equation 3.1), the agreement between DATA and MC is remarkable, as figure 3.10 demonstrates. The energy resolution can be parametrized with a function

$$\frac{\sigma}{E_{\rm vis}} = \sqrt{\frac{p_{\rm a}^2}{E_{\rm vis}} + p_{\rm b}^2 + \frac{p_{\rm c}^2}{E_{\rm vis}^2}}, \tag{3.19}$$

where the $p_{\rm a}$, $p_{\rm b}$ and $p_{\rm c}$ are free parameters to be figured by a fit, which represent the contributions of statistical fluctuations, systematic uncertainties (e.g. non-uniformity) and electronic noise, respectively [10]. The best-fit results for both DATA and MC are shown in table 3.1, which confirm the high degree of similarity achieved in the energy scale of the two.

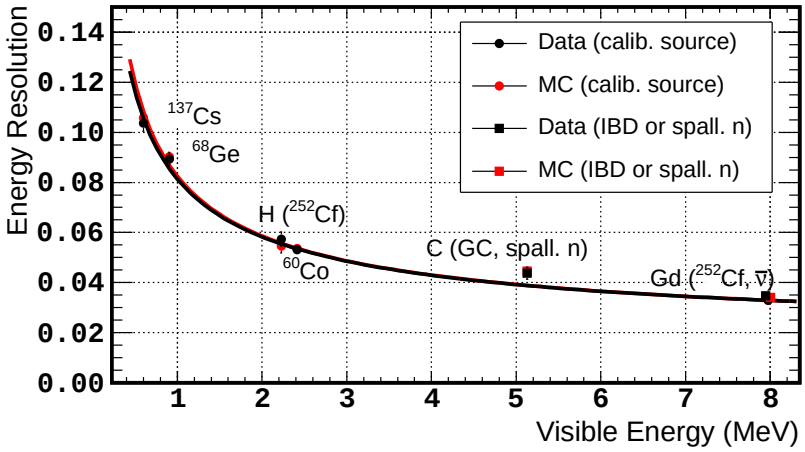


Figure 3.10: Energy resolution as a function of the peak visible energy for DATA (black) and MC (red) after the calibration. The circles correspond to data from calibration sources (specified with a label close to the data point) deployed at the center of the detector. The squares represent data from neutron captures from extended sources such as spallation neutrons or $\overline{\nu}_e$ IBD interactions. The solid lines show the best fit of the parametrization of the resolution (eq. 3.19). From [2].

Parameter	DATA	MC
p_{a} ($\mathrm{MeV}^{1/2}$)	0.077 ± 0.002	0.077 ± 0.002
p_{b}	0.018 ± 0.001	0.018 ± 0.001
p_{c} (MeV)	0.017 ± 0.011	0.024 ± 0.006

Table 3.1: Best-fit values of the parameters of eq. 3.19 for the data of fig. 3.10.

Parameter	Gd selection	H selection
b_{u}	1.0000 ± 0.0036	1.0000 ± 0.0025
b_{s}	1.0000 ± 0.0050	1.0000 ± 0.0033
b_{qnl}	1.004 ± 0.004	1.004 ± 0.004
$c_{\mathrm{qnl}}\,(\mathrm{MeV}^{-1})$	-0.0001 ± 0.0006	-0.0001 ± 0.0006
$a_{\mathrm{lnl}}\,(\mathrm{MeV})$	-0.027 ± 0.006	0.000 ± 0.066
b_{lnl}	1.008 ± 0.003	1.000 ± 0.021

Table 3.2: Summary of initialization values and uncertainties of the parameters of eq. 3.21 for the two $\overline{\nu}_e$ selections.

3.8 Energy scale in the Rate+Shape analysis

Even after all the corrections described in this chapter are applied, a residual bias in the energy of the MC simulation might exist; which must be taken into account when comparing the DATA and MC positron spectra to measure θ_{13} with the Rate+Shape analysis (section 6.2). This is achieved by treating the visible energy of the MC simulation not as a fixed variable but as a floating variable which can be adjusted during the θ_{13} analysis, constrained by its uncertainty. In order to do so, the visible energy used in the oscillation analysis, $E^{\mathrm{MC}}_{\mathrm{vis,fit}}$, is written as

$$E^{\mathrm{MC}}_{\mathrm{vis,fit}} = E^{\mathrm{MC}}_{\mathrm{vis}} \cdot \delta f^{\mathrm{MC}}_{\mathrm{u}} \cdot \delta f^{\mathrm{MC}}_{\mathrm{s}} \cdot \delta f^{\mathrm{MC}}_{\mathrm{qnl}}(E^{\mathrm{MC}}_{\mathrm{vis}}) \cdot \delta f^{\mathrm{MC}}_{\mathrm{lnl}}(E^{\mathrm{MC}}_{\mathrm{vis}}) \tag{3.20}$$

where $E^{\mathrm{MC}}_{\mathrm{vis}}$ is the visible energy of equation 3.1. $\delta f^{\mathrm{MC}}_{\mathrm{u}}$ is the residual correction to the MC energy due to the non-uniformity with respect to the DATA. Since the uniformity calibration (section 3.4) should have already corrected for this effect and no dependence on the visible energy is expected, $\delta f^{\mathrm{MC}}_{\mathrm{u}}$ can be written as $\delta f^{\mathrm{MC}}_{\mathrm{u}} = b_{\mathrm{u}}$, with b_{u} being a floating parameter which is initialized in the fit to 1.0 (no correction). The uncertainty on the b_{u} parameter is the residual relative non-uniformity between DATA and MC (see table 3.2). Analogously, $\delta f^{\mathrm{MC}}_{\mathrm{s}}$ is the residual correction to the MC energy due to the non-stability of the DATA in time. As before, the temporal dependence across all the energy range is thought to be corrected by the stability calibration (section 3.6), so $\delta f^{\mathrm{MC}}_{\mathrm{s}} = b_{\mathrm{s}}$, where b_{s} is a floating parameter originally set to unity (no correction), with an uncertainty given by the residual non-stability of the DATA energy (see table 3.2). $\delta f^{\mathrm{MC}}_{\mathrm{qnl}}(E^{\mathrm{MC}}_{\mathrm{vis}})$ is the parametrization of the residual charge non-linearity (equation 3.17) already presented in section 3.7. $\delta f^{\mathrm{MC}}_{\mathrm{lnl}}(E^{\mathrm{MC}}_{\mathrm{vis}})$ is the implementation of the light non-linearity correction (equation 3.18), which had been postponed until now.

Rewriting equation 3.20 explicitly in terms of the parameters

$$E^{\mathrm{MC}}_{\mathrm{vis,fit}} = E^{\mathrm{MC}}_{\mathrm{vis}} \cdot b_{\mathrm{u}} \cdot b_{\mathrm{s}} \cdot \left(b_{\mathrm{qnl}} + c_{\mathrm{qnl}} \cdot E^{\mathrm{MC}}_{\mathrm{vis}}\right) \cdot \left(\frac{a_{\mathrm{lnl}}}{E^{\mathrm{MC}}_{\mathrm{vis}}} + b_{\mathrm{lnl}}\right) \tag{3.21}$$

and rearranging them, the visible energy in the fit can be modelized as a quadratic polynomial

$$E^{\mathrm{MC}}_{\mathrm{vis,fit}} = a' + b' \cdot E^{\mathrm{MC}}_{\mathrm{vis}} + c' \cdot \left(E^{\mathrm{MC}}_{\mathrm{vis}}\right)^2 \tag{3.22a}$$

	Gd selection		H selection	
Coefficient	Value	Uncertainty	Value	Uncertainty
a' (MeV)	-0.027	0.006	0.000	0.067
b'	1.012	0.008	1.004	0.022
c' (MeV^{-1})	-0.0001	0.0006	-0.0001	0.0006

Table 3.3: Initialization values and uncertainties of the coefficients of the quadratic polynomial (eq. 3.22) which modelizes the visible energy of the MC simulation in the θ_{13} Rate+Shape fit for the two $\overline{\nu}_e$ selections. They are computed from the values of tab. 3.2, using the definitions of eq. 3.22.

where the three coefficients are defined as

$$a' \equiv b_{\mathrm{u}} \cdot b_{\mathrm{s}} \cdot b_{\mathrm{qnl}} \cdot a_{\mathrm{lnl}} \tag{3.22b}$$

$$b' \equiv b_{\mathrm{u}} \cdot b_{\mathrm{s}} \cdot (b_{\mathrm{qnl}} \cdot b_{\mathrm{lnl}} + c_{\mathrm{qnl}} \cdot a_{\mathrm{lnl}}) \tag{3.22c}$$

$$c' \equiv b_{\mathrm{u}} \cdot b_{\mathrm{s}} \cdot c_{\mathrm{qnl}} \cdot b_{\mathrm{lnl}} \tag{3.22d}$$

Table 3.3 shows the initialization values and uncertainties of the three coefficients, which are computed from those of the parameters of table 3.2.

As it was stated before, some of the parameters of equation 3.21 are correlated. Consequently, the three coefficients of equation 3.22a will be correlated too. The correlation matrix ρ for them is determined to be

$$\rho_{\mathrm{Gd}} = \begin{array}{c} \\ a' \\ b' \\ c' \end{array} \begin{array}{c} \begin{array}{ccc} a' & b' & c' \end{array} \\ \begin{pmatrix} 1 & -0.30 & 7.1 \cdot 10^{-3} \\ -0.30 & 1 & -0.29 \\ 7.1 \cdot 10^{-3} & -0.29 & 1 \end{pmatrix} \end{array} \tag{3.23a}$$

$$\rho_{\mathrm{H}} = \begin{array}{c} \\ a' \\ b' \\ c' \end{array} \begin{array}{c} \begin{array}{ccc} a' & b' & c' \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -0.10 \\ 0 & -0.10 & 1 \end{pmatrix} \end{array} \tag{3.23b}$$

where the subscript refers to the Gd or H selections. Further details about how the visible energy model (equation 3.22a) and the correlation matrix are included in the Rate+Shape statistical analysis are given in section 6.2.

Chapter 4

Antineutrino selections

The Double Chooz collaboration has developed two selections of the electron antineutrinos which interact in the detector via inverse β-decay, $\overline{\nu}_e + p \rightarrow e^+ + n$, depending on which nucleus captures the neutron produced, gadolinium or hydrogen (other nuclei have negligible neutron capture rates). The selections presented here were published in [2, 159], and they correspond to an endeavor to optimize previous selections using Gd captures [1] and H captures [72]. The two selections have been kept as similar as possible in order to facilitate the combination of the results on θ_{13}.

The dichotomy posed by the neutron capture nucleus is reflected in the chapter layout: it is divided into two long sections, the first covering the Gd selection (section 4.1) and the second devoted to the H selection (section 4.2). Within each selection, a similar structure is followed: firstly, the data analyzed are introduced. Secondly, the trigger preselection before searching for $\overline{\nu}_e$ is described (sections 4.1.1 and 4.2.1). Thirdly, the cuts targeted to find the trigger coincidence which is characteristic of the IBD process are explained (sections 4.1.2 and 4.2.2). Finally, the backgrounds hiding in the IBD selection are described together with the specific vetoes developed against them, and the estimation of the remaining contamination (sections 4.1.3 and 4.2.3). Figure 4.1 shows a scheme with both antineutrino selections to be used as a guide for the chapter.

4.1 Gadolinium selection

The detector design is optimized for the IBD neutrons to be captured on Gd. The Gd is only present in the scintillator of the innermost volume, the Target, and it is surrounded by an unloaded scintillator, the Gamma Catcher, to improve the calorimetric response. In addition, the extremely high thermal neutron capture cross section of Gd offers a high signal detection efficiency. Finally, the energy released upon a radiative neutron capture on Gd is well above the typical energies of the natural radioactivity background, giving this selection the highest signal to background ratio.

The Far Detector data used for the analysis amounts to 12305 runs taken between April 13^{th} 2011 and January 15^{th} 2013, totaling 489.51 days of runtime. After subtraction of the dead-time introduced by the muon and OV vetoes (described in sections 4.1.1.1 and 4.1.3.1, respectively), a live-time of 467.90 days

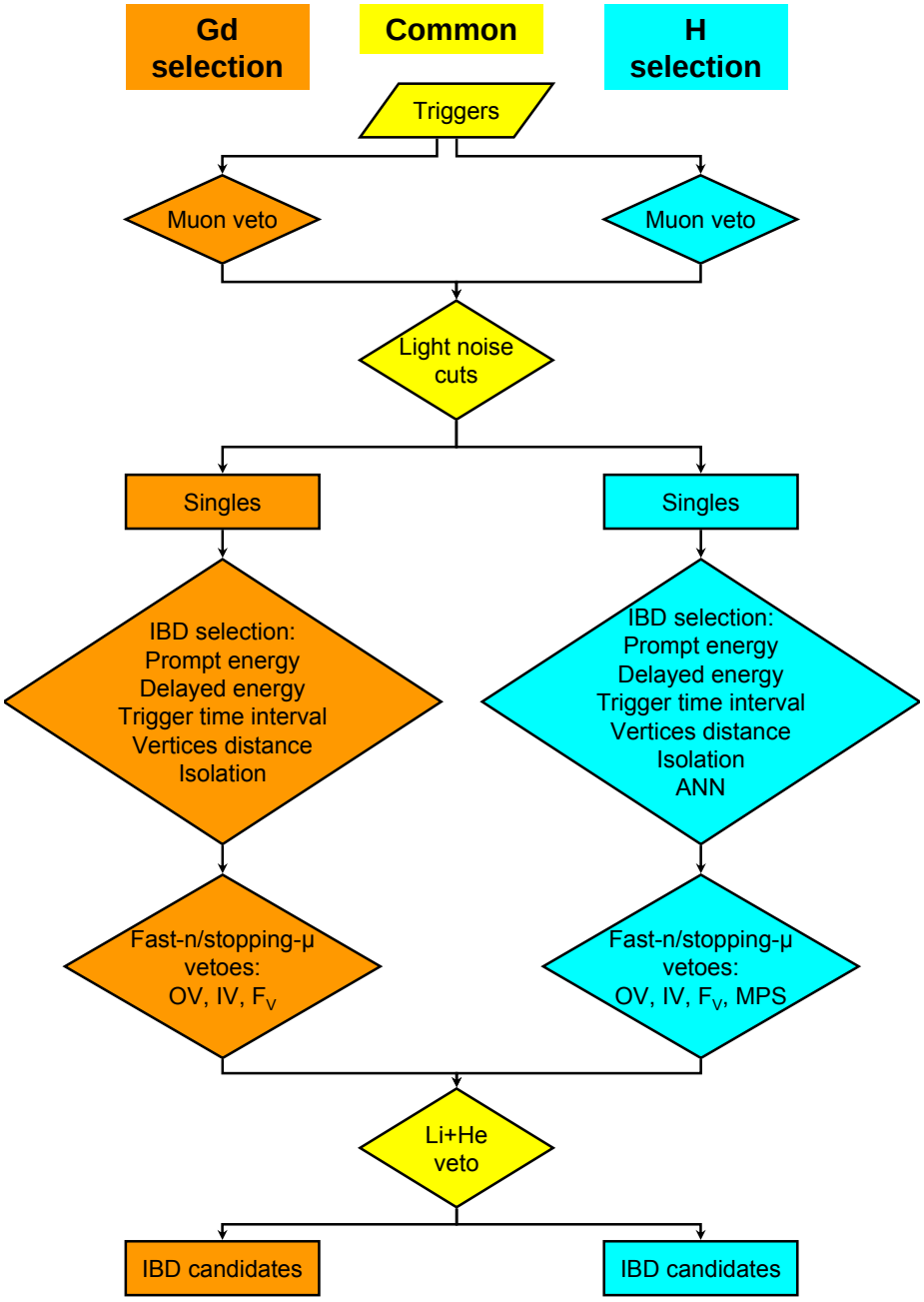


Figure 4.1: Antineutrino selection flowchart. Only the flow of accepted events is drawn. The central column shows the elements common to Gd and H selections. The leftmost (rightmost) column shows the elements which belong to the Gd (H) selection only.

Cut	Inefficiency (%)	Correction factor
Muon veto	$4.49 \pm (< 0.01)$	$0.9551 \pm (< 0.0001)$
Light noise	$0.01 \pm (< 0.01)$	$0.9999 \pm (< 0.0001)$
Isolation	$1.06 \pm (< 0.01)$	$0.9894 \pm (< 0.0001)$
OV veto	$0.06 \pm (< 0.01)$	$0.9994 \pm (< 0.0001)$
IV veto	0.04 ± 0.01	0.9997 ± 0.0001
F_V veto	0.06 ± 0.11	0.9994 ± 0.0011
Li+He veto	0.50 ± 0.02	0.9950 ± 0.0002

Table 4.1: Inefficiencies in the Gd-based $\overline{\nu}_e$ selection due to the the background-oriented cuts and the associated correction factors to the normalization of the MC $\overline{\nu}_e$ simulation. Since these cuts are not applied to the MC, the correction factor is simply the cut efficiency in the DATA (which is the complementary of the inefficiency).

is obtained, corresponding to an exposure of 66.5 GW-ton-years (reactor power $\times$ detector mass $\times$ live-time). This doubles approximately the statistics from the previous Gd-based θ_{13} measurement [1].

4.1.1 Single trigger selection

The IBD process provides a trigger coincidence which is a powerful discriminant against backgrounds. However, before searching for the coincidence of the two signals, a set of valid triggers (known as *singles*) must be assembled.

In order to keep this sample free of trigger uncertainty, the minimum energy of the singles is established at 0.4 MeV, where the trigger efficiency is 100% with negligible uncertainty. This sample must be cleaned from known backgrounds to minimize the risk of spurious coincidences faking the IBD signature. This is achieved by requiring the singles to pass both the muon veto and the light noise cuts described in what follows.

4.1.1.1 Muon veto

In spite of being located under a 300 m.w.e. overburden, the Far Detector is still subjected to a cosmic muon flux. Any event with a visible energy larger than 20 MeV in the Inner Detector or an energy larger than 16 MeV in the Inner Veto is tagged as a muon and rejected. The rate of those events is $45.25\,\mathrm{s}^{-1}$. Since the muons can produce spallation neutrons and cosmogenic isotopes, which might be mistaken for IBD signals, a 1 ms veto time is imposed after each muon in which all the triggers are discarded [1]. This veto acts as a detector dead-time, which in the current data set amounts to 21.6 days. This is equivalent to an inefficiency of 4.49% in the $\overline{\nu}_e$ selection with uncertainty $< 0.01\%$, which must be taken into account when comparing the observed DATA to the MC simulation (see table 4.1), which does not include muons.

[1] The effective veto time is longer than 1 ms since, if it affects one of the two triggers of a real IBD event, the surviving trigger will also be effectively rejected since the coincidence cannot be found. This can be regarded as if the veto time were $1\,\mathrm{ms} + \langle \Delta T \rangle$, where $\langle \Delta T \rangle$ is the mean prompt-delayed coincidence time interval.

4.1.1.2 Light noise cuts

The *light noise* is an unexpected background caused by the spontaneous emission of light from the Inner Detector PMT bases. Tests in the laboratory [160, 161, 162] have revealed that it is caused by the glowing of the epoxy used in the bases when it is subjected to heat and high voltage. The rate of this background is not stable (see figure 4.2) and has evolved from 28 events/s up to 65 events/s, becoming the most frequent cause of triggers. It is thought that the seasonal temperature rises in the detector favor the polarization of the epoxy molecules, enhancing corona discharges which produce the light.

Because the origin is localized in a PMT base, the light noise events tend to concentrate the light on few PMTs, and the arrival times of the light to the PMTs vary widely as the light spreads out. These characteristics make them to be very different from the events occurring in the scintillator volumes, which as a result of their centrality within the detector, produce a more uniform signal distribution in the PMTs.

A trigger is considered not to be light noise if it satisfies all the following cuts:

$$q_{\max}/q_{\text{tot}} < 0.12 \tag{4.1a}$$

$$\sigma_t < 36\,\text{ns or } \sigma_q < 464 - 8\sigma_t\,\text{charge units} \tag{4.1b}$$

$$Q_{\text{diff}} < 3 \cdot 10^4\,\text{charge units}, \tag{4.1c}$$

where $q_{\max}$ corresponds to the maximum charge recorded by a single PMT and q_{tot} is the total charge of the event. The cut in eq. 4.1a is equivalent to requesting that any PMT should not record more than 12% of the event charge.

The variables σ_t and σ_q are the standard deviations of the PMTs hit time and charge distributions, respectively. The cuts in eq. 4.1b ensure that any accepted event delivers the light homogeneously and simultaneously between PMTs. In figure 4.3 these two variables are plotted against each other, showing the two-dimensional cut which separates the two event populations, one with a small spread in the PMT hit times, corresponding to the events in the scintillator volumes, and another with a larger spread which corresponds to the light noise events.

The Q_{diff} variable is defined as

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

where q_i is the charge recorded by one of the N PMTs surrounding the PMT with the maximum charge at a distance less than 1 m. The Q_{diff} variable measures how the light is shared in the neighborhood of the maximum charge PMT (light noise events tend to distribute the light unevenly, causing a large Q_{diff} value as shown in fig. 4.4), and it is especially useful for rejecting light noise events at energies close to the one of the neutron capture on Gd (~ 8 MeV) and above.

The light noise cuts reject 77.12% of the triggers in the $(0.4, 20)$ MeV range which passed the muon veto. The visible energy spectrum of the rejected events (figure 4.5) shows that this background extends over the whole energy range, piling up at low energies. In the same figure it is seen how the expected features,

such as the low energy peaks of the natural radioactivity or the Gd peak emerge after the light noise cuts are applied. The bulk of the light noise is rejected by the cuts described above, and the remaining light noise events are rejected by the F_V cut described in section 4.1.3.1. Therefore, the presence of this background in the IBD sample used to measure θ_{13} is considered negligible. Finally, the light noise cuts cause an inefficiency in the $\overline{\nu}_e$ selection of $(0.0124 \pm 0.0008)\%$, which is quantified using the MC simulation since there is not such background in it. This inefficiency is used later to correct the MC normalization (see table 4.1).

4.1.2 Inverse beta-decay event selection

The singles selection just described yields a sample with a $13.22\,\mathrm{s}^{-1}$ valid trigger rate, in which the signature of an $\overline{\nu}_e$ interacting via IBD is sought: the coincidence of a *prompt* trigger given by the kinetic energy deposit and annihilation of the positron, and a *delayed* trigger given by the radiative capture of the neutron on a Gd nucleus a few microseconds later. Therefore, a pair of triggers is considered an *IBD candidate* if it satisfies the following cuts:

$$\text{Prompt visible energy: } 0.5 < E_{\text{prompt}} < 20\,\text{MeV}. \tag{4.2a}$$
$$\text{Delayed visible energy: } 4 < E_{\text{delayed}} < 10\,\text{MeV}. \tag{4.2b}$$
$$\text{Prompt-delayed time interval: } 0.5 < \Delta T < 150\,\mu\text{s}. \tag{4.2c}$$
$$\text{Prompt-delayed vertices distance: } \Delta R < 100\,\text{cm}. \tag{4.2d}$$
$$\text{No valid triggers in the } 200\,\mu\text{s before the prompt and only one (the delayed) in the } 600\,\mu\text{s after the prompt.} \tag{4.2e}$$

Although with the current statistics the $\overline{\nu}_e$ prompt spectrum is expected to finish at $\sim 8.5\,\text{MeV}$, the energy window is extended up to 20 MeV in eq. 4.2a to include purposefully background events of cosmogenic ^{9}Li/^{8}He (which have a maximum Q-value of 13.6 MeV [38]), fast neutrons and stopping muons, so as to have a sideband free of signal which helps to constrain further the background in the Rate+Shape oscillation analysis (cf. figure 6.9).

The delayed energy distribution is shown in figure 4.6. The cut in eq. 4.2b was shifted from $(6, 12)\,\text{MeV}$ used in the previous Gd analysis [1] to $(4, 10)\,\text{MeV}$ for this analysis since the number of $\overline{\nu}_e$ above 10 MeV is negligible; whereas by decreasing the minimum energy to 4 MeV, the signal selection efficiency is increased and the normalization uncertainty induced by shifts in the energy scale is made insignificant, yet keeping the accidental coincidences below 0.1 events/day.

The time interval between triggers (eq. 4.2c) spans $(0.5, 150)\,\mu\text{s}$. In previous analyses [69, 1] the minimum time had to be kept at $2\,\mu\text{s}$ to reduce the stopping muon contamination (muon mean lifetime is $2.2\,\mu\text{s}$ [10]), but this is achieved by other means in this analysis (explained in section 4.1.3.1); so the minimum time can be lowered to integrate over the neutron thermalization regime, where the neutron capture probability increases rapidly (cf. first $\sim 10\,\mu\text{s}$ in figure 4.7) as the neutron loses its kinetic energy [2]. Thus, the sensitivity to the neutron thermalization model is reduced, so does the systematic uncertainty on the MC $\overline{\nu}_e$

[2] The neutron capture cross-section in Gd has a strong dependence on energy, as can be seen in figure 5.1.

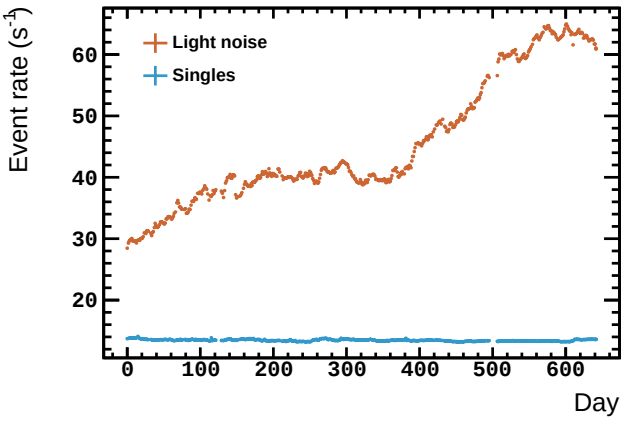


Figure 4.2: Light noise (orange) and singles (blue) rates. The light noise are the events rejected by the cuts in eq. 4.1. The trigger rate after the rejection of the light noise (singles) is stable. Edited from [163].

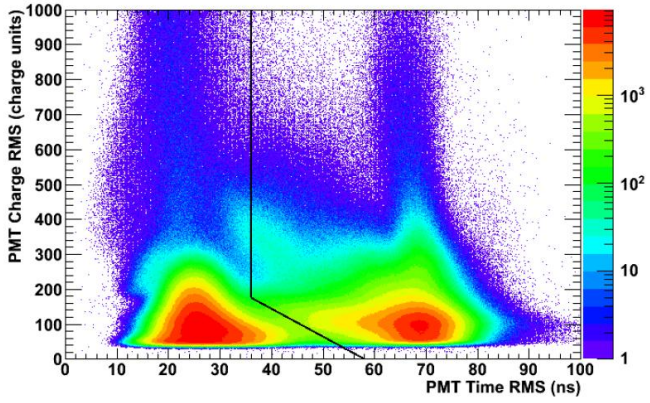


Figure 4.3: Standard deviation of the PMTs charge distribution versus standard deviation of the PMTs hit time distribution for a DATA subsample. The cuts in equations 4.1a and 4.1c have been applied. The accepted singles lie on the left side of the black line showing the cut in equation 4.1b. From [163].

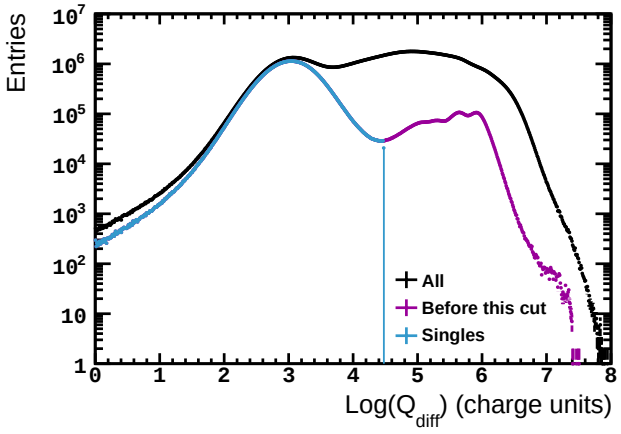


Figure 4.4: Distribution of the light noise cut variable Q_{diff}. The black histogram shows the DATA without light noise cuts applied. DATA distribution resulting from the application of the two cuts in equations 4.1a and 4.1b is shown in violet. The final singles distribution is shown in blue. Edited from [163].

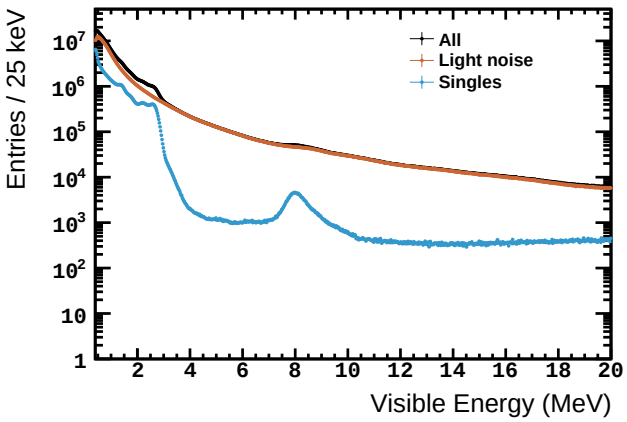


Figure 4.5: Visible energy spectra of the triggers passing the muon veto before (black) and after (blue) the light noise cuts in eq. 4.1 are applied. The spectrum of the rejected light noise events is shown in orange. Edited from [163].

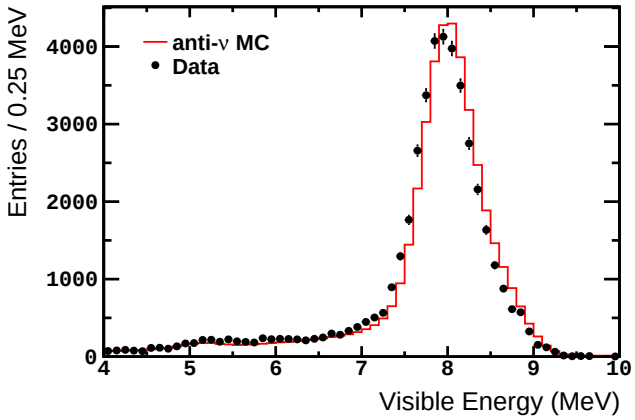


Figure 4.6: Delayed trigger visible energy spectra for IBD candidates in DATA (black circles) and MC (red histogram). MC normalized to the number of DATA candidates. Edited from [164].

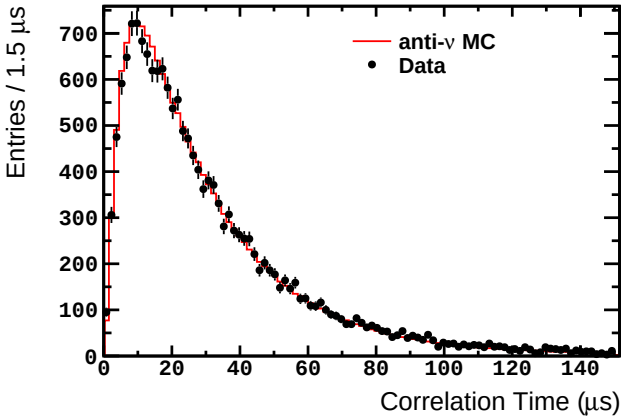


Figure 4.7: Prompt-delayed trigger time interval for IBD candidates in DATA (black circles) and MC (red histogram). MC normalized to the number of DATA candidates. Edited from [164].

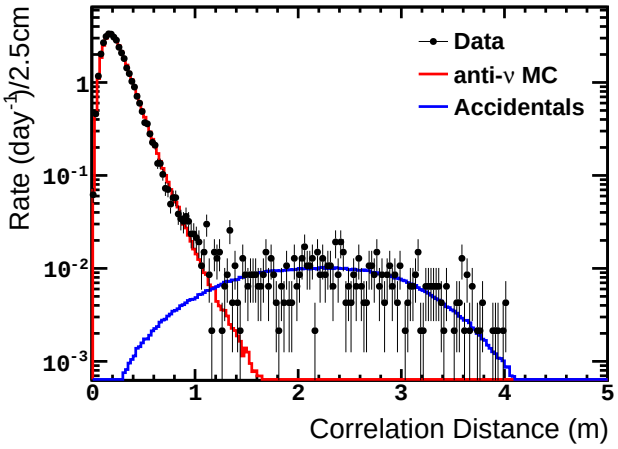


Figure 4.8: Prompt-delayed reconstructed vertices distance for IBD candidates in DATA (black circles), MC (red histogram) and the off-time accidental sample (blue histogram) without the distance cut applied. MC normalized to the live-time of DATA. The accidental sample is collected in 2000 off-time windows and rescaled accordingly. Edited from [2].

normalization. Regarding the maximum time, it has been optimized to increase the signal detection efficiency but still keeping the systematic uncertainty due to the neutron transport model low, since if the neutron is left to diffuse for a long time, the uncertainty caused by neutrons migrating into or out of the Target starts to become relevant (discussed in section 5.4.3).

The distance cut between `RecoBAMA`-reconstructed vertices (eq. 4.2d) is a novelty in the Gd selection, and it is introduced to reject accidental coincidences resulting from lowering the minimum delayed energy cut to 4 MeV (see figure 4.8).

The isolation cut of equation 4.2e removes events caused when multiple fast neutrons, which have been created by spallation in the matter surrounding the detector by the passage of a cosmic muon, enter the detector together and cause series of triggers which might be taken as false IBD coincidences: either by the association of a proton recoil and a neutron capture on Gd, or by two neutron captures occurring close in time. In addition, this cut ensures that the prompt-delayed coincidence is unambiguous, so no trigger can be simultaneously considered prompt and delayed.

The inefficiency of the isolation cut in the $\overline{\nu}_e$ selection is calculated as the probability of a valid trigger to happen within the 800 μs isolation window opened around the prompt trigger, and it equals 1.06% with $< 0.01\%$ uncertainty. This inefficiency only affects the DATA, so the MC normalization needs to be corrected to account for it (see table 4.1).

No inefficiency is computed for the time interval, vertices distance or delayed energy cuts, since they are also applied to the MC simulation. Instead, a dedicated analysis is performed on both DATA and MC to measure the efficiency of these cuts (see section 5.4.2), and correct the MC when it is needed.

4.1.3 Background vetoes and estimations

After the search for IBD-like coincidences using the cuts in equations 4.2, the surviving backgrounds are processes capable of mimicking the IBD interaction signature. Among these, the main causes are the accidental coincidence of two uncorrelated triggers or the coincidence of two correlated triggers caused by fast neutrons, stopping muons or some cosmogenic isotopes. All these sources are described in what follows, adding how their rates and prompt energy spectra can be estimated. Moreover, the ones which cause correlated triggers are originated by the interaction of cosmic muons with matter. For them, a series of vetoes are applied in order to further reduce the background contamination.

4.1.3.1 Fast neutrons and stopping muons

These two backgrounds are capable of causing trigger coincidences which resemble the IBD interaction. The rate of these fake IBD coincidences can be decreased by using the subdetectors which provide active background rejection (OV and IV) as vetoes or by using the trigger charge-hit time PMT information (F_V veto).

The *fast neutrons* are high energy neutrons entering the detector (mean kinetic energy $\sim 20\,\text{MeV}$ [66]), which have been produced by muon-induced spallation of the nuclei in the rock surrounding the detector. Since the muon is missed by the Inner Detector and the Inner Veto, it does not fire the muon veto. The prompt trigger is given by a neutron interacting with a proton in the Target or Gamma Catcher scintillator volumes, causing it to recoil. The delayed trigger is given by the radiative capture on Gd of one of the neutrons (not necessarily the one which caused the proton recoil) after it has thermalized.

The *stopping muons* are muons which stop inside the detector. If the energy deposited is low enough ($E_{\text{vis}} < 20\,\text{MeV}$), they do not fire the muon veto and are mistaken as prompt triggers. The delayed trigger is given by the Michel electron/positron created when the muon decays:

$$\mu^- \to e^- \overline{\nu}_e \nu_\mu, \qquad \mu^+ \to e^+ \nu_e \overline{\nu}_\mu.$$

OV veto

The Outer Veto is a muon-oriented subdetector (see section 2.3.2) which sits on top of the main detector and which extends beyond the area occupied by it; so it is capable of detecting impinging muons, including those which do not go through the detector but pass close by it. Any IBD candidate whose prompt trigger is coincident with an OV trigger is discarded (because of the different timing in the Inner Detector and the Outer veto, coincident is defined as "within 224 ns").

The dead-time introduced by this veto is estimated using a fixed rate trigger. Counting the number of those triggers which are coincident with an OV trigger gives a $(0.058 \pm 0.001)\%$ inefficiency in the $\overline{\nu}_e$ selection due to the OV veto.

IV veto

Any particle (fast neutrons, stopping muons, but also γ-rays) entering the detector from outside will have to cross the Inner Veto on its way to the Inner

Detector, so there is a chance that it will leave part of its energy in the Inner Veto, yet below the muon veto threshold. Therefore, any prompt trigger which shows a certain coincident activity in the Inner Veto is considered to be caused by an incoming background particle and rejected. All the following conditions must be fulfilled to execute this veto:

$$\text{IV PMT hit multiplicity} \geq 2 \tag{4.3a}$$

$$\text{Total charge in the IV} > 400\,\text{charge units} \; (\approx 0.2\,\text{MeV}) \tag{4.3b}$$

$$\text{ID-IV vertices distance} < 3.7\,\text{m} \tag{4.3c}$$

$$\text{ID-IV time interval: } -110 < \Delta T_{\text{ID}-\text{IV}} < -10\,\text{ns}. \tag{4.3d}$$

The cuts above are tuned to reduce the background while keeping the rejection of $\overline{\nu}_e$ due to accidental Inner Detector-Inner Veto coincidences low (see figure 4.9). This is predominantly achieved by requiring both time (eq. 4.3d) and spatial (eq. 4.3c) coincidences; where the reconstructed vertex in the Inner Veto is obtained from an artificial neural network which uses as inputs the relative numbers of photoelectrons in each Inner Veto PMT with respect to the total number of photoelectrons observed by the 78 Inner Veto PMTs.

The inefficiency in the $\overline{\nu}_e$ selection because of the IV veto is estimated to be $(0.035 \pm 0.014)\%$ from the fraction of accidental candidates selected with the off-time method (section 4.1.3.3) which are rejected by the IV veto.

No veto is applied on the delayed trigger since the deexcitation γ rays emitted upon the radiative neutron capture on Gd reach the Inner Veto with a non-negligible rate; which would lead to an excessive rejection of $\overline{\nu}_e$ interactions.

F_V veto

The chimney used to fill the Target and Gamma Catcher volumes contains liquid scintillator, so any muon entering through it will produce scintillation light which can cause a trigger. Moreover, it is not protected by the Inner Veto, and despite the Upper Outer Veto being designed to cover it, 72.4% of the DATA were taken without the Upper Outer Veto operative. Therefore, it is a natural entrance of stopping muons. In order to reject those events, a veto based on the delayed trigger F_V, the negative log-likelihood of the event reconstruction (equation 3.13) was devised,

$$E_{\text{delayed}} < 0.068 \exp(F_V/1.23), \tag{4.4}$$

exploiting the fact that F_V shows a high value for events with a hit pattern not compatible with a point-like source in the Target or the Gamma Catcher, such as the chimney stopping muons or the light noise events which survived the cuts in eqs. 4.1.

In figure 4.10 it is shown how the cut in equation 4.4 separates the IBD candidates (the left distribution) from the background events, which do not exhibit clustering around the Gd neutron capture energy. By looking at the vertex distribution in figure 4.11, it is observed how the rejected events pile up below the chimney (stopping muons) or in the center of the detector (mis-reconstructed light noise). This is further confirmed by examining the prompt-delayed time distribution of the rejected events (cf. figure 4.12), which shows a fast decay time constant which is characteristic of the decay of stopped muons or a prolonged light noise emission which gives two consecutive triggers.

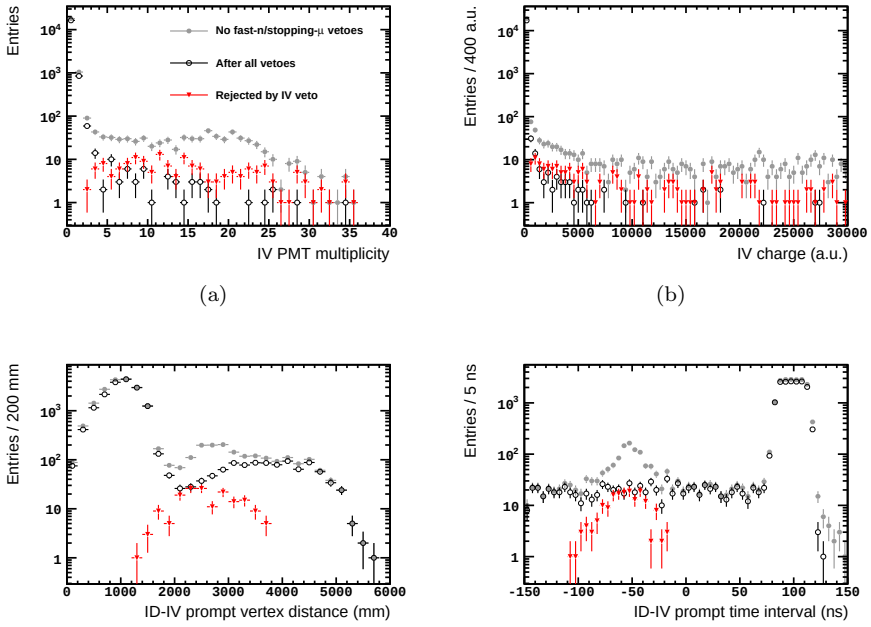


Figure 4.9: Distributions used in the definition of the IV veto. All show the data before (gray circles) and after (empty black circles) applying the OV, IV and F_V vetoes against the fast neutron and stopping muon background. The events rejected exclusively by the IV veto are shown as down-pointing red triangles. (a) shows the IV PMT hit multiplicity. The excess in the first two bins is caused by the $\overline{\nu}_e$ signal in coincidence with accidental hits, while the background spreads over higher PMT multiplicities. (b) shows the IV charge spectrum. Most of the $\overline{\nu}_e$ are piled up below 400 charge units, corresponding to coincidences with charge fluctuations in the IV PMTs or harmless very low energy events in the Inner Veto. On the other hand, the background extends up to 30000 units (where the muon veto threshold is set). (c) and (d) show the ID-IV prompt vertex distance and trigger time interval, respectively. The trigger time definition is charge-weighted, so events with low charge in the Inner Veto have positive time intervals corresponding to the Inner Detector trigger times (time offsets between Inner Detector and Inner Veto are not calibrated). In both of them, the maximum peak corresponds to the $\overline{\nu}_e$ signal while the secondary peak is caused by the background. It can be seen how the IV-vetoed events lie below those secondary peaks.

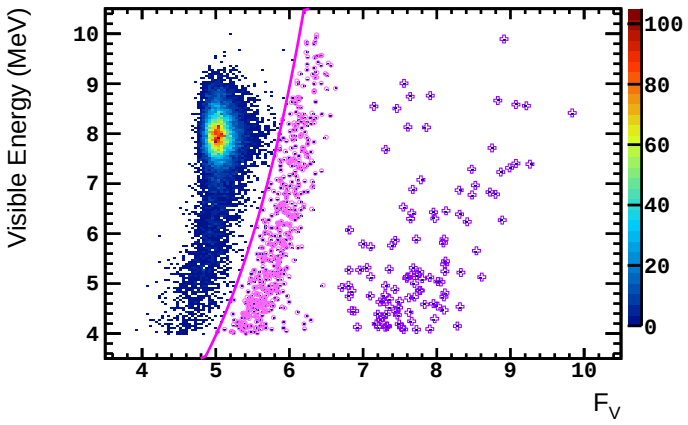


Figure 4.10: Two-dimensional histogram showing the correlation between the delayed trigger visible energy and the F_V value before the F_V veto is applied. The selected IBD candidates lie on the left of the pink line, which shows the F_V veto of equation 4.4. The rejected events form two distributions which can be roughly separated at $F_V \approx 6.7$: the stopping muon-like events are marked with a light pink circle, and the light noise-like events are marked with a violet cross.

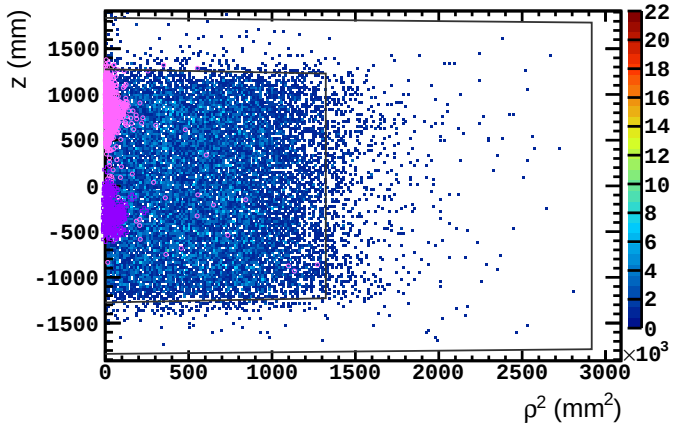


Figure 4.11: Delayed trigger reconstructed vertex distribution before the F_V veto is applied. The vetoed stopping muon-like events are marked with a light pink circle, and the vetoed light noise-like events are marked with a violet cross, as in figure 4.10. The inner and outer lines delimit the Target and Gamma Catcher vessels, respectively.

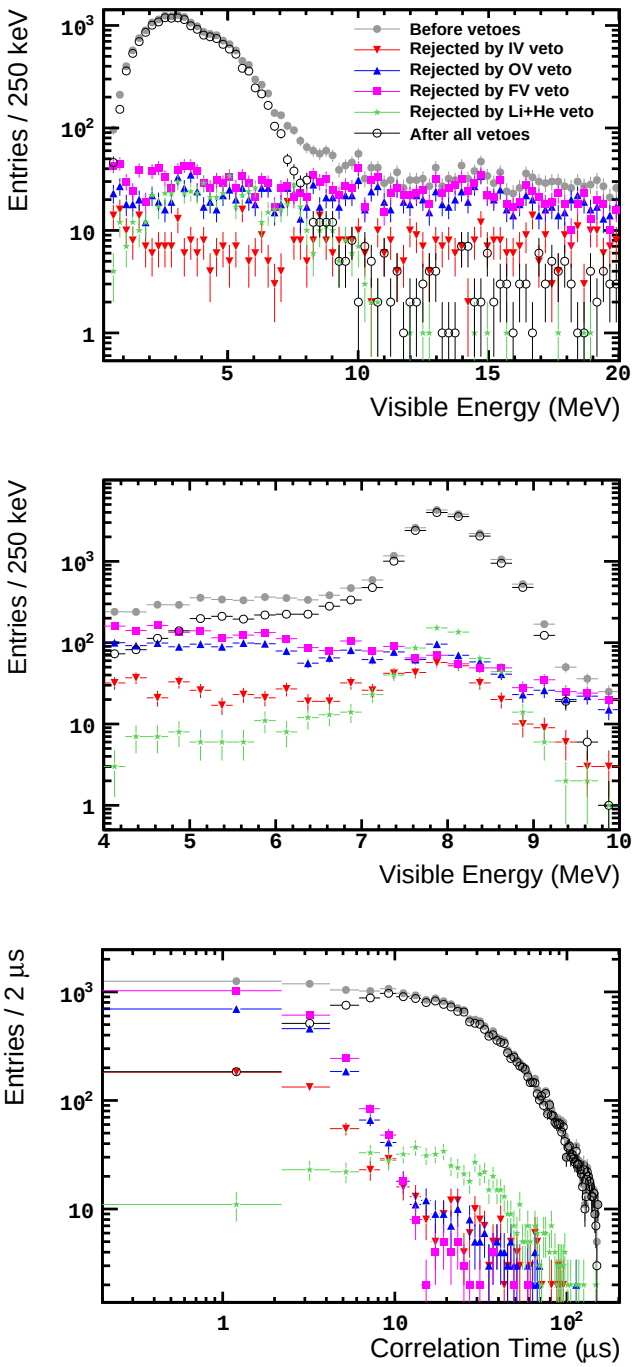


Figure 4.12: Prompt visible energy (top), delayed visible energy (middle) and prompt-delayed trigger time interval (bottom) distributions for IBD candidates before (gray circles) and after (empty black circles) applying the background vetoes described in sections 4.1.3.1 and 4.1.3.2. The events rejected individually by each veto are also shown: IV veto (down-pointing red triangles), OV veto (up-pointing blue triangles), F_V veto (pink squares) and Li+He veto (green stars). Error bars show the statistical uncertainty. Edited from [2].

The inefficiency in the $\overline{\nu}_e$ selection due to the F_V veto is estimated to be $(0.06 \pm 0.11)\%$ from the fraction of IBD candidates which pass all the IBD selection criteria except the F_V veto. Because this sample of rejected events also contains background events, the number of $\overline{\nu}_e$ lost is estimated after subtracting the accidental coincidences, the fast neutron events (which are measured between 11 and 20 MeV and then extrapolated to the delayed energy range assuming a flat spectrum), and the events which are distant from the cut displayed in figure 4.10 (and hence are not likely to be $\overline{\nu}_e$).

An analogous cut to eq. 4.4 is not applied to the prompt trigger so as not to bias the prompt energy spectrum which is used to measure θ_{13} in the Rate+Shape analysis.

The aforesaid OV, IV and F_V vetoes individually reject 62%, 24%, 71% of the IBD candidates above 12 MeV, respectively, where $\overline{\nu}_e$ are not expected and the fast neutron and stopping muon background is dominant. When the three vetoes act together, a 90% rejection is obtained. The combined use of the vetoes cause an overall inefficiency of $(0.15 \pm 0.11)\%$ in the $\overline{\nu}_e$ selection, which is taken into account when the DATA is compared to the MC simulation, where these vetoes are not applied.

Rate and shape estimation

The shape of the prompt energy spectrum of the remnant fast neutron and stopping muon contamination in the IBD sample is determined using the *IV-tagged events* (IBD candidates passing all the selection criteria but which are rejected by the IV veto of eq. 4.3; corresponding to the down-pointing red triangles in fig. 4.12). Since the IV veto is sensitive to both fast neutrons and stopping muons, the spectrum obtained is representative of these two backgrounds. A linear fit to the observed spectrum gives a slope of -0.02 ± 0.11 events/ MeV^2, which is consistent with a flat spectrum. This is further confirmed by the shapes of the OV-tagged events and the remaining IBD candidates beyond 12 MeV in the top plot of figure 4.12. Therefore, this background is modeled as flat in the Rate+Shape oscillation analysis.

The rate of the background events still present in the IBD sample is extrapolated assuming a flat spectrum from the measured rate of coincidences in the range $(20, 30)$ MeV, which are obtained modifying the prompt energy cut of equation 4.2a and raising to 30 MeV the muon veto of section 4.1.1.1. The estimated values are 0.604 ± 0.051 events/day for the reactor-on period and 0.529 ± 0.089 events/day for the reactors-off period due to a different configuration of the Outer Veto in the two periods.

4.1.3.2 Cosmogenic isotopes

The passage of cosmic muons through the detector creates unstable cosmogenic isotopes by nuclear spallation, mostly of ^{12}C nuclei. Among these radioisotopes, ^{9}Li and ^{8}He have a decay mode in which a neutron is emitted along with an electron (branching ratios 50.80% and 16%, respectively), causing a trigger coincidence which is indistinguishable from an IBD interaction. Moreover, the lifetimes of ^{9}Li and ^{8}He are 257.2 ms and 171.8 ms [38], respectively, so the 1 ms after-muon veto is not effective against them.

Li+He veto

In order to reduce the rate of spurious coincidences caused by ^{9}Li and ^{8}He, a likelihood was developed [165] to estimate the probability that an IBD-like event is the result of a cosmogenic isotope decay. The variables used as inputs are the distance between the muon track and the prompt vertex and the number of neutron captures either on H ($1.8 - 2.6\,$MeV) or on Gd ($4 - 10\,$MeV) following the muon in a 1 ms window. A high neutron multiplicity is characteristic of a showering muon, which is more likely to produce cosmogenic isotopes [166]. Other variables were considered but they showed non-significant improvement.

The probability density functions for the cosmogenic events needed to build the likelihood are obtained from the distributions of ^{12}B decays in DATA since they are confirmed to be very similar to the ^{9}Li/^{8}He ones, but the expected low rates of the latter make them less precise. ^{12}B is the most frequently produced short-lived (29.1 ms) cosmogenic isotope; and as ^{9}Li (^{12}C(n, n3p)^{9}Li), it has also a production mode requiring an initial neutron (^{12}C(n, p)^{12}B). Since ^{12}B does not emit a neutron when it β^- decays, a sample is collected using a selection analogous to the IBD one in which a delayed trigger is not required. In addition, the event must be within the 50 ms following a muon which has been well reconstructed [167].

A background sample (coincidences of a muon and prompt trigger not correlated to it) is collected using an off-time selection in which the muon is searched in multiple windows opened a minimum of 10 s before the prompt trigger. The distributions of this off-time sample are used to build the probability density functions associated to the background, and are also subtracted from the on-time ^{12}B distributions in order to make them background-free.

The likelihood is computed for all the muons preceding the prompt trigger in a 700 ms window, and the maximum value is assigned to the event (see figure 4.13). The Li+He veto is set at a likelihood value of 0.4, corresponding to the minimum value in which it is expected that more ^{9}Li/^{8}He events are vetoed than $\overline{\nu}_e$ ones. Any event with a likelihood greater than 0.4 is rejected. In figure 4.12, the prompt energy, delayed energy and prompt-delayed time interval distributions of the rejected events are shown, which have the expected features of a βn emitter isotope.

The inefficiency in the $\overline{\nu}_e$ selection as a result of the Li+He veto is computed from the number of IBD candidates paired with an off-time muon which are vetoed, and it amounts to $(0.504 \pm 0.018)\%$. Since the $\overline{\nu}_e$ simulation does not include backgrounds, this inefficiency must be considered when computing the MC normalization (see table 4.1) .

Rate and shape measurement

The ^{9}Li/^{8}He spectrum shape is measured directly from the DATA using the IBD candidates rejected by the Li+He veto. In order to boost statistics, the neutron is allowed to capture either on Gd or on H [3]. A background sample consisting mostly of $\overline{\nu}_e$ in accidental coincidence with a muon is obtained using an equivalent selection, except the muon is sought in the aforementioned off-

[3] In this context, the H capture is defined as a delayed trigger with a visible energy in the range $(1.8, 2.6)$ MeV, a prompt-delayed time interval within $(0.5, 600)\,\mu$s and a prompt-delayed vertex distance smaller than 60 cm.

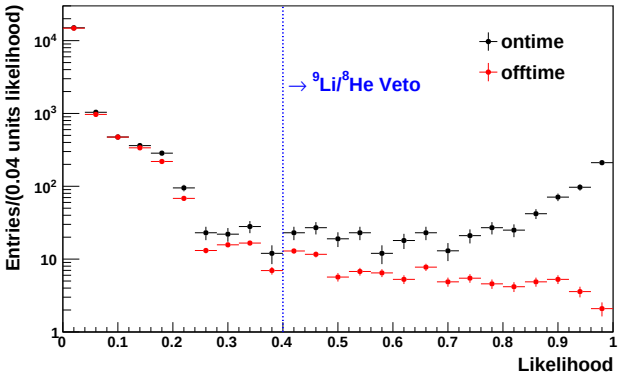


Figure 4.13: Distribution of the highest likelihood found for muon-prompt trigger pairs, for the on-time (black) and off-time (red) samples. The excess seen in the on-time distribution shows the presence of ^{9}Li/^{8}He events. From [165].

time windows, and then subtracted. The resulting spectrum is shown in figure 4.14. In the same figure the MC simulation of the ^{9}Li decay is shown as a reference. It implements the β^- decay to ^{9}Be, which then breaks up into two α and one neutron, either directly or via ^{5}He or ^{8}Be intermediate states. The uncertainty in the spectrum is taken from a covariance matrix obtained from varying the branching ratios of the possible processes.

The rate of the ^{9}Li/^{8}He events in the IBD sample is estimated from a fit to the distribution of time intervals between the prompt trigger and a preceding muon. The presence of ^{9}Li/^{8}He decays would manifest as an excess at short times which decreases exponentially. In order to gather data to constrain the accidental coincidences of a muon and an uncorrelated trigger, time intervals up to 20 s are allowed. Two rate estimations are made: an upper limit and lower limit. The rates are measured without applying the Li+He veto in a first instance, and then the vetoed events are subtracted afterwards.

To obtain the upper limit, the muons are sorted by the energy deposited in the Inner Detector, since the probability to produce cosmogenic isotopes scales with it. With the exception of the sample with a muon energy $E_\mu > 600\,\mathrm{MeV}^*$ [4], all the other samples require a cut on the distance of the prompt trigger to the muon track, $\Delta R_\mu < 75$ cm, to reduce the rate of the accidental coincidences and reach enough purity to get a precise result. The inefficiency introduced by the distance cut is estimated and corrected for each energy range from the geometrical detector acceptance and the lateral profile describing the distance of the cosmogenic candidate prompt vertex to the muon track. The acceptance depends on the deposited energy through the impact parameter and is derived from the data using uncorrelated muon-IBD prompt candidate pairs (since the uncorrelated IBD candidates are distributed homogeneously). The lateral profile is obtained from the pure $E_\mu > 600\,\mathrm{MeV}^*$ sample as figure 4.15 shows. The total estimated rate with this method is $2.20^{+0.35}_{-0.27}$ events/day.

[4] MeV* should be understood as an approximate unit since the energy scale cannot be extended accurately up to such energies because of the non-linearity induced by the flash-ADC saturation.

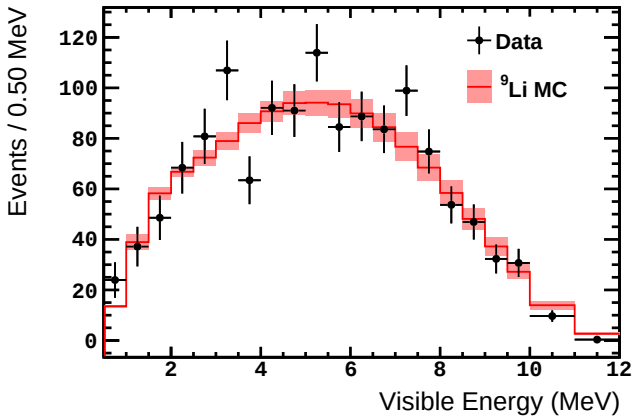


Figure 4.14: Prompt trigger energy spectrum of the ^{9}Li/^{8}He candidates selected by the Li+He veto. Points show the DATA after subtraction of the background (coincidences of a muon and an IBD event), which is measured with an off-time selection. Error bars show the statistical uncertainty. The predicted spectrum from the ^{9}Li MC simulation is overlaid (red histogram, where the band represents the uncertainty). The MC is normalized to the DATA. From [2].

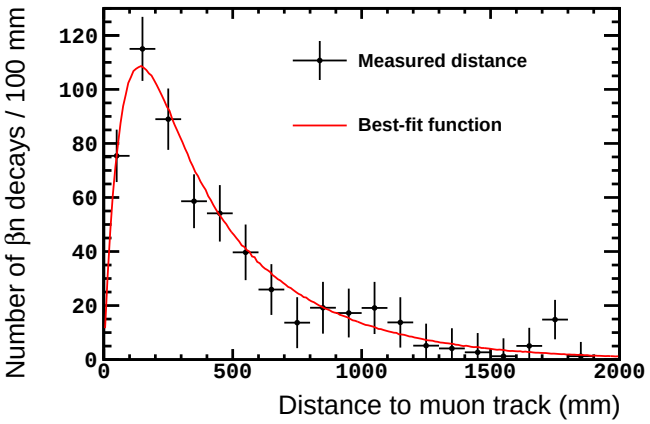


Figure 4.15: Lateral distance of the βn candidate prompt trigger vertex to the muon track. Points show the data for muons depositing more than 600 MeV* (error bars represent the statistical uncertainty). The best-fit to an exponential function modelizing the distribution along the distance convoluted with a Gaussian accounting for the spatial resolution of the muon track and prompt vertex is overlaid (red line). With the resolution of the prompt vertex fixed at 10 cm, the fit gives a "decay length" of 42 ± 4 cm and a muon track resolution of 15 ± 4 cm. From [2].

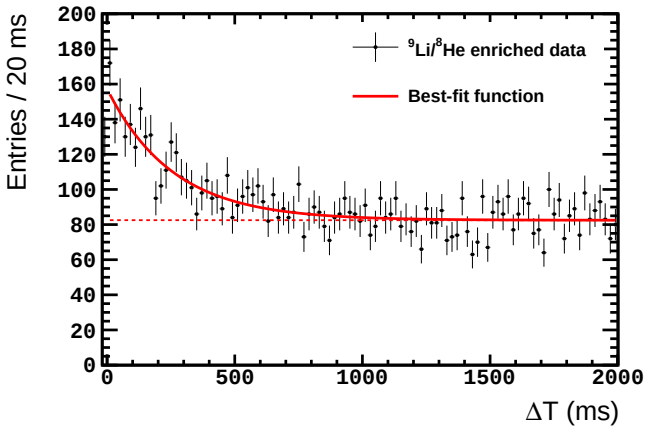


Figure 4.16: Muon-prompt trigger time interval distribution obtained with the selection used for the rate lower limit. Error bars show the statistical uncertainty. The excess at short times is interpreted as the cosmogenic isotope decays. The best-fit to an exponential function plus a constant term in the range $[10^{-3}, 20]$ s is shown (red solid line). The exponential describes the radioactive decay (the decay constant is fixed at 257 ms, the ^{9}Li lifetime) and the flat (red dashed line) component accounts for the accidental muon-prompt trigger coincidences. From [2].

The lower limit is computed from a muon sample which is selected by the cut $E_\mu > 300\,\mathrm{MeV}^*$ if the muon is followed by at least one neutron capture in the subsequent 1 ms, or $E_\mu > 500\,\mathrm{MeV}^*$ and $\Delta R_\mu < 75$ cm if there is not a neutron. The values of the energy cuts are targeted to select the maximum of ^{9}Li/^{8}He events but keeping the accidental coincidences low. In figure 4.16 the distribution of the muon-prompt trigger time interval is shown. The fit yields a rate of 2.05 ± 0.13 events/day.

When the upper and lower limits are combined, the resulting rate is $2.08^{+0.41}_{-0.15}$ events per day, where the uncertainty includes the systematic component which has been estimated varying the cut on the distance to the muon track, the value of the "decay length" parameter, the binning in the time interval distribution, considering a possible slope in the flat component and the impact of the ^{8}He fraction ($8 \pm 7\%$ from [166] after rescaling due to the different muon energy).

Finally, when the Li+He veto is applied, 1.12 ± 0.05 events/day are rejected. After subtraction of those events, the remaining rate of ^{9}Li/^{8}He decays is estimated to be $0.97^{+0.41}_{-0.16}$ events/day.

4.1.3.3 Accidental coincidences

This background is caused by the random coincidence of two triggers which are not physically correlated. The prompt trigger is mostly given by natural radioactivity originated in the PMTs, the detector materials or the surrounding rock; while the delayed trigger is mainly due to decays of ^{12}B (Q-value 13.4 MeV [38]), neutron captures on Gd or other nuclei which yield similar energies, and

proton recoils.

Because of its random origin, this background can be measured directly from the DATA by applying the same selection cuts in eq. 4.2 and background vetoes of sections 4.1.3.1 and 4.1.3.2, except for the prompt-delayed time interval cut (eq. 4.2c), which is substituted by:

$$(0.5\,\mu\text{s} + T_{\text{off}}) < \Delta T < (150\,\mu\text{s} + T_{\text{off}}) \tag{4.5}$$

where T_{off} is a time offset which must be chosen long enough so as to guarantee that no correlation exists between the two triggers. In addition, because the two triggers are widely separated in time, the isolation cut in equation 4.2e must be replaced by two cuts:

$$\begin{aligned} &\text{No valid triggers in the } 200\,\mu\text{s before} \\ &\text{and in the } 600\,\mu\text{s after the prompt.} \end{aligned} \tag{4.6a}$$

$$\begin{aligned} &\text{No valid triggers in the } 200\,\mu\text{s before the time } t_{prompt} + T_{\text{off}} \\ &\text{and only one (the delayed) in the } 600\,\mu\text{s after it.} \end{aligned} \tag{4.6b}$$

where t_{prompt} is the prompt trigger time.

The time offset is given by the formula:

$$T_{\text{off}} = 1\,\text{s} + n \times T_{\text{iso}} \tag{4.7}$$

where $T_{\text{iso}} = 800\,\mu\text{s}$ is the length of the isolation window of equation 4.2e and n is an integer in the interval $[0, 1999]$. The minimum offset is 1 s, which ensures only uncorrelated triggers are considered. The use of 2000 consecutive time windows allows to increase the statistics of this background to have a more precise measurement.

The technique to measure the accidental background from the DATA just described is known as the *off-time method*. It provides the rate and the spectrum of the accidental background for the *off-time* sample defined by equation 4.5. However, in order to have a measurement corresponding to the *on-time* sample defined by equation 4.2c, the off-time results must be corrected to account for the different effect that some of the cuts have in the two samples.

Muon veto correction From section 4.1.1.1 it is understood that, in order to consider a trigger valid, the 1 ms preceding it must be free of triggers tagged as muons. Whereas the probability of having a prompt trigger vetoed by a muon is the same in the on-time and off-time samples, the delayed one is not. In the on-time sample, the proximity in time required for the delayed trigger results in the muon-free 1 ms windows before the prompt and delayed triggers being mostly overlapped; whereas in the off-time sample the muon-free windows before the prompt and delayed triggers do not overlap at all, thus increasing the probability of vetoing the delayed with respect to the on-time case.

The resulting underestimation is corrected using a correction factor, f_μ, which is calculated as the ratio of the Poissonian probabilities of non-vetoing the delayed trigger (i.e. having 0 muons before it) in the on-time and off-time samples, P_μ^{on} and P_μ^{off}, respectively; once the vetoed events in which the muon occurs before the prompt trigger are removed:

$$f_\mu = \frac{P_\mu^{\text{on}}}{P_\mu^{\text{off}}} \tag{4.8a}$$

with

$$P_{\mu}^{\text{on}} = \frac{\int_{\Delta T_{\text{min}}}^{\Delta T_{\text{max}}} \mathrm{e}^{-R_{\mu} t'} \, \mathrm{d}t'}{\int_{\Delta T_{\text{min}}}^{\Delta T_{\text{max}}} \mathrm{d}t'} \tag{4.8b}$$

$$P_{\mu}^{\text{off}} = \mathrm{e}^{-R_{\mu} T_{\mu}} \tag{4.8c}$$

where R_{μ} is the rate of the muon-tagged events and $T_{\mu} = 1\,\text{ms}$ is the length of the muon veto window. P_{μ}^{on} corresponds to the averaged probability in the time interval from equation 4.2c, so the integral limits are $\Delta T_{\text{min}} = 0.5\,\mu\text{s}$ and $\Delta T_{\text{max}} = 150\,\mu\text{s}$. Computing the expressions in eqs. 4.8 gives $f_{\mu} = 1.04272 \pm 0.00002$.

Isolation cut correction In the on-time sample only one isolation cut is applied (eq. 4.2e) whereas in the off-time sample two cuts are used (eqs. 4.6a and 4.6b). Therefore, there is an additional inefficiency in the off-time sample which must be corrected. The correction factor, f_{iso}, is computed as the ratio of the probabilities in the on-time and off-time samples of the delayed trigger, d, passing the isolation cut provided that the prompt trigger, p, has passed the isolation cut:

$$f_{\text{iso}} = \frac{P^{\text{on}}(d \text{ is isolated} | p \text{ is isolated})}{P^{\text{off}}(d \text{ is isolated} | p \text{ is isolated})} \tag{4.9}$$

In the on-time sample, no additional isolation cut is requested for the delayed, so $P^{\text{on}}(d \text{ is isolated} | p \text{ is isolated}) = 1$. In the off-time sample, the probability is given by:

$$P^{\text{off}}(d \text{ is isolated} | p \text{ is isolated}) = \mathrm{e}^{-R_{\text{s}} T_{\text{iso}}} + h(R_{\text{s}}, T_{\text{iso}}, R_{\mu}, T_{\mu}, \Delta T) \tag{4.10}$$

where R_{s} is the singles rate. The first term in 4.10 corresponds to the Poisson probability of having 0 additional valid triggers in the isolation window covering the delayed trigger. The second term is a correction to the previous one which must be computed numerically (see [168] for the computation) and which accounts for the case in which there is a preceding trigger which would make the delayed fail the isolation cut, but this preceding trigger has been removed from the sample by the muon veto. The value of f_{iso} is found to be 1.0106 ± 0.0002.

Run length correction In the off-time method, the trigger coincidences are searched within multiple time windows shifted by a time offset T_{off}. Hence, the live-time in which those coincidences can be found is reduced with respect to the on-time coincidences, since the last seconds of each run produce less coincidences because the time windows cannot be opened beyond the end of the run. This is corrected using the following correction factor:

$$f_{\text{run}} = \frac{\overline{T}_{\text{run}}}{\overline{T}_{\text{run}} - \overline{T}_{\text{off}}} \tag{4.11}$$

where $\overline{T}_{\rm run}$ is the average length of the run and

$$\overline{T}_{\rm off} = \sum_{n=0}^{1999} (1\,{\rm s} + n \times T_{\rm iso})/2000$$

is the averaged time offset of equation 4.7. After calculation, $f_{\rm run} = 1.00055$ with negligible uncertainty.

Li+He veto correction The Li+He veto, which is only applied to the prompt trigger to reduce the IBD-like coincidences caused by the decay of cosmogenic isotopes (section 4.1.3.1), also has an effect on the on-time accidentals by rejecting those which have a cosmogenic isotopes decay (typically ^{12}B) as delayed. The mechanism is as follows: the prompt trigger is given by a single trigger unrelated to the muon. Because the IBD selection features a prompt-delayed distance cut (eq. 4.2d), this prompt must be close to the delayed. Moreover, the delayed is close to the muon, since it is correlated with it. Therefore, the prompt is also close to the muon; so it has a high probability of being vetoed by the Li+He veto, which uses the distance to the muon track as one of the variables. This does not happen in the off-time sample, in which the delayed trigger is so far in time that it has no correlation with any muon preceding the prompt trigger. As a result, the off-time measurement is overestimated and requires a correction estimated by:

$$f_{\rm Li} = \frac{\epsilon_{\rm Li}^{\rm on}}{\epsilon_{\rm Li}^{\rm off}} \tag{4.12}$$

where $\epsilon_{\rm Li}^{\rm on}$ and $\epsilon_{\rm Li}^{\rm off}$ are the Li+He veto efficiencies computed for accidental events in the on-time and off-time samples, respectively. The on-time accidental sample is collected running the IBD selection with the cut $\Delta R > 1.4\,{\rm m}$ instead of the one in eq. 4.2d, since at that distance only accidental coincidences are expected (cf. figure 4.8). Requiring a large prompt-delayed separation is against the mechanism described above, but it is the only way to select on-time accidentals. The possible bias introduced is incorporated as a systematic uncertainty on the result: $f_{\rm Li} = 0.875 \pm 0.017({\rm stat}) \pm 0.027({\rm syst})$.

The total correction factor applied to the accidental rate is the product of the ones above:

$$f_{\rm acc} = f_\mu \times f_{\rm iso} \times f_{\rm run} \times f_{\rm Li} = 0.923 \pm 0.034 \tag{4.13}$$

The final accidental background rate is $0.0701 \pm 0.0003({\rm stat}) \pm 0.0026({\rm syst})\ {\rm day}^{-1}$, where the systematic uncertainty corresponds to the conversion from the off-time measurement to the on-time estimation using the factor in equation 4.13. In figures 4.17 and 4.18, the prompt spectrum used in the θ_{13} Rate+Shape analysis (section 6.2.1) and the delayed spectrum used to subtract the accidental background for the estimation of the neutron detection efficiency (section 5.4.2.1), are shown.

There is no need for a specific veto against this background, as it was the case of the other backgrounds, since the distance cut of the IBD selection (equation 4.2d) acts already as an accidental veto, rejecting 91% of the accidental events as seen in figure 4.8.

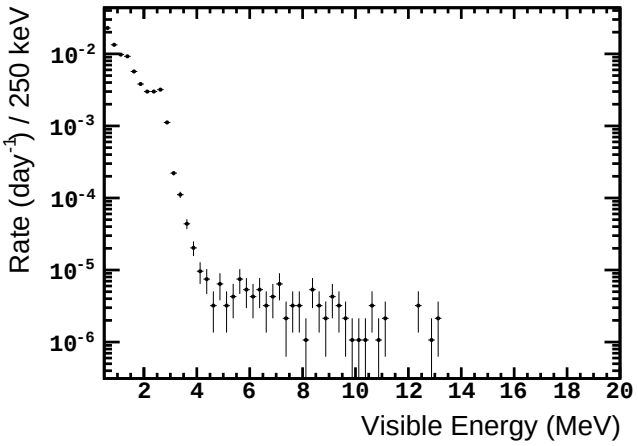


Figure 4.17: Prompt trigger energy spectrum for the accidental background. The accidental sample is collected in 2000 off-time windows and rescaled accordingly. From [169].

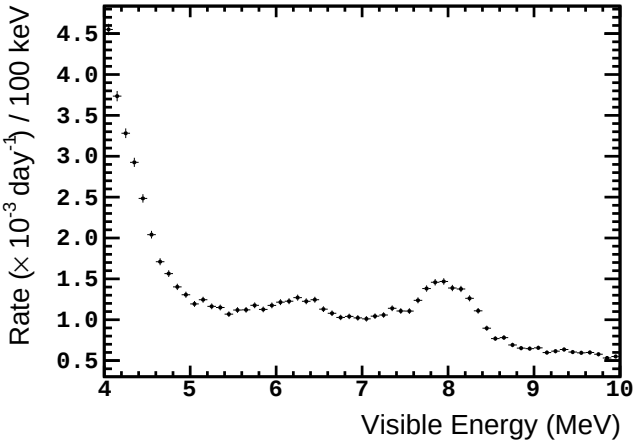


Figure 4.18: Delayed trigger energy spectrum for the accidental background. The accidental sample is collected in 2000 off-time windows and rescaled accordingly. From [169].

Other backgrounds

As it was stated in the explanation about the Li+He veto construction (section 4.1.3.2), ^{12}B is the short-lived cosmogenic isotope most abundantly produced in the detector, but since it is not a βn emitter, it cannot cause IBD-like coincidences on its own. Moreover, the coincidences caused by the decay of a ^{12}B nucleus and any uncorrelated trigger (and vice versa) are already taken into account in the accidental background measurement (see section 4.1.3.3). The mechanism that requires a dedicated calculation is the one in which the muon which originates the ^{12}B also produces an associated particle capable of acting as a prompt or delayed trigger to make the coincidence, that is, another ^{12}B nucleus or a neutron. The latter possibility is highly suppressed since it requires the neutron to remain uncaptured for more than 1 ms (the muon veto window). In order to make an estimation of the $^{12}\mathrm{B} - {}^{12}\mathrm{B}$ events, the off-time method used to measure the accidental background is employed, decreasing the 1 s time offset of equation 4.7 to a few microseconds so as to select ^{12}B decays (since the lifetime is 29.1 ms). A rate < 0.03 events/day is found, so this background can be ignored.

Another correlated background is given by the reaction $^{13}\mathrm{C}(\alpha, \mathrm{n})^{16}\mathrm{O}$, where the α is mostly produced in the decay of one nucleus from the Uranium or Thorium chains. The rate of such reaction is estimated from the contamination of α emitters in the detector and found to be negligible.

Table 4.2 summarizes the event rates after the full Gd-based $\overline{\nu}_e$ selection is over. The novel background rejection techniques just described have enabled a reduction of the background rates with respect to the previous Gd selection [1]: the fast neutron and stopping muon rate has decreased by 48% thanks to the newly added IV and F_V vetoes; the cosmogenic isotope rate has decreased by 22% because of the Li+He veto, which in addition has produced the first data-driven measurement of the ^{9}Li and ^{8}He spectrum in this experiment; and the accidental coincidence rate has decreased by 73% benefiting from the new cut on the distance between the prompt and delayed trigger vertices. The reduction in the background rates has not been at expense of the signal efficiency thanks to the wider cut on the time interval between triggers and the lowered cut on the minimum energy of the delayed trigger. Thus, the signal-to-background ratio has passed from ~ 16 in the previous Gd-based selection to the current ~ 22.

4.1.3.4 Reactors-off measurement

The rare occassion when the two reactors are off allows to perform a direct measurement of the backgrounds in the detector, up to a few residual $\overline{\nu}_e$ from the stopped reactors. In the current dataset, 7.24 live-days have been accumulated in which 7 events survive all the selection cuts and vetoes explained in the previous sections. The total prediction in this circumstances amounts to $12.9^{+3.1}_{-1.4}$ events, including 1.57 ± 0.47 events due to the residual $\overline{\nu}_e$ which are computed in a simulation [170] corresponding to a null θ_{13}. The compatibility of the measurement with the prediction is 9.0% (1.7σ), and the fact that less events have been observed than predicted disfavors the existence of an additional unknown background. The number of events is too low to extract any valuable spectral information, but it is used as a constraint to the total background rate

Event type	Rate (events/day)	Gd-III/Gd-II
IBD candidates (average)	37.66 ± 0.29	—
Backgrounds		
Fast-n and stopping-μ	0.604 ± 0.051	0.52
^{9}Li and ^{8}He	$0.97^{+0.41}_{-0.16}$	0.78
Accidental	$0.0701^{\pm 0.0003(\mathrm{stat})}_{\pm 0.0026(\mathrm{syst})}$	0.27
^{12}B	< 0.03	N/A
$^{13}\mathrm{C}(\alpha, \mathrm{n})^{16}\mathrm{O}$	< 0.1	N/A

Table 4.2: Summary of the rates of the events passing the Gd-based $\overline{\nu}_e$ selection. The IBD candidate (including backgrounds) average rate during the periods with at least one reactor on is shown; the actual IBD rate changes in time as a function of the thermal power of the nuclear reactors as in equation 2.7. The background rates correspond to the expected rates remaining in the IBD candidate sample after the full $\overline{\nu}_e$ selection is performed, estimated from samples rejected by it as explained in the text. The column Gd-III/Gd-II shows the reduction of the background rate in the current selection [2] with respect to the previous publication [1], after correcting for the different prompt energy range.

in the θ_{13} analyses of chapter 6.

4.2 Hydrogen selection

The main motivation of a selection using the IBD neutrons captured on H is to obtain an independent data sample to measure θ_{13} with increased $\overline{\nu}_e$ statistics with respect to the Gd-based measurement, and eventually combine both results to improve the precision on θ_{13}. As it is pointed out in section 2.1, Double Chooz has the smallest detector among the current reactor antineutrino experiments to measure θ_{13}. Therefore, the addition of IBD events in which the neutron is captured on H partially compensates the size disadvantage. Since the Target scintillator is loaded with Gd, which makes it very effective in capturing neutrons, most of the captures on H will occur in the Gamma Catcher. Therefore, a factor ~ 2 more $\overline{\nu}_e$ with respect to the number in the Gd selection are expected to be collected using H captures, based on the volume of the Gamma Catcher ($22.5\,\mathrm{m}^3$) and the Target ($10.3\,\mathrm{m}^3$).

The major difficulty of this selection is the background events produced by the accidental coincidences of triggers. Unlike the neutron captures on Gd, which release a total energy $\sim 8\,\mathrm{MeV}$ which is far above the typical energies of natural radioactivity; the capture on H yields a single γ-ray with $2.22\,\mathrm{MeV}$, which is immersed in the natural radioactivity spectrum. Moreover, the H nucleus has a neutron capture cross section of $0.33\,\mathrm{b}$, much smaller than Gd ($\sim 48769\,\mathrm{b}$, averaged over all isotopes), which results in a longer capture mean time, increasing the probability of accidental coincidences and reducing the selection efficiency.

Double Chooz demonstrated that this analysis was feasible producing the first measurement of θ_{13} using H captures [72]. However, this selection had a

Cut	Inefficiency (%)	Correction factor
Muon veto	$6.01 \pm (< 0.01)$	$0.9399 \pm (< 0.0001)$
Light noise	$0.06 \pm (< 0.01)$	$0.9994 \pm (< 0.0001)$
Isolation	$2.12 \pm (< 0.01)$	$0.9788 \pm (< 0.0001)$
OV veto	$0.06 \pm (< 0.01)$	$0.9994 \pm (< 0.0001)$
IV veto	0.00 ± 0.17	1.0000 ± 0.0017
F_V veto	0.05 ± 0.02	0.9995 ± 0.0002
MPS veto	0.00 ± 0.10	1.0000 ± 0.0010
Li+He veto	0.51 ± 0.01	0.9949 ± 0.0001

Table 4.3: Inefficiencies in the H-based $\overline{\nu}_e$ selection due to the the background-oriented cuts and the associated correction factors to the normalization of the MC $\overline{\nu}_e$ simulation. Since these cuts are not applied to the MC, the correction factor is simply the cut efficiency in the DATA (which is the complementary of the inefficiency).

signal to background ratio of $\sim 1:1$ due to the accidental background. The goal of this new selection is to improve it.

The data set used for the hydrogen selection is the same as in the gadolinium selection (presented in section 4.1). The live-time amounts to 462.72 days, after the subtraction of the dead-time caused by a slightly longer muon veto (discussed in section 4.2.1) and the OV veto, which is identical to the one in the Gd selection. This results in an exposure of 209.4 GW-ton-years, roughly a factor two more with respect to the first H-based analysis [72].

The selection of IBD candidates is organized similarly to the Gd-based one, as figure 4.1 shows.

4.2.1 Single trigger selection

As it happens in the Gd case (section 4.1.1), a preliminary selection of valid triggers (singles) is carried out before looking for the coincidence of two of them characteristic of the IBD reaction. The minimum energy cut is kept at 0.4 MeV, ensuring a 100% trigger efficiency with negligible uncertainty. Furthermore, the muon veto and the light noise cuts (described next) reduce the number of triggers unrelated to $\overline{\nu}_e$, so the coincidence search is made over a cleaner sample.

4.2.1.1 Muon veto

The muon definition is maintained as a trigger with a visible energy larger than 20 MeV in the Inner Detector or 16 MeV in the Inner Veto. Upon such an event, that trigger and all the following triggers in a 1.25 ms window are rejected. The muon veto time has been extended by 25% compared to the 1 ms of the Gd selection (section 4.1.1.1) to gain a factor ~ 10 in the background rejection, at the expense of increasing 1.5% the inefficiency in the $\overline{\nu}_e$ selection; which now is 6.01% with $< 0.01\%$ uncertainty. As before, the normalization of the $\overline{\nu}_e$ MC must be corrected for this inefficiency since this cut is not applied because the muon background is not simulated (see table 4.3).

4.2.1.2 Light noise cuts

The light noise is a background consisting in the emission of light from the PMT bases, which was already introduced in section 4.2.1.2 of the Gd selection. The same cuts based on the homogeneity of the observed charge and the simultaneity of the hit times of the PMTs are used here (equations 4.1). The inefficiency in the $\overline{\nu}_e$ selection is estimated using the IBD MC simulation again, where this background is absent. Most of the IBD interactions leading to a neutron capture on H occur in the Gamma Catcher, the volume which surrounds the central volume (Target). In consequence, these events tend to illuminate the PMTs less homogeneously and simultaneously than those within the Target (mostly Gd captures), which results in an inefficiency of $(0.0604 \pm 0.0012)\%$ that is bigger than the Gd one ($(0.0124 \pm 0.0008)\%$), but still well below 0.1%.

4.2.2 Inverse beta-decay event selection

Once the singles sample is assembled (giving a trigger rate of $13.2\,\mathrm{s}^{-1}$), the search for the coincidence within a few microseconds of the two triggers which evidence an $\overline{\nu}_e$ underwent an IBD interaction can start: a *prompt* trigger produced by the positron kinetic energy loss and ultimate annihilation with an electron from the detector, and a *delayed* trigger given by the radiative capture of the neutron. Hence, the traditional strategy relies on a set of one-dimensional cuts on the observables of the IBD coincidence: visible energy of the prompt trigger, visible energy of the delayed trigger, time interval between the prompt and delayed triggers, distance between the prompt and delayed reconstructed vertices, and an isolation cut around the trigger pair to reject coincidences of more than two triggers caused by backgrounds and ensure the coincidence is univocal (see equations 4.2 for the Gd case). While this strategy works nicely for the Gd selection thanks to the very distinctive features of the neutron capture on Gd ($\sim 8\,\mathrm{MeV}$ energy release above the natural radioactivity, $\sim 30\,\mu\mathrm{s}$ capture time); in the case of the H selection ($\sim 2.2\,\mathrm{MeV}$ energy release obscured by the natural radioactivity, $\sim 200\,\mu\mathrm{s}$ capture time) it forces us to choose between a high signal efficiency but a poor signal to background ratio or the contrary. However, the dilemma can be circumvented by abandoning the one-dimensional cut strategy, which fails to grasp the background rejection power available in the multidimensional space of the observables, and adopting a multivariate analysis.

The approach chosen for the H selection is to develop an Artificial Neural Network (ANN) classifier using as inputs the visible energy of the delayed trigger (E_{delayed}), the time interval (ΔT) and the reconstructed distance (ΔR) between the prompt and delayed triggers. These three variables are shown to exhibit a different interrelationship for the IBD coincidences and the accidental coincidences (see figure 4.19). The IBD coincidences cluster around the H capture peak energy, with short capture times and distances characteristic of the common origin of the positron and the neutron. On the other hand, the accidental coincidences tend to concentrate at lower energies (where the natural radioactivity is more intense), with a flat distribution in time as it is expected from random coincidences, and with large distances due to the uncorrelated origins. The visible energy of the prompt trigger is not included in the ANN to preserve the prompt spectrum shape which is used to measure θ_{13}.

The ANN is a feed-forward multilayer perceptron from the `TMVA` package

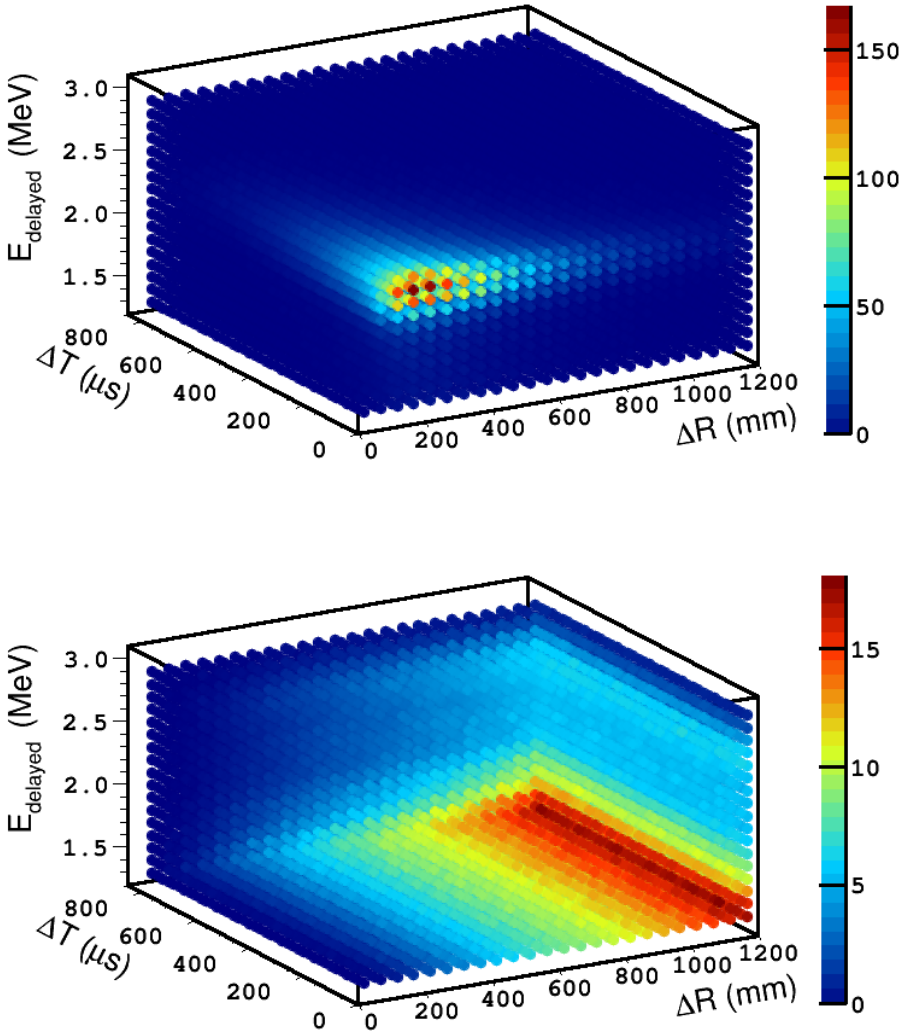


Figure 4.19: Three-dimensional distribution of the input variables to the ANN classifier: visible energy of the delayed trigger (E_{delayed}), time interval (ΔT) and the reconstructed distance (ΔR) between the prompt and delayed triggers. The IBD MC simulation (top) and an accidental background sample (bottom) collected using the off-time method (see section 4.2.3.3; with the time offset subtracted for the plot) are shown. The number of events is coded as color. Both distributions are normalized to the expected number of events in the DATA sample (oscillation not included). Points are drawn in ascending order of events.

[171] included in the ROOT framework [172]. The input layer consists of $3+1$ neurons (for the three input variables plus one bias node with a constant value); one hidden layer which is composed of $8+1$ neurons (one being a bias node); and the output layer which has a single neuron which gives the classifier variable. The activation function of the 8 hidden neurons is a hyperbolic tangent, while the rest have a linear function, resulting in a continuous output ranging from -1.2 (most accidental-like) to $+1.2$ (most IBD-like). The ANN is trained with a back-propagation method using an IBD MC sample as signal and an accidental background sample obtained with the off-time method as background. Since the training samples must be independent of the data samples used for the actual analysis, the IBD MC corresponds to an independently generated one not used in the θ_{13} analysis, and the accidental background is collected using 200 coincidence time windows shifted by a different time offset, T_{off}, from the one used for the oscillation analysis (the minimum offset used is 2 s instead of the standard 1 s). In addition to the valid trigger cuts described in section 4.2.1, the following cuts are used to select the training samples:

- Prompt visible energy: $0.5 < E_{\text{prompt}} < 20\,\text{MeV}$.
- Delayed visible energy: $1.3 < E_{\text{delayed}} < 3.5\,\text{MeV}$.
- Prompt-delayed time interval:

  $0.25 < \Delta T < 1000\,\mu\text{s}$ (for the IBD MC sample),
  $0.25\,\mu\text{s} + T_{\text{off}} < \Delta T < 1000\,\mu\text{s} + T_{\text{off}}$ (for the accidental sample).

- Prompt-delayed vertices distance: $\Delta R < 150\,\text{cm}$.

Moreover, since the accidental sample is extracted from the DATA, it must pass the vetoes against fast neutrons, stopping muons and cosmogenic background:

- No valid triggers in the $1000\,\mu\text{s}$ before and in the $900\,\mu\text{s}$ after the prompt trigger time, t_{prompt}.
  No valid triggers in the $1000\,\mu\text{s}$ before the time $t_{prompt} + T_{\text{off}}$ and only one (the delayed) in the $900\,\mu\text{s}$ after it.
- OV veto: reject prompt triggers coincident with an OV one.
- IV veto: reject triggers (prompt or delayed) which satisfy these cuts:

  IV PMT hit multiplicity ≥ 2
  Total charge in the IV > 400 charge units
  ID-IV vertices distance $< 4\,\text{m}$
  ID-IV time interval: $-110 < \Delta T_{\text{ID}-\text{IV}} < -20\,\text{ns}$.

- F_V veto: reject delayed triggers with $E_{\text{delayed}} < 0.2208 \exp(F_V/1.818)$.
- Li+He veto: reject prompt triggers with a likelihood > 0.4.

These cuts are already presented in the context of the Gd selection (see sections 4.1.2, 4.1.3.1, 4.1.3.2), with some of their values adapted here to the H selection. Figure 4.20 shows the distributions of the ANN input variables in the training

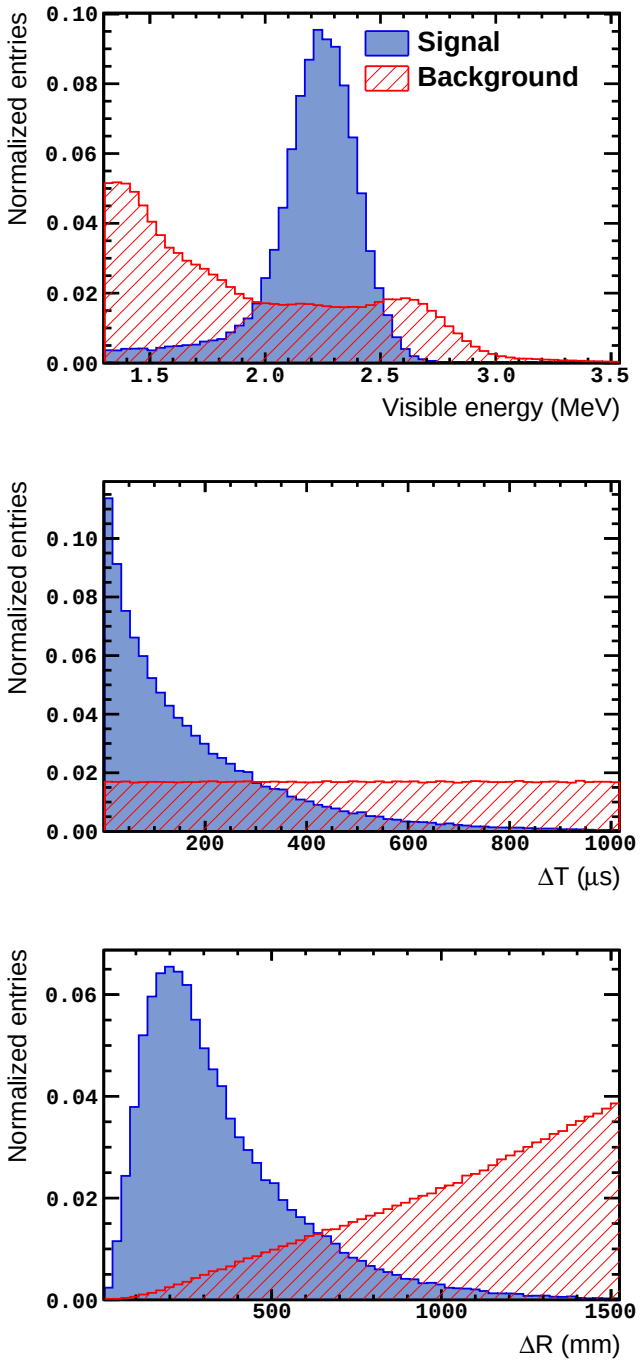


Figure 4.20: Distributions of the ANN input variables in the training samples: visible energy of the delayed trigger (E_{delayed}, top), time interval (ΔT, middle) and the reconstructed distance (ΔR, bottom) between the prompt and delayed triggers. The signal training sample (blue filled histograms) is an IBD MC simulation. The background training sample (red hatched histograms) are accidental coincidences obtained from DATA using the off-time method (the time offset has been subtracted in the middle plot). All histogram integrals are normalized to unity. From [173].

samples. The ranges of the three variables are enlarged with respect to the actual ones used for the IBD selection for oscillation analysis described later, so as the ANN can be used in the estimation of the neutron selection efficiency (presented in section 5.5.2), in which these cuts are loosened.

The ANN can be viewed as a mathematical function which maps $\mathbb{R}^{+3} \to \mathbb{R}$, where the domain is the set of $(E_{\text{delayed}}, \Delta T, \Delta R)$ triads and the image is the ANN classifier output, $ANN = f(E_{\text{delayed}}, \Delta T, \Delta R)$. Once the ANN is trained, its response can be characterized by a numerical scan over the allowed ranges of the input variables, which are defined by the selection cuts of the training samples. The result of such scan is shown in figure 4.21. It is observed how the ANN considers as signal-like the volume where the IBD events cluster as seen in the three-dimensional distribution of figure 4.19. The shape of the signal-like volume (resembling a tetrahedron) anticipates the improved performance brought by the multivariate analysis: for a fixed E_{delayed} (an horizontal plane in figure 4.21), the cut on ΔT depends on the value of ΔR (or vice versa). If an event has a short ΔR (typical of an IBD pair), the ΔT cut is loose to boost the signal efficiency. On the contrary, if the event has a large ΔR (accidental-like), the ΔT cut is tightened to have a good signal to background ratio. The area of the signal-like triangle in the $\Delta T - \Delta R$ plane increases with the delayed energy, reaching the maximum at $E_{\text{delayed}} = 2.2\,\text{MeV}$ corresponding to the energy of the neutron capture on H, and decreasing sharply after. A one-dimensional cut strategy would produce an hexahedron, in which the signal-like region in the $\Delta T - \Delta R$ plane would be a constant rectangle, which is always the same irrespective of the values of the three variable. Clearly, such approach does not exploit all the information available. Figure 4.21 shows an ambiguous region ($ANN \sim 0$, neither signal nor background-like) for $E_{\text{delayed}} > 3.1\,\text{MeV}$ where no signal is expected (cf. figure 4.19). It seems to be an artifact of the ANN, but it can be removed easily by imposing a $E_{\text{delayed}} \leq 3\,\text{MeV}$ cut on the IBD selection. The discovery of this artifact advocates for the convenience of performing this kind of study before relying on the multivariate classifier whenever it is possible (i.e. the number of input variables is manageable).

When applied to actual events (either DATA or MC), the ANN classifier is treated as another variable to cut on (see figure 4.22). Thus, the IBD selection using neutrons captured on H becomes

$$\text{Prompt visible energy: } 1 < E_{\text{prompt}} < 20\,\text{MeV}. \tag{4.14a}$$

$$\text{Delayed visible energy: } 1.3 < E_{\text{delayed}} < 3.0\,\text{MeV}. \tag{4.14b}$$

$$\text{Prompt-delayed time interval: } 0.5 < \Delta T < 800\,\mu\text{s}. \tag{4.14c}$$

$$\text{Prompt-delayed vertices distance: } \Delta R < 120\,\text{cm}. \tag{4.14d}$$

$$\text{ANN classifier: } ANN > -0.23. \tag{4.14e}$$

$$\text{No valid triggers in the } 800\,\mu\text{s before the prompt and only one (the delayed) in the } 900\,\mu\text{s after the prompt.} \tag{4.14f}$$

The values of the cuts on the E_{delayed}, ΔT, ΔR and ANN are chosen to achieve a signal efficiency of 80%, which is close to the 85% in the previous H selection [72], in which only one-dimensional cuts were used: $1.5 < E_{\text{delayed}} < 3.0\,\text{MeV}$, $10 < \Delta T < 600\,\mu\text{s}$, $\Delta R < 90\,\text{cm}$. However, this selection features a signal to accidental background ratio of 11.9, which outperforms the previous selection which had 1.1.

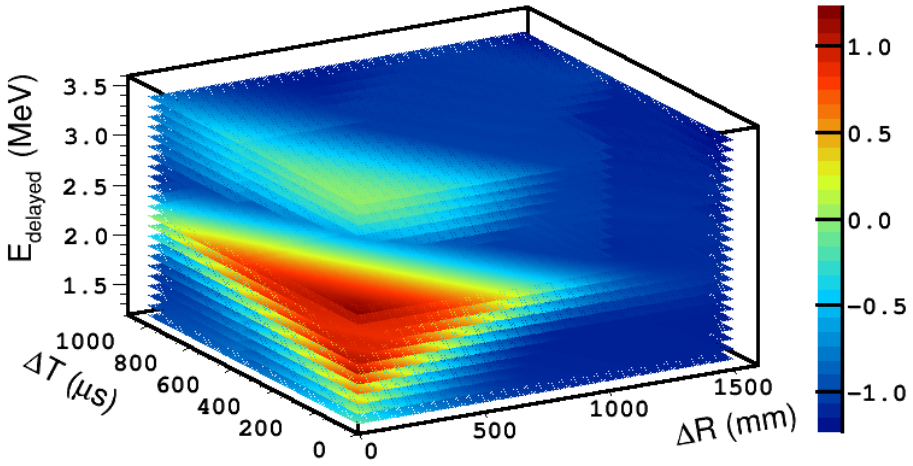


Figure 4.21: Numerical scan over the Artificial Neural Network input variables: visible energy of the delayed trigger (E_{delayed}), time interval (ΔT) and the reconstructed distance (ΔR) between the prompt and delayed triggers. The ANN classifier (ANN) value is represented as color, from dark blue (most background-like) to dark red (most signal-like). Points are drawn in ascending order of ANN.

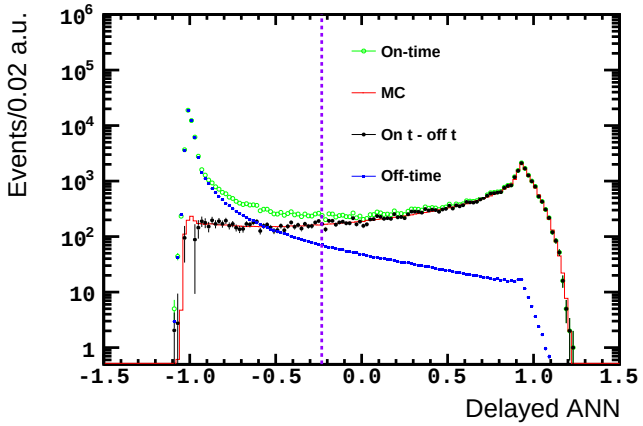


Figure 4.22: Distribution of the ANN classifier output. The DATA collected using the IBD selection (labeled as on-time) are shown as empty green circles. Blue squares show an accidental background sample collected in 200 off-time windows and rescaled accordingly. The DATA resulting from the subtraction of the off-time sample to the on-time one are shown as black circles. The IBD MC simulation (red histogram) is shown, normalized to the integral of the accidental-background subtracted DATA. The cut on the ANN classifier is displayed as a purple dashed line (events to the right of the line are accepted).

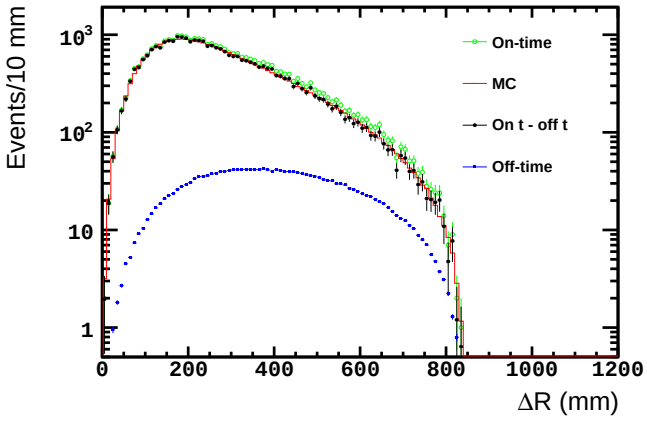


Figure 4.23: Distribution of the prompt-delayed reconstructed vertices distance after all cuts are applied. The color and shape of the data markers follow the ones of fig. 4.22.

Actually, the ANN cut makes some of the cuts in equations 4.14 redundant. The $ANN > -0.23$ imposes an effective cut on $\Delta R \lesssim 84\,\text{cm}$ (see figure 4.23), which is stronger than equation 4.14d. Likewise, the upper delayed energy cut of equation 4.14b is only needed to exclude a disconnected allowed region due to the ANN artifact described before, since the cut $ANN > -0.23$ implies an effective cut $E_{\text{delayed}} \lesssim 2.8\,\text{MeV}$ (see figure 4.24).

In addition to the required changes to the IBD selection due to the fact that the neutron is being captured on a different nucleus, there are two differences more with respect to the Gd selection discussed in section 4.1.2. The first one is a simple adjustment of the values of the isolation cut (eq. 4.14f), already presented for the Gd selection (see eq. 4.2e), to accommodate it for the different time interval allowed between triggers (see figure 4.25). This modifies the inefficiency in the $\overline{\nu}_e$ selection due to this cut, which now is 2.12% with $< 0.01\%$ uncertainty. The second difference is found in the minimum visible energy of the prompt, which has been raised from 0.5 to 1 MeV. The reason for this increment is to exclude the 0.511 MeV peak in the prompt energy resulting from IBD interactions in the Buffer volume (see figure 4.26). The Buffer is filled with a non-scintillating liquid, so the $\overline{\nu}_e$ which undergo an IBD reaction there produce a positron which loses its kinetic energy invisibly. However, one of the annihilation γ-rays can reach the scintillators and be absorbed there, resulting in a prompt trigger peaking at 0.511 MeV. If the neutron is subsequently captured either in the Gamma Catcher or in the Buffer but with the 2.2 MeV γ-ray depositing energy in the scintillators, the IBD coincidence is fulfilled. The presence of this peak would cause a significant normalization uncertainty due to small shifts of the energy scale, so it was decided to remove it. These Buffer IBDs do not represent a problem for the Gd selection since the $(4, 10)\,\text{MeV}$ delayed energy window already excludes them.

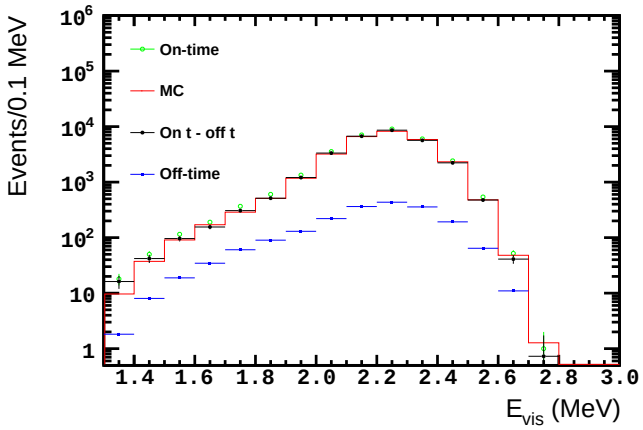


Figure 4.24: Delayed energy spectrum after all cuts are applied. The color and shape of the data markers follow the ones of fig. 4.22.

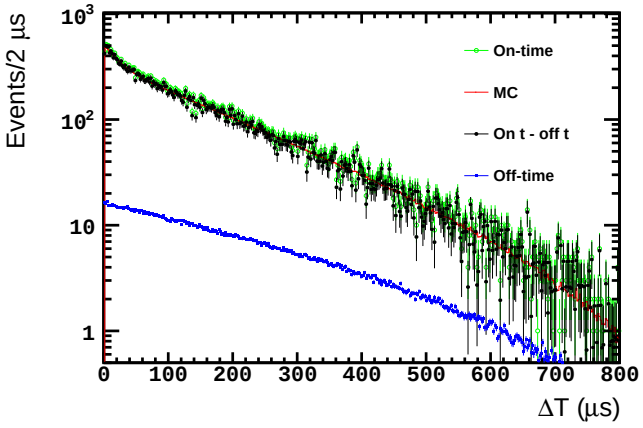


Figure 4.25: Distribution of the prompt-delayed trigger time interval after all cuts are applied. The color and shape of the data markers follow the ones of fig. 4.22. The time offset of the off-time sample has been subtracted for this plot.

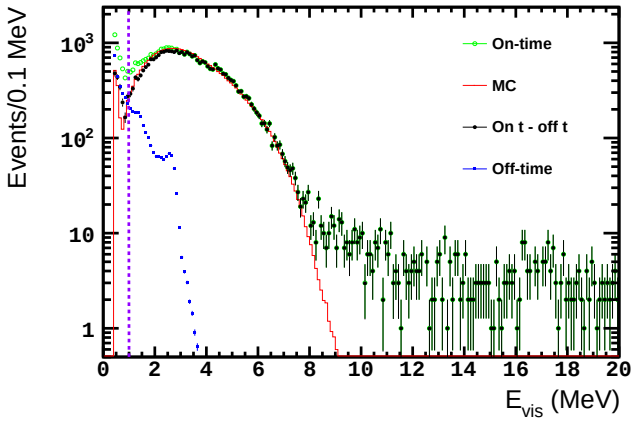


Figure 4.26: Prompt energy spectrum before the $E_{\text{prompt}} > 1\,\text{MeV}$ cut (purple dashed line). The color and shape of the data markers follow the ones of fig. 4.22.

4.2.3 Background vetoes and estimations

There are non-$\overline{\nu}_e$ events which survive the single trigger selection and are capable of causing fake IBD-like coincidences, so they remain in the selected sample after the cuts in equations 4.14 are applied. Some of them are the offspring of muons which interacted within or close by the detector or exhibit additional features which provide a handle to reduce them through additional cuts, as it was shown in the Gd selection (section 4.1.3). Nevertheless, they cannot be suppressed completely and hence the rate and the prompt spectrum shape of these background events must be estimated in order to take them into account when comparing the observed DATA to the prediction to measure θ_{13}.

4.2.3.1 Fast neutrons and stopping muons

These two backgrounds were already introduced in the context of the Gd selection (section 4.1.3.1). In the case of the fast neutrons, in which the delayed trigger is given by a neutron capture, the rate is expected to increase in the H selection; not only because of the higher number of hydrogen nuclei in the detector with respect to gadolinium ones (which makes up for the smaller capture cross-section) but also because the Gamma Catcher acted as passive shield against neutrons for the Target. Regarding the stopping muons, the delayed trigger is given by the Michel electron, so they are not affected directly by the change of nuclei; but since the number of Michel electrons descends as the energy is lowered, the interval used to select H captures yields a smaller number of stopping muons than in the Gd case.

As in the Gd selection, the rate of these two backgrounds can be reduced by vetoing the Inner Detector triggers which show coincident activity in the Outer Detector (OV or IV) or rejecting those which have an ill reconstruction likelihood (F_V veto). In addition, since the completion of the Gd selection, a new cut based on the PMT pulse shape (MPS veto) was devised and is featured

for the first time for the H selection [159].

OV veto

The OV veto aims at reducing muon-induced triggers by rejecting the prompt triggers which follow an Outer Veto trigger. Therefore, it is not affected by the choice of neutron captures on H as delayed triggers, and the condition already described for the Gd selection (veto of 224 ns after the Outer Veto hit) can be maintained, leading to the same inefficiency in the $\overline{\nu}_e$ selection: $(0.058 \pm 0.001)\%$.

IV veto

The IV veto (rejection of prompt triggers which show correlated activity in the Inner Veto) was already presented for the Gd selection (see equations 4.3) to reduce the rate of triggers due to stopping muons, fast neutrons and γ rays entering the detector. However, a similar veto was not implemented for the delayed trigger to avoid the rejection of $\overline{\nu}_e$ interactions, since the γ rays released after the radiative capture of an IBD neutron on a Gd nucleus were occasionally observed in the Inner Veto. However, the fact that the H-based $\overline{\nu}_e$ selection has the delayed energy window (equation 4.14b) placed where the natural radioactivity lies motivates the application of the IV veto also to the delayed trigger in order to decrease the rate of accidental IBD-like coincidences caused by this background. The IV veto helps to reduce the number of triggers given by external γ rays by cutting on those which Compton-scatter in the Inner Veto before reaching the Inner Detector; achieving a reduction approximately of 14.7% (IV veto on the prompt trigger), 14.8% (IV veto on the delayed trigger) and 27% (IV veto on both prompt and delayed triggers) of the remaining accidental background.

The definition of the veto is similar to the Gd case, with the prompt or delayed trigger being rejected if it satisfies all the following cuts:

$$\text{IV PMT hit multiplicity} \geq 2 \tag{4.15a}$$

$$\text{Total charge in the IV} > 400\,\text{charge units} \; (\approx 0.2\,\text{MeV}) \tag{4.15b}$$

$$\text{ID-IV vertices distance} < 4\,\text{m} \tag{4.15c}$$

$$\text{ID-IV time interval: } -110 < \Delta T_{\text{ID}-\text{IV}} < -20\,\text{ns} \tag{4.15d}$$

The Inner Detector-Inner Veto reconstructed vertex distance and trigger time interval cuts have been modified slightly with respect to the Gd version to minimize the rejection of $\overline{\nu}_e$ events due to accidental coincidences of an $\overline{\nu}_e$ interaction in the Inner Detector and an uncorrelated signal in the Inner Veto, or an $\overline{\nu}_e$ interaction in the Inner Detector which produces a γ ray which deposits energy in the Inner Veto.

The inefficiencies in the $\overline{\nu}_e$ selection caused by the IV vetoes are estimated from the Inner Detector-Inner Veto trigger time interval distributions of the selected IBD candidates after all vetoes have been applied (see figure 4.27). Since the shape in the cut region defined by equation 4.15d) is consistent with a flat distribution (no deficit is observed compared to the sideband), and assuming this flat distribution to be entirely due to IBD events, an upper limit on the inefficiency in the $\overline{\nu}_e$ selection is set at 0.169%.

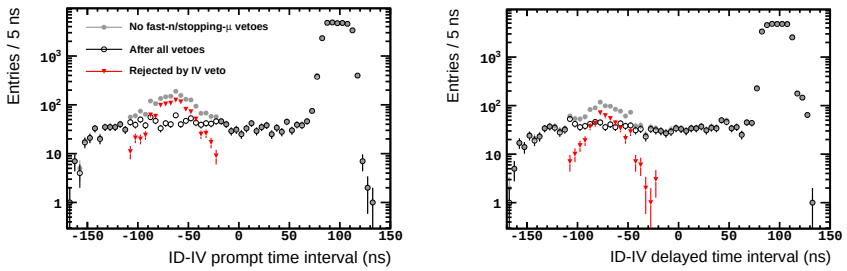


Figure 4.27: ID-IV prompt (left) and delayed (right) trigger time interval distributions. The trigger time is determined by a charge-weighted sum of the individual PMT start times. In both plots, the data before applying the OV, IV, F_V and MPS vetoes (gray circles) shows two peaks: the rightmost one corresponds to events with low charge in the Inner Veto (mostly $\overline{\nu}_e$) and the leftmost one is caused by the fast neutron and stopping muon contamination (ID-IV time offset is not corrected). The data after vetoes (empty black circles) does not exhibit an excess in the background region but a flat distribution corresponding to coincidences of uncorrelated events in the Inner Detector and the Inner Veto, and is used to estimate the inefficiency of the IV veto as explained in the text. The events rejected exclusively by the IV veto are shown as down-pointing red triangles. Edited from [174].

F_V veto

The Gd-based $\overline{\nu}_e$ selection showed that the value of the vertex reconstruction likelihood (specifically its negative logarithm, called F_V, equation 3.13) of the delayed trigger could be used to distinguish between the events occurring in the Target and Gamma Catcher volumes, which exhibit a low F_V value (i.e. a good reconstruction) and where neutron captures from IBD interactions are dominant, from those which occur out of the expected regions, thereby being badly reconstructed, such as stopping muons (decaying in the chimney) or remaining light noise events (originating from one of the PMTs on the Buffer wall).

In the Gd case, it was observed that a good separation between signal-like and background events was achieved using a two-dimensional cut in the $E_{\text{delayed}} - F_V$ space (see figure 4.10). Therefore, a similar cut was proposed for the H selection. However, in this case, the signal and background distributions are not so clearly detached (see figure 4.28), which required to adapt the cut parameters. The optimal veto in terms of minimal rejection of $\overline{\nu}_e$ and maximal rejection of background events was determined to be:

$$E_{\text{delayed}} < 0.2755 \exp(F_V/2.0125), \tag{4.16}$$

The inefficiency in the $\overline{\nu}_e$ selection as a result of this veto is estimated from the number of IBD candidates rejected by it with respect to the sample with the veto not applied. Since the rejected events are mostly background events due to the stopping muons, accidental coincidences and fast neutrons, a supplementary set of cuts is introduced for this estimation: stopping muons are removed by requiring a prompt-delayed time interval $\Delta T > 10\,\mu$s (~ 4.55 times the muon mean lifetime) and accidentals are suppressed by cutting on the prompt en-

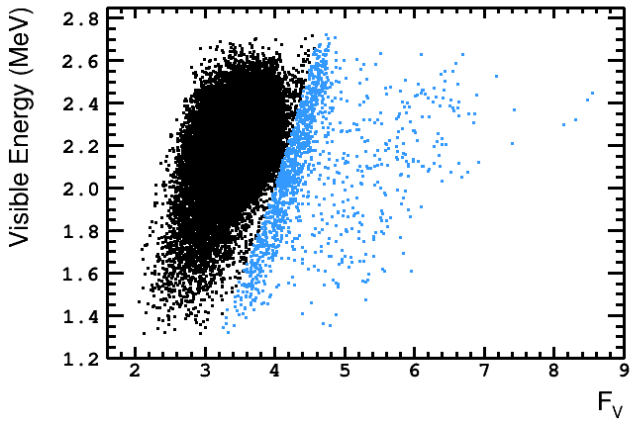


Figure 4.28: Two-dimensional distribution of the IBD candidates in the delayed trigger visible energy$-F_V$ space before the F_V veto (eq. 4.16) is applied. The selected IBD candidates after the veto are shown in black and the rejected events by the veto are shown in blue. Edited from [175].

ergy ($E_{\text{prompt}} > 2\,\text{MeV}$ or $E_{\text{prompt}} > 3\,\text{MeV}$, the difference between the two is taken as systematic uncertainty). Finally, fast neutrons are subtracted using an estimation in the $\overline{\nu}_e$-free prompt energy region between $(10, 20)\,\text{MeV}$ which is extrapolated to the $(3, 10)\,\text{MeV}$ (or $(2, 10)\,\text{MeV}$) region assuming a flat spectrum. The resulting inefficiency is measured to be $(0.046 \pm 0.015)\%$.

MPS veto

The MPS (stands for *multiple pulse shapes*) veto is an innovative cut against fast neutrons which was developed after the Gd analysis was completed, so it is used for the first time in the H $\overline{\nu}_e$ selection. It relies on the fact that fast neutrons can produce multiple proton recoils within the prompt trigger digitization window. If low-energy proton recoils precede the one with the higher energy (otherwise they are obscured by the larger one), the resulting multiple pulse shapes can be observed in the PMT waveform (see figure 4.29), exploiting the information provided by the flash-ADC. This allows to assign several start times to the various pulses in the same flash-ADC window. When all the pulse start times from all Inner Detector PMTs are pictured for a single event (after correcting for the different time of flight of the light from the reconstructed vertex and the channel offset), the preceding proton recoils appear as short start times to the left of the peak given by the larger energy deposition, which appears shifted to longer start times (see figure 4.30). The magnitude of this shift is quantified by fitting a Gaussian function with mean μ_{MPS} and standard deviation σ_{MPS} to the peak of the distribution, defining a shift estimator as

$$\Delta_{\text{MPS}} = \mu_{\text{MPS}} - 1.8 \cdot \sigma_{\text{MPS}}. \tag{4.17}$$

A prompt trigger is considered to be a fast neutron and thereby rejected if

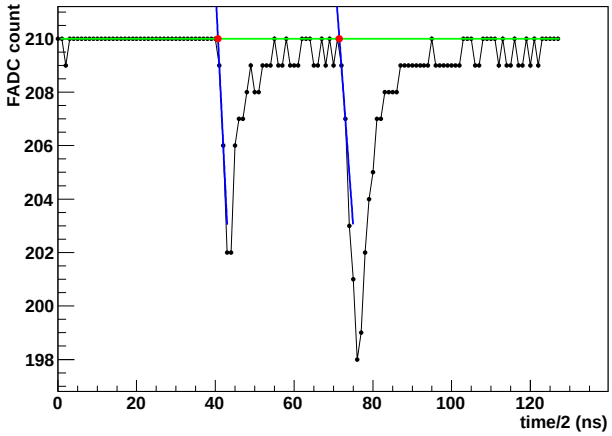


Figure 4.29: Example of one PMT waveform showing 2 pulses within the flash-ADC digitization window. The green horizontal line marks the reference baseline. The blue oblique lines show the results of linear fits to the leading edges of the pulses. The interceptions of the oblique lines with the horizontal line determine the start times of the pulses (marked as red dots). From [176].

it satisfies

$$E_{\text{prompt}} \geq 3\,\text{MeV} \text{ and } \Delta_{\text{MPS}} > 5\,\text{ns} \tag{4.18a}$$

or

$$1.2\,\text{MeV} < E_{\text{prompt}} < 3\,\text{MeV} \text{ and } \Delta_{\text{MPS}} > 5\,\text{ns} \\ \text{and is not tagged as ortho-positronium} \tag{4.18b}$$

The cut at 5 ns on Δ_{MPS} was tuned to minimize the rejection of $\overline{\nu}_e$ events. It was established analyzing the ^{60}Co calibration data, which has negligible fast-neutron contamination. The ^{60}Co source, upon β^- decay, releases two sequential γ-rays with energies 1.173 and 1.333 MeV [38] totaling $\approx$ 2.5 MeV, which is close to the mean prompt energy of the IBD events.

The energy dependence of the MPS veto is introduced to take into account the formation of ortho-positronium (o-Ps) by the positron emitted in the IBD reaction. o-Ps is the triplet state (^{3}S$_1$) of the positronium, and has a $\sim$ 45% formation fraction in the Double Chooz scintillators, with a mean lifetime of $\sim$ 3.4 ns [5] [178]. o-Ps events also produce multiple pulse shapes, in which the first pulse is given by the positron ionization and the second pulse corresponds to the o-Ps decay into γ-rays totaling 1.022 MeV. In order to minimize the loss of IBD events, it is required that the o-Ps tagging algorithm (described in reference [178]) does not identify the event as o-Ps. This algorithm tries to fit the distribution of pulse start times to two time profiles, the first one

[5] The mean lifetime of o-Ps in vacuum is 142 ns, but the interaction with the surrounding medium accelerates the decay. There is another positronium state, the singlet (^{1}S$_0$) known as para-positronium, but it is too short lived ($\mathcal{O}$(ps)) to be relevant here.

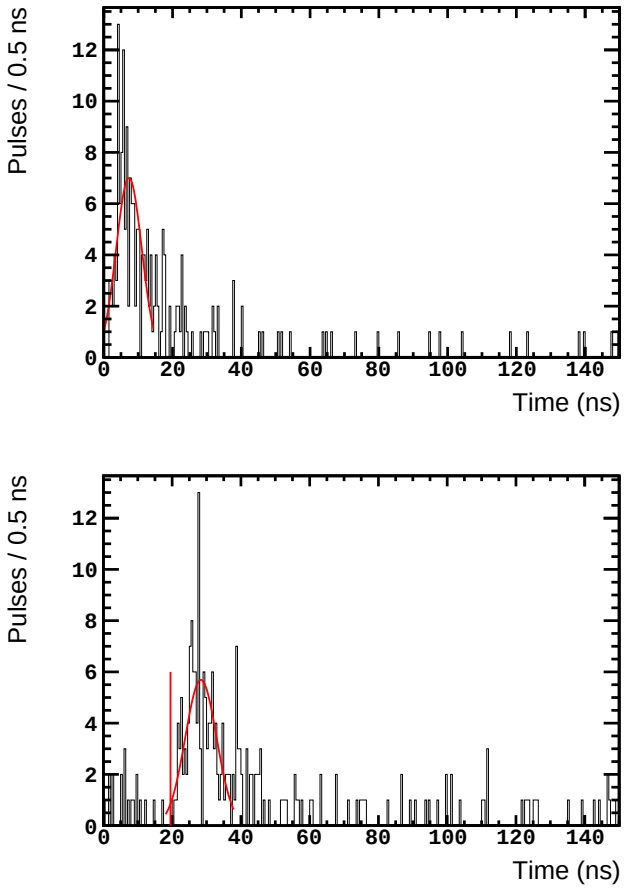


Figure 4.30: Distribution of pulse start times from all Inner Detector PMTs for an IBD-like prompt trigger (top) and a fast-neutron-like prompt trigger (bottom). In order to compare different events, the earliest pulse sets the origin of start times. The result of the Gaussian fit described in the text to estimate the shift induced by the multiple pulses is shown in red. In the IBD-like case, no shift is observed and the shift estimator Δ_{MPS} (eq. 4.17) is negative. In the fast-neutron-like case, a clear shift is observed, with $\Delta_{\mathrm{MPS}} = 19\,\mathrm{ns}$ (marked by a red vertical line). The pulses before the main peak are interpreted as low-energy proton recoils. From [177].

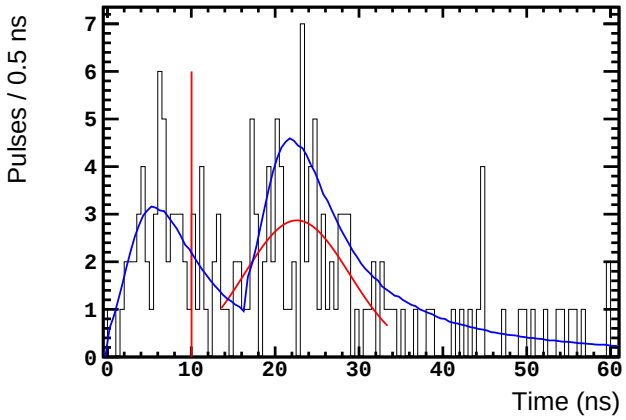


Figure 4.31: Distribution of pulse start times from all Inner Detector PMTs for a prompt trigger with visible energy $E_{\text{prompt}} = 1.67\,\text{MeV}$ which has been tagged as ortho-positronium. The blue line shows the successful fit to the two profiles of the ortho-positronium tagging algorithm: the first peak is given by the ionization induced by the positron ($\sim 0.6\,\text{MeV}$) and the second peak is given by the ortho-positronium decay releasing 1.022 MeV. The Gaussian fit used by the MPS veto to estimate the shift is shown in red, which finds $\Delta_{\text{MPS}} = 10\,\text{ns}$ (marked by a red vertical line). If the MPS veto did not contemplate the possibility of ortho-positronium formation (eq. 4.18b), this would lead to the rejection of this event. From [177].

describing a positron ionization in the range of $0.2 - 2\,\text{MeV}$ and the second one corresponding to the positron annihilation releasing 1.022 MeV, with a floating temporal separation between the two (which is null in case of no o-Ps is formed). An example of an event tagged as o-Ps is displayed in figure 4.31. The o-Ps identification only needs to be applied to prompt triggers with energies below 3 MeV because for higher energies the ionization peak is too large and hides the annihilation peak with its tail, so the two shapes cannot be distinguished. The o-Ps tagging does not work for energies below 1.2 MeV because the ionization pulse is too small ($< 200\,\text{keV}$) to be properly identified, hence it cannot be told whether the event is fast neutron or o-Ps.

The MPS veto rejects $\sim 25\%$ of the fast neutron background. This is estimated from the number of events rejected with a prompt energy greater than 12 MeV (where fast neutrons are the dominant background). The same rejection factor is found from the fraction of OV-tagged events (IBD candidates which pass all cuts but the OV veto) with a prompt-delayed trigger time interval longer than $10\,\mu\text{s}$ (to suppress the stopping muons which are also present in the sample).

The inefficiency in the $\overline{\nu}_e$ selection due to the MPS veto is estimated from the number of IBD candidates in the prompt energy region $(2, 10)\,\text{MeV}$ which pass all vetoes except the MPS one. Only the IBD candidates reconstructed in the lower half of the detector are used in order to reduce the background contamination. Furthermore, the remaining fast neutron contamination is subtracted

using the background shape described in the next paragraphs, with a normalization scaled from the number of rejected IBD candidates above 12 MeV, which are assumed to be all fast neutrons. This results in a 0.1% upper limit on the inefficiency. This upper limit was cross-checked measuring the number of events rejected by the MPS veto in the calibration data from the ^{60}Co source deployed in different positions of the Target and Gamma Catcher; and the number of delayed triggers rejected in the IBD samples using captures either on H or on Gd. In all cases, the upper limit was found to be valid.

The total inefficiency in the $\overline{\nu}_e$ selection caused by the four vetoes against fast neutrons and stopping muons (IV, OV, F_V and MPS) is $(0.10 \pm 0.20)\%$, which must be taken into account when comparing the DATA to the MC simulation, since these vetoes are not applied to it (see table 4.3).

Rate and shape estimation

In the Gd-based $\overline{\nu}_e$ selection, the prompt energy spectrum of the fast neutron and stopping muon background was modeled as a flat spectrum from the observed shape of the distribution of the IBD candidates which passed all cuts but the IV veto. However, this cannot be done for the H selection in such straightforward manner since the IV-vetoed events include also accidental background events induced by the external γ radioactivity which deposit energy in the Inner Veto before reaching the Inner Detector. Therefore, in order to obtain an accidental-background-free sample of fast neutrons and stopping muons to estimate the shape of their prompt energy spectrum, supplementary cuts on the prompt trigger are added to the condition that it must be rejected by the IV veto (equations 4.15):

$$\text{Energy in the IV} > 6\,\text{MeV} \tag{4.19a}$$

$$\text{ID-IV vertices distance: } 1.1 < \Delta R_{\text{ID}-\text{IV}} < 3.5\,\text{m} \tag{4.19b}$$

$$\text{ID-IV time interval: } -80 < \Delta T_{\text{ID}-\text{IV}} < -20\,\text{ns} \tag{4.19c}$$

The requirement of an energy deposition larger than 6 MeV in the Inner Veto practically eliminates the natural radioactivity contamination. As a result, the cuts on the vertices distance and the time interval between the Inner Detector and Inner Veto signals are updated.

The events that satisfy both the IV veto and the previous cuts are said to be *IV-tagged*. The prompt energy spectrum of these is shown in figure 4.32. In order to modelize the shape to be used in the Rate+Shape analysis, several functions were tried: constant (as in the Gd case), linear and exponential. To enhance the discrimination between the different possibilities, an extended IBD selection up to 60 MeV was performed (by raising to that value the muon veto of section 4.2.1.1 and the maximum prompt trigger energy in equation 4.14a). The best fit was found using the function

$$\frac{\mathrm{d}N_{\text{IV-tagged}}(E_{\text{prompt}})}{\mathrm{d}E_{\text{prompt}}} = p_0 \cdot \exp\left(-p_1 \cdot E_{\text{prompt}}\right) + p_2 \tag{4.20}$$

with the parameters $p_0 = 12.52 \pm 1.36$ events/MeV, $p_1 = 0.042 \pm 0.015\,\text{MeV}^{-1}$, $p_2 = 0.79 \pm 1.39$ events/MeV, and a χ^2/d.o.f. = 85.08/57.

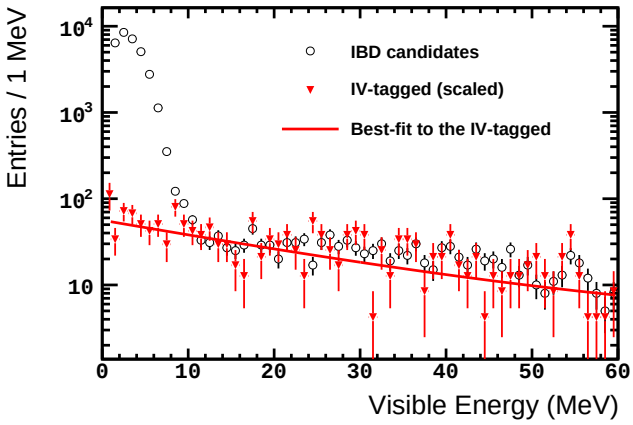


Figure 4.32: Prompt energy spectrum of the IV-tagged events (down-pointing red triangles, selection described in the text) used to estimate the shape of the fast neutron and stopping muon background. The IV-tagged events are normalized to the number of IBD candidates (empty black circles) in the interval $[20, 40]$ MeV obtained from an extended IBD selection up to 60 MeV with a relaxed muon veto and prompt energy cut. The red line shows the scaled best-fit function (eq. 4.20) used to parametrize the shape of the IV-tagged events. From [179].

The rate of the fast neutron and stopping muon background which remains in the sample is estimated by integrating the function of equation 4.20 over the $(1, 20)$ MeV interval defined by the prompt energy cut (equation 4.14a) and dividing by the live-time. The result found is corrected for the efficiency of the IV-tag selection by multiplying it by a factor (0.2363 ± 0.0147) which is computed as the ratio of the number of IBD candidates in the extended range $(20, 60)$ MeV to that of IV-tagged events. The final result amounts to 1.55 ± 0.15 events/day for the reactor-on period. The reactors-off period has a slightly lower rate, 1.45 ± 0.20 events/day, since the Outer Veto was fully operative by then.

A consistent rate and spectrum shape are found using the *OV-tagged* events (events which pass all vetoes except the OV veto).

4.2.3.2 Cosmogenic isotopes

The role of the cosmogenic isotopes ^{9}Li and ^{8}He as IBD impostors through their βn decay mode has been already discussed in the Gd selection (section 4.1.3.2). Since the neutron emitted can be also captured on hydrogen, they are also a background source for the H-based $\overline{\nu}_e$ selection. Therefore, similar strategies to reduce the number of these background events and to estimate the rate and the spectrum of the surviving contamination are followed.

Li+He veto

In order to reject some of the cosmogenic background events in the IBD candidates, the Gd selection features a cut on a likelihood which estimates the

probability that the IBD event is due to a ^{9}Li/^{8}He decay. This likelihood is computed for each prompt trigger using the distance between a previous muon track and the reconstructed vertex and the number of neutron captures within 1 ms after that muon (the likelihood is computed for all the muons in the 700 ms preceding the prompt trigger and the maximum value is retained). Since the probability density function of neutron captures used in the likelihood was built using both captures on H and on Gd, the likelihood can be readily applied to the H-based $\overline{\nu}_e$ selection without any modification. The Li+He veto value is kept at 0.4, which means that any prompt trigger with a likelihood value bigger than it will be rejected.

This veto causes some inefficiency in the $\overline{\nu}_e$ selection which must be taken into account when comparing the DATA and the MC simulation, where this cut is not applied. Analogously to the Gd case, a sample of off-time muons and IBD candidates is assembled, in which the IBD prompt trigger happens at least 2 s after the muon. The inefficiency is measured as the fraction of these muon-IBD candidate pairs which are rejected by the Li+He veto, and it is found to be $(0.508 \pm 0.012)\%$, similar to the Gd one.

Rate and shape measurement

The same prompt energy spectrum of the ^{9}Li/^{8}He background which has been used in the Gd-based $\overline{\nu}_e$ selection (see figure 4.14) is used here, except that only the part above 1 MeV is considered due to the tighter prompt energy cut in the H-based selection (equation 4.14a). This spectrum is obtained from the IBD candidates, with the neutron captured either on H or on Gd, which pass all cuts but are rejected by the Li+He veto. The background spectrum due to the non-cosmogenic rejected events (mainly $\overline{\nu}_e$) is measured from a sample of IBD candidates paired with an off-time muon (occurring at least 10 s before the prompt trigger) and subtracted.

The estimation of the rate of the ^{9}Li/^{8}He events which contaminate the final IBD sample is done similarly to the Gd case. A fit to the distribution of time intervals between the prompt trigger and previous muons is used to determine the rate (intervals up to 20 s are counted in order to measure the rate of accidental muon-prompt trigger coincidences as well). The ^{9}Li/^{8}He events appear as an excess in the distribution at short times, which decreases exponentially with a time constant given by the lifetimes of the nuclei (257.2 ms for ^{9}Li and 171.8 ms for ^{8}He [38]). Two measurements of the background rate are made using different muon cuts to obtain an upper and a lower limit, which are then combined to improve the precision of the measurement. The background rates are first estimated without the Li+He veto applied, and then the rate of vetoed events is subtracted to get the final rate.

The lower limit on the rate is set from the analysis of a cosmogenic-enriched sample which is selected from muons with an energy $E_\mu > 400\,\mathrm{MeV}^*$ and one or more neutrons in the following 1 ms, and muons with an energy $E_\mu > 500\,\mathrm{MeV}^*$ which are not followed by neutron captures in the subsequent 1 ms but have a distance between the muon track and the prompt vertex smaller than 100 cm. A fit to the distribution of time intervals between the prompt trigger and previous muons (see figure 4.33) yields 2.26 ± 0.15 events/day.

The upper limit on the rate is obtained from the analysis of muons with energies larger than 20 MeV. Among these, only the sample with $E_\mu > 600\,\mathrm{MeV}^*$

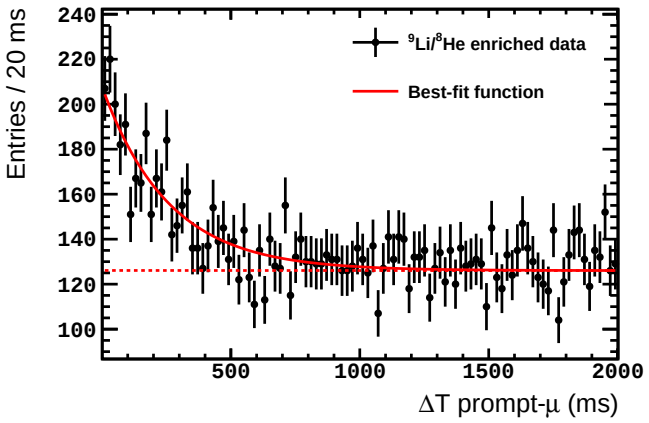


Figure 4.33: Distribution of time intervals between a muon and a prompt trigger which is used to set a lower bound on the cosmogenic background rate. The black circles show the data (error bars show the statistical uncertainty). The excess at short times corresponds to the decays of the cosmogenic background. The best-fit to a sum of two exponential functions and a constant term is shown (red solid line). The exponentials describe the radioactive decay of ^{9}Li and ^{8}He (the decay constants are fixed at 257 ms and 172 ms, the ^{9}Li and ^{8}He lifetimes, respectively) and the flat component (red dashed line) represents the accidental muon-prompt trigger coincidences. From [180].

is pure enough not to require an additional cut on the distance between the muon track and the reconstructed prompt vertex, ΔR_μ. For the others, a cut $\Delta R_\mu < d_{\max}$ is used, where $d_{\max}$ is the maximum distance allowed. In order to take into account the inefficiency due to this cut, a lateral distance profile (LDP) which represents the expected separation between the cosmogenic prompt vertex and the muon track is extracted using the sample with $E_\mu > 600\,\mathrm{MeV}^*$ (see figure 4.34). The LDP is fitted with an exponential function (which describes the probability of finding a cosmogenic event as a function of the distance to the muon track) convoluted with a Gaussian (to model resolution effects of the muon track reconstruction). Assuming a resolution on the prompt vertex of 10 cm, this gives a characteristic "decay length" of 49.1 ± 1.1 cm and a muon track resolution of 11.0 ± 3.3 cm which are used as inputs to a toy MC simulation to estimate the correction due to the use of the cut on ΔR_μ and the finite detector size. The value of $d_{\max}$ is varied between 40 and 100 cm, and the standard deviation of the resulting rates is taken as the systematic uncertainty, yielding an upper bound of $2.74^{+0.43}_{-0.39}(\mathrm{stat}) \pm 0.23(\mathrm{syst})$ events/day.

When the previous two rates are combined, and taking into account the corrections and systematic uncertainties due to the variations of the cut on the distance to the muon track, the "decay length" parameter, the resolution of the muon reconstruction, the bin width of the time interval distribution, the possible slope in the accidental muon-prompt trigger coincidences due to the finite run length, and the contribution of ^{8}He (which corresponds to $7.9 \pm 6.5\%$ of the βn candidates, from a measurement by the KamLAND experiment [166]

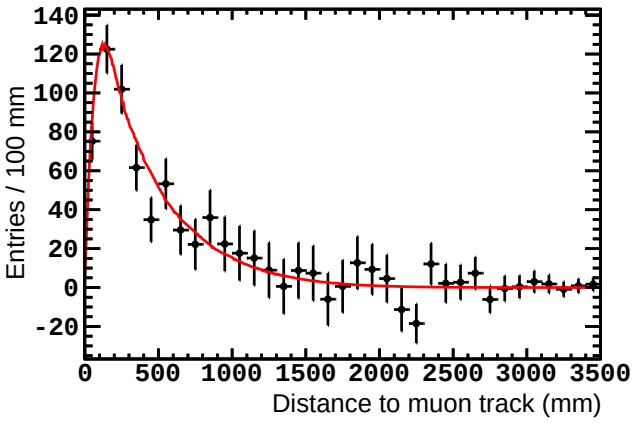


Figure 4.34: Lateral distance of the cosmogenic βn candidate prompt trigger vertex to the muon track. A signal sample is collected using muons depositing more than 600 MeV* and with a time interval to the prompt trigger within $(1.25, 500)$ ms. A background distribution due to accidental-coincidences of a muon and a non-cosmogenic event is obtained using the $\overline{\nu}_e$ MC simulation. Points show the background-subtracted data (error bars represent the statistical uncertainty). The best-fit to a model function (description in the text) is overlaid (red line). From [180].

rescaled to the Double Chooz Far Detector overburden), the resulting rate of ^{9}Li/^{8}He amounts to $2.58^{+0.57}_{-0.32}$ events/day.

In order to compute the background rate surviving the Li+He veto, the rate of the vetoed events (1.63 ± 0.06 events/day) must be subtracted. This results in a final rate of $0.95^{+0.57}_{-0.33}$ events/day.

4.2.3.3 Accidental coincidences

The estimation of the accidental background rate and spectrum follows closely that of the Gd-based $\overline{\nu}_e$ selection, which relies on the *off-time method* (see section 4.1.3.3). In this case, the prompt and delayed triggers are required to pass all cuts and vetoes of the H selection except the time coincidence of equation 4.14c, which is replaced by a version of the Gd accidental-selection cut (equation 4.5),

$$(0.5\,\mu\text{s} + T_{\text{off}}) < \Delta T < (800\,\mu\text{s} + T_{\text{off}}) \tag{4.21}$$

where the maximum time has been updated to the value used in the H selection[6]. Again, the isolation cut (equation 4.14f) must be split into two cuts

$$\begin{aligned} &\text{No valid triggers in the } 800\,\mu\text{s before} \\ &\text{and in the } 900\,\mu\text{s after the prompt.} \end{aligned} \tag{4.22a}$$

$$\begin{aligned} &\text{No valid triggers in the } 800\,\mu\text{s before the time } t_{prompt} + T_{\text{off}} \\ &\text{and only one (the delayed) in the } 900\,\mu\text{s after it.} \end{aligned} \tag{4.22b}$$

These are a version of the Gd ones (equation 4.6) with the values adapted to the H selection.

The time offset, T_{off}, is given by the same formula (equation 4.7) as in the Gd case, but in the H selection the total span of the isolation window is $T_{\text{iso}} = 1700\,\mu$s. Due to the inclusion of the natural radioactivity in the delayed energy window, the rate of the accidental background is expected to be much higher in the H selection, even after the cut on the ANN classifier value. Consequently, n is restricted to the interval $[0, 199]$; so only 200 off-time windows (instead of the 2000 of the Gd) are sufficient to collect the desired statistics.

As in the Gd estimation, the accidental background rate resulting from the off-time method must be corrected due to the different effect that some cuts have in the on-time sample. This is achieved through a series of correction factors.

Muon veto correction For the same reason explained in the Gd case, the muon veto has a higher probability to reject the delayed trigger in the off-time sample than in the on-time one. To compensate this deficiency, a correction factor, f_μ, is computed as the ratio of the probabilities of non-vetoing the delayed trigger in the on-time and off-time samples, P_μ^{on} and P_μ^{off}, respectively (removing first the events in which the muon veto rejects the prompt trigger too). So far, the procedure is identical to the Gd selection. However, unlike the Gd accidental background, in which the distribution of time intervals between triggers is flat (characteristic of random coincidences); the cut on the ANN classifier output makes this distribution non-flat in the H selection (see figure 4.25). While this has no effect on the off-time probability, which simply requires the delayed trigger to be preceded by a muon-free window as long as the muon veto time (see equation 4.8c, with $T_\mu = 1.25\,$ms); the non-flatness changes the on-time probability, in which the probability of a muon-free window is averaged over all the possible time intervals between the prompt and delayed triggers. Therefore, the expression of the P_μ^{on} probability derived for the Gd case (equation 4.8b) is modified to

$$P_\mu^{\text{on}} = \frac{\displaystyle\int_{\Delta T_{\min}}^{\Delta T_{\max}} f(t')\mathrm{e}^{-R_\mu t'}\,\mathrm{d}t'}{\displaystyle\int_{\Delta T_{\min}}^{\Delta T_{\max}} f(t')\mathrm{d}t'} \tag{4.23}$$

[6] In order to make it equivalent to the on-time selection, the time interval between prompt and delayed triggers which is fed to the ANN classifier has the time offset subtracted: $\Delta T - T_{\text{off}}$.

where $f(t')$ is a third degree polynomial fitted to the accidental background distribution of figure 4.25 to modelize its shape. Taking the integral limits to be $\Delta T_{\rm min} = 0.5\,\mu$s and $\Delta T_{\rm max} = 800\,\mu$s corresponding to the extremes of the trigger time interval cut of the IBD selection (equation 4.14c), the correction factor is found to be $f_\mu = 1.036 \pm 0.002$.

Isolation cut correction The need for this correction arises from the fact that while the on-time selection has the isolation cut applied once (equation 4.14f), the off-time selection has it applied twice (equations 4.22a and 4.22b). The calculation of the correction factor, $f_{\rm iso}$, follows that of the Gd case (see equation 4.9 and the description there), but with the corresponding variables set to the values of the H selection. This yields a correction factor $f_{\rm iso} = 1.02149 \pm 0.00007$.

Run length correction As a consequence of the finiteness of the data runs, the live-time in which on-time coincidences can be found is longer than that for off-time coincidences. In order to correct for this, the same formula as in the Gd selection is used (see equation 4.11); but in this case the averaged time offset is

$$\overline{T}_{\rm off} = \sum_{n=0}^{199} (1\,{\rm s} + n \times T_{\rm iso})/200.$$

This results in a correction factor $f_{\rm run} = 1.00036$ with negligible uncertainty.

Li+He veto correction While this is the leading correction in the Gd, it is insignificant in the H case since the contribution of the cosmogenic radionuclei to the delayed trigger of the accidental background is negligible, which is largely dominated by the natural radioactivity.

When all the correction factors are multiplied together, the total correction to the off-time accidental background rate measured is

$$f_{\rm acc} = f_\mu \times f_{\rm iso} \times f_{\rm run} = 1.059 \pm 0.002 \tag{4.24}$$

which results in a final on-time accidental background rate of 4.334 ± 0.007(stat)$\pm$ 0.008(syst) events/day, where the systematic uncertainty corresponds to the uncertainty on the correction factor.

The off-time method also provides the prompt energy spectrum of the accidental background (see figure 4.26), in which the natural radioactivity peaks at $\sim 1.4\,$MeV due to the deexcitation γ-ray emitted upon electron capture on ^{40}K and at $\sim 2.6\,$MeV due to the deexcitation γ-ray emitted upon β^- decay of ^{208}Tl (which is one of the next-to-last radionuclei of the ^{232}Th chain) are clearly observed.

Other backgrounds

As in the Gd case, the background events due to ^{12}B – ^{12}B, ^{12}B – n or the reaction ^{13}C$(\alpha, {\rm n})^{16}$O are estimated to contribute negligibly to the background contamination in the IBD selected sample.

Event type	Rate (events/day)	H-III/H-II
IBD candidates (average)	69.88 ± 0.39	—
Backgrounds		
Fast-n and stopping-μ	1.55 ± 0.15	0.49
^{9}Li and ^{8}He	$0.95^{+0.57}_{-0.33}$	0.34
Accidental	$4.334^{\pm 0.0007(\mathrm{stat})}_{\pm 0.0008(\mathrm{syst})}$	0.06

Table 4.4: Summary of the rates of the events passing the H-based $\overline{\nu}_e$ selection. The IBD candidate (including backgrounds) average rate during the periods with at least one reactor on is shown; the actual IBD rate changes in time as a function of the thermal power of the reactors as in equation 2.7. The background rates correspond to the expected rates remaining in the IBD candidate sample after the full $\overline{\nu}_e$ selection is performed, estimated from samples rejected by it as explained in the text. The column H-III/H-II shows the reduction of the background rate in the current selection [159] with respect to the previous publication [72].

When this thesis was being finalized, an additional background contribution due to double neutron captures on H was found. This background is created when two spallation neutrons created outside the detector manage to reach the Gamma Catcher and mimic an IBD coincidence there, without revealing their presence first through an energetic proton recoil in the Inner Veto, which could reject the coincidence because of the IV veto explained in section 4.2.3.1; or a proton recoil in the Inner Detector, which could make them to be rejected by the isolation cut in equation 4.14f. In the latter case, it must be noted that the isolation cut is only applied over valid triggers (i.e. not within the 1.25 μs after a muon, not tagged as light noise, with an energy larger than 0.4 MeV). This opens the possibility to avoid the rejection by the isolation cut if the proton recoil occurs in the time still vetoed by the muon veto window, or if the proton recoil looks like a light noise event (e.g. if it happens very close to a PMT) or if it has an energy below 0.4 MeV. In all this cases, what would have been a rejected triple coincidence becomes an acceptable IBD-like coincidence.

Studying triple coincidences in which the first trigger is deemed invalid for the abovesaid reasons, and after correction for the efficiency of detecting this first trigger, the rate of this background was estimated to be $0.15^{+0.05}_{-0.04}$ events/day. Given that the rate is much lower than the usual background sources already known (see table 4.4), and since the inclusion of this background in the oscillation fits of chapter 6 proved that the change on the θ_{13} measurement is not significant at all, and the Rate+Shape analysis pulls its rate even to lower values, it is not further considered for the rest of the thesis.

4.2.3.4 Reactors-off measurement

When the H-based $\overline{\nu}_e$ selection is applied to the periods when the two reactors are off, the data amounts to a live-time of 7.15 days in which 63 IBD candidates passing all cuts are found. The predicted number of IBD interactions due to residual $\overline{\nu}_e$ from the stopped reactors is 2.73 ± 0.82 according to

the simulation [170] (oscillation not included). When the expected number of IBD background events are included according to their estimated rates, the prediction raises to $50.8^{+4.4}_{-2.9}$ events, which agrees within 1.4σ with the observation, demonstrating the consistency of the background estimations.

Chapter 5

Neutron detection efficiency

In the one-detector phase, Double Chooz relies on a Monte Carlo simulation to predict the $\overline{\nu}_e$ flux at the Far Detector which will be compared to the DATA to measure the oscillation-induced deficit. Therefore, the accuracy of the simulation in reproducing the $\overline{\nu}_e$ generation and detection is mandatory in this phase of the experiment. The $\overline{\nu}_e$ detection efficiency is one of the variables of the simulation that must be ensured to match the one in the actual DATA. However, because either certain effects are not included or they are not well reproduced in the simulation, the normalization of the prediction must be corrected a posteriori, which adds an additional uncertainty source.

This chapter is dedicated to the detection efficiency of the neutron produced in the inverse β-decay: its measurement, determination of its correction factor and uncertainty. The motivation for an in-depth study results from the neutron detection efficiency being the dominant contribution to the $\overline{\nu}_e$ detection uncertainty in the Double Chooz θ_{13} measurements [69, 1].

The structure of the chapter is as follows: firstly, the $\overline{\nu}_e$ detection efficiency and the factor used to correct the Monte Carlo simulation to meet the observed date are introduced in section 5.1. Secondly, a synopsis on how the neutron physics is simulated in Double Chooz is presented in section 5.2, which helps to underline the difficulties associated to it. Thirdly, the two neutron sources used for the studies are described in section 5.3. Then, the study of the neutron detection efficiency bifurcates corresponding to the two $\overline{\nu}_e$ selections developed, the one using neutron captures on gadolinium and the one using the captures on hydrogen, which are presented sequentially in sections 5.4 and 5.5. For each one, the three components in which the neutron detection efficiency is decomposed are discussed thoroughly.

5.1 Monte Carlo correction factor

The antineutrinos emitted by the two cores of the Chooz nuclear power plant are detected through the inverse β-decay reaction: $\overline{\nu}_e + p \rightarrow e^+ + n$. As seen in chapter 2, in the null-oscillation hypothesis, the expected rate of $\overline{\nu}_e$ from a reactor R detected in one detector through the IBD reaction at time t is given

by:

$$\frac{\mathrm{d}N_R(t)}{\mathrm{d}t} = \epsilon N_p \left[\frac{1}{4\pi L_R^2} \frac{P_{\mathrm{th},R}(t)}{\langle E_\mathrm{f} \rangle_R(t)} \langle \sigma_\mathrm{f} \rangle_R(t) \right], \tag{5.1}$$

where ϵ denotes the $\overline{\nu}_e$ detection efficiency and N_p denotes the number of protons in the detector considered. The remaining variables between the square brackets are reactor-related and they have been discussed in section 2.4.1.

The Double Chooz MC simulation is designed to replicate step by step the DATA, from the generation of $\overline{\nu}_e$ in the reactor to their interaction in the detector, including the posterior conversion of the event into numerical variables that are used in the selection. Thus, DATA and MC are analyzed in an identical way. This means that there is already a built-in detection efficiency in the MC simulation, ϵ^{MC}, which can be identified with the ϵ in equation 5.1. However, in order to make the simulation outcome comparable to the actual DATA which has its own detection efficiency, ϵ^{DATA}, ϵ^{MC} must be corrected to account for known effects which have not been included in the simulation or for shortcomings in the modeling. This is achieved using a MC normalization correction factor, c_{MC} which is defined as:

$$c_{\mathrm{MC}} \equiv \frac{\epsilon^{\mathrm{DATA}}}{\epsilon^{\mathrm{MC}}}. \tag{5.2}$$

Notice that, because c_{MC} is defined as the ratio of the two efficiencies, they do not need to be absolute efficiencies (i.e. the fraction of selected $\overline{\nu}_e$ with respect to the total number of $\overline{\nu}_e$ that interacted in the detector), which in the case of the DATA is impossible to know, but they can be *relative efficiencies*, that is, the fraction of $\overline{\nu}_e$ passing the selection cuts with respect to a parent sample which satisfies looser cuts.

The signal detection efficiency can be factorized into three components corresponding to the selection steps followed in the search for IBD events: first assemble a set of valid triggers (singles), and then look for coincidences of a prompt and a delayed trigger:

$$\epsilon \equiv \epsilon_{\mathrm{single}} \cdot \epsilon_{\mathrm{prompt}} \cdot \epsilon_{\mathrm{delayed}}. \tag{5.3}$$

Each factor can be further decomposed into the intrinsic sources which cause the inefficiencies:

$$\epsilon_{\mathrm{single}} \equiv \epsilon_\mu \cdot \epsilon_{\mathrm{LN}} \tag{5.4}$$

$$\epsilon_{\mathrm{prompt}} \equiv \epsilon_{\mathrm{trig}} \cdot \epsilon_{\mathrm{iso}} \cdot \epsilon_{\mathrm{veto}} \tag{5.5}$$

$$\epsilon_{\mathrm{delayed}} \equiv f_X \cdot \epsilon_{\mathrm{sel}} \cdot k_{\mathrm{spill}}. \tag{5.6}$$

$\epsilon_{\mathrm{single}}$ denotes the efficiency of the single selection and it is divided into two factors: ϵ_μ represents the efficiency of the muon veto and ϵ_{LN} represents the efficiency of the light noise cuts.

The positron detection efficiency, $\epsilon_{\mathrm{prompt}}$, consists of three components: ϵ_{trig} corresponds to the trigger efficiency. Because the minimum energy of the prompt triggers is set where the trigger has a 100% efficiency with insignificant uncertainty, this factor is trivial but it is written for completeness. ϵ_{iso} corresponds to the efficiency of the isolation cut. ϵ_{veto} corresponds to the total efficiency of

the background vetoes: Li+He, OV, IV, F_V[1] and MPS (only in the H analysis).

The neutron detection efficiency, $\epsilon_{\text{delayed}}$, is factorized into three components: the gadolinium or hydrogen fraction (depending on the nuclei used to capture the neutron) f_X, where X denotes Gd or H, a neutron selection efficiency, ϵ_{sel}, and a neutron mobility factor, k_{spill}. There is no need to include a trigger efficiency factor in $\epsilon_{\text{delayed}}$ because, regardless the nucleus used to capture the neutron, the energy interval begins when the trigger is 100% efficient with negligible uncertainty.

The MC normalization correction factor, c_{MC}, can be factorized analogously as the signal detection efficiency.

$$c_{\text{MC}} \equiv c_{\text{single}} \cdot c_{\text{prompt}} \cdot c_{\text{delayed}}, \tag{5.7}$$

where:

$$c_{\text{single}} \equiv c_{\mu} \cdot c_{\text{LN}} \tag{5.8}$$

$$c_{\text{prompt}} \equiv c_{\text{trig}} \cdot c_{\text{iso}} \cdot c_{\text{veto}} \tag{5.9}$$

$$c_{\text{delayed}} \equiv c_{X} \cdot c_{\text{sel}} \cdot c_{\text{spill}}. \tag{5.10}$$

Since the MC simulation used for the prediction contains exclusively electron antineutrino IBD interactions; the muon, light noise, isolation and background veto cuts are unnecessary and not applied to it. Therefore, the expected $\overline{\nu}_e$ number must be corrected to account for the inefficiencies in the DATA $\overline{\nu}_e$ selection caused by these cuts. The computation of the inefficiencies and their uncertainty has already been described in chapter 4. The results are summarized in tables 4.1 (Gd selection) and 4.3 (H selection).

Regarding the neutron detection efficiency, it is the subject of this chapter. Since the detailed explanation of each component and the methods used to measure them and compute their correction factors and uncertainties are dependent on the nuclei used to capture the neutron, they are given in the dedicated section of each analysis.

5.2 Neutron physics simulation

The neutron detection efficiency correction (eq. 5.10) originates from discrepancies between the DATA and the MC simulation regarding the neutron physics. Therefore, before discussing how this correction is calculated, it is useful to describe briefly the simulation of the neutron physics in Double Chooz.

The $\overline{\nu}_e$ generator, `DCRxtrTools` [145], also generates the neutron in the detector with the corresponding momentum according to the IBD kinematics [123]. Then, the interaction with the detector components is simulated using the `Geant4`-based software, `DCGLG4sim` (section 2.4.2), in which the detector geometry and materials have been implemented.

The neutron resulting from an IBD interaction is produced with a kinetic energy in the range of tens of keV, which is quickly reduced mainly by elastic scattering off protons (H nuclei, which have a relative abundance of 65.3% in the

[1] Although the F_V veto uses delayed variables, and in the H analysis there is also a IV veto to the delayed trigger, they are grouped together with the other background vetoes on the prompt trigger for clarity and because they do not rely on the neutron physics modelization.

Target and 67% in the Gamma Catcher scintillators [181]). The loss of the neutron kinetic energy in these elastic scatterings is initially very efficient because the neutron and proton masses are similar, and the protons can be considered at rest. When the neutron approaches the thermal regime, protons can no longer be considered at rest, and the thermalization process is characterized by a time constant $\mathcal{O}(\,\mu\text{s})$.

When the neutron kinetic energy falls below 4 eV, the interaction with H nuclei, which so far has been simulated using `Geant4`'s `NeutronHPElastic` (which treats the H nuclei as a free gas of protons), is substituted with the model used in `NeutronTH`. `NeutronTH` is the Double Chooz custom implementation of the low energy neutron scattering [182, 183] and is based on the parametrization of experimental data following the effective analytical model of reference [184], which takes into account the effect of the molecular bounds, which cannot be neglected anymore since the neutron energies are of the same order as the molecular binding energies. This causes a rise in the cross section and makes the energy loss mechanism to become less efficient, since the neutron interacts with the whole molecule rather than with a single H nucleus. The increase in the cross section, which also changes the angular distribution of the scattered neutron, decreases the neutron mean free path, causing the ΔR (cf. figure 4.8) distribution to fall faster. The molecular bonds are modeled as in dodecane, which is a representative molecule since dodecane and dodecane-like molecules are the major constituents of the scintillator. When the neutron is finally thermalized, the capture is governed by a diffusion process in which the characteristic time depends on the scintillator considered: $\sim 30\,\mu$s for the Target and $\sim 200\,\mu$s for the Gamma Catcher. In figure 5.1, the radiative neutron capture cross section is shown as a function of the neutron kinetic energy for the two Gd isotopes with the largest cross section, ^{157}Gd and ^{155}Gd (which have an isotopic abundance of 15.7% and 14.8%, respectively), together with the ones for ^{12}C and ^{1}H. Since the cross section is inversely proportional to the neutron energy, the increase in its recoil energy caused by taking into consideration the H molecular bounds is translated into a longer neutron capture time, which shifts the maximum of the ΔT distribution (cf. figure 4.7) to higher values.

Eventually, the neutron will be captured. A thermal neutron (kinetic energy $\sim 0.03\,$eV) within the Target will be preferentially captured on a Gd nucleus as a result of the higher cross section ($2.54 \cdot 10^5\,$b for ^{157}Gd, $6.09 \cdot 10^4\,$b for ^{155}Gd), or on a ^{1}H nucleus in which the lower cross section (0.33 b) is partially compensated with its larger relative abundance $N_{\mathrm{H}}/N_{\mathrm{Gd}} = 1.72 \cdot 10^4$; whereas in the Gamma Catcher, the neutron will be almost exclusively captured on H. The capture on ^{1}H yields a single γ ray with an energy of 2.22 MeV, and the capture on Gd yields multiple γ rays totaling $\sim 8\,$MeV. The energy distribution of these γ rays depends on the structure of the energy levels of the Gd nucleus which captured the neutron, which are populated differently according to the neutron incident energy. This is specially relevant for neutron energies below 300 eV, for which gadolinium exhibits resonances in the capture cross section (see fig. 5.1) corresponding to different structures of excited levels. `NeutronTH` takes this into consideration by substituting the `Geant4` distribution of γ rays which is used above that energy with an experimental γ ray emission spectrum from the `ENSDF` database (see figure 5.2) which is used as a proxy for all the energies below 300 eV, since the lack of data and the small variations for low energy neutrons do not advocate a more complex treatment. Despite the radiative capture on

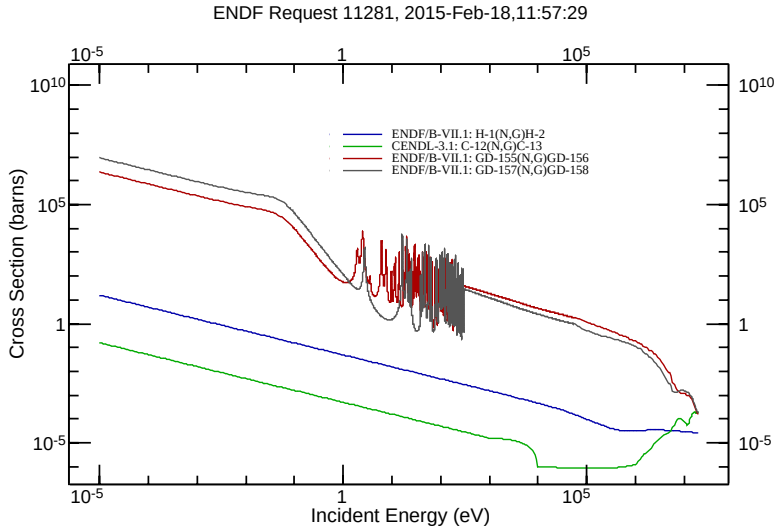


Figure 5.1: Neutron capture cross section data as a function of the neutron kinetic energy for ^{157}Gd (gray), ^{155}Gd (red), ^{1}H (blue) and ^{12}C (green). From [185]

^{12}C represents a minuscule fraction of the neutron captures, the energies and intensities for that reaction are also modified (cf. figure 5.3).

5.3 Neutron sources

The study of the gadolinium and hydrogen fractions and the neutron selection efficiencies were carried using two neutron sources, the $\overline{\nu}_e$ and the ^{252}Cf, which are described next.

5.3.1 Electron antineutrino source

Electron antineutrinos are a source of neutrons in the detector through the IBD process. The major advantage of these neutrons is that they are produced homogeneously inside the detector, so they are especially appropriate for volume-wide studies of the detector performance.

In the MC, the $\overline{\nu}_e$ simulation is used as the neutron source; while in the DATA the observed IBD candidates are used. The DATA sample is known to include backgrounds, which in any case have smaller rates than the $\overline{\nu}_e$ rates in the Gd and H analyses. Nevertheless, it must be noticed that, in contrast to the oscillation analyses in which only the $\overline{\nu}_e$ are signal; in this analysis the signal are the neutrons, disregarding if they actually come from IBD interactions. Thus,

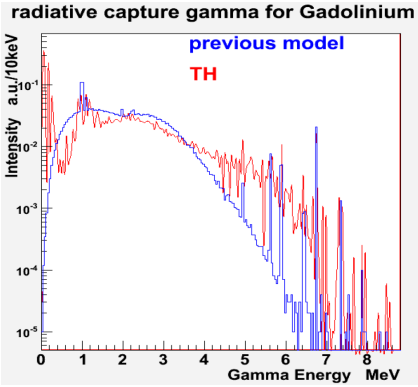


Figure 5.2: Intensities of the emitted γ ray energies in the radiative capture of a neutron by a Gd nucleus in `Geant4` (blue) and `NeutronTH` (red). From [182].

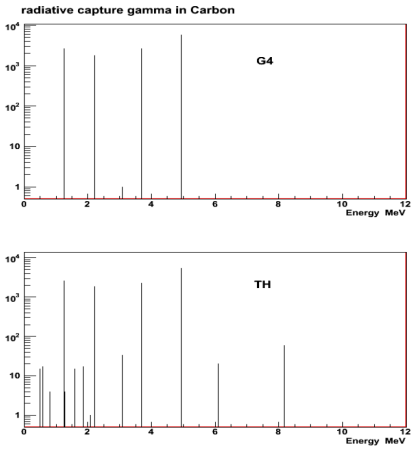


Figure 5.3: Energies and intensities of the emitted γ rays in the radiative capture of a neutron by a ^{12}C nucleus in `Geant4` (top) and `NeutronTH` (bottom). From [182].

the cosmogenic βn emitters (^{9}Li and ^{8}He) are part of the signal, so do the fast neutrons. Therefore, the background is defined as the trigger coincidences in which the delayed trigger is not a neutron. Then, the main background sources are two: the stopping muons and the majority of the accidental background.

In the accidental background the neutron captures represent a small fraction ($< 5\%$) of the delayed triggers [168]. Because the off-time method (sections 4.1.3.3 and 4.2.3.3) used to measure the accidental background can collect an arbitrarily large number of events through the use of multiple coincidence windows, this background is measured with high precision and it can be subtracted with negligible statistical uncertainty. In the analysis of the neutron detection efficiency the number of coincidence windows used is the same as for the oscillation analyses: 2000 windows for Gd, and 200 for the H (the Gd number being bigger to compensate the lower rate of accidentals in the Gd selection).

The stopping muon contamination, although greatly reduced by the combined action of the OV, IV and F_V vetoes, will remain as the dominant background. Because there is no way of measuring this background with enough precision, its effect is estimated thanks to its tendency to accumulate in the shallower part of the detector, especially below the chimney (which is its natural entrance due to the lack of active background rejection there, see figure 4.11), and it is accounted for as a systematic uncertainty in the correction factor.

The event selection is very similar to the ones for the oscillation analysis and, in fact, the single definition is identical to the ones in sections 4.1.1 and 4.2.1. In the DATA sample, to avoid including gratuitously background, only the runs in which at least one reactor was operative are used, so the two reactors off period is excluded, totaling a live-time of 460.67 days for the Gd analysis and 455.57 days for the H analysis; while the full MC $\overline{\nu}_e$ sample is used. Furthermore, the extension of the prompt energy window up to 20 MeV to accumulate background

events to help to constrain the background rate in the Rate+Shape oscillation analysis is not useful in this study, so the prompt energy windows are restricted to the interval in which antineutrinos are expected: $0.5\,\mathrm{MeV} < E_{\mathrm{prompt}} < 8\,\mathrm{MeV}$ for the gadolinium selection and $0.7\,\mathrm{MeV} < E_{\mathrm{prompt}} < 9\,\mathrm{MeV}$ for the hydrogen selection. The respective Li+He, OV, IV, F_V [2] vetoes are maintained, and an additional F_V cut is imposed on the prompt trigger, $F_V < 5.8$, to increase the rejection of stopping muons (this cut cannot be used for the oscillation analysis because it distorts the prompt energy spectrum used to measure θ_{13}). Concerning the isolation cut, it is conserved for the study of the neutron selection efficiency in the gadolinium selection, but it is enlarged to $(-800\,\mu\mathrm{s}, +900\,\mu\mathrm{s})$ for the Gd fraction, and to $(-1000\,\mu\mathrm{s}, +1000\,\mu\mathrm{s})$ in the hydrogen selection to allow the use of longer coincidence times. The values of the ΔT, E_{delayed}, ΔR (and ANN) cuts depend on the efficiency or fraction being measured, so they are reported below when they are introduced.

In order to obtain the accidental background sample to subtract it from the DATA, the same selection cuts are applied except the time coincidence, which is obtained from the off-time method already presented in the selection for the oscillation analysis. Generically, the Gd or H fraction, or the neutron selection efficiencies in the Gd-based and H-based $\overline{\nu}_e$ selections can be computed as

$$\begin{aligned}\mathcal{E} &= \frac{k_{\mathrm{on}}(\Delta T, E_{\mathrm{delayed}}, \Delta R) - f_{\mathrm{acc}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)\frac{1}{w_{\mathrm{acc}}} k_{\mathrm{off}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)}{N_{\mathrm{on}}(\Delta T', E'_{\mathrm{delayed}}, \Delta R') - f_{\mathrm{acc}}(\Delta T', E'_{\mathrm{delayed}}, \Delta R')\frac{1}{w_{\mathrm{acc}}} N_{\mathrm{off}}(\Delta T', E'_{\mathrm{delayed}}, \Delta R')} \\ &\equiv \frac{N(\Delta T, E_{\mathrm{delayed}}, \Delta R)}{N(\Delta T', E'_{\mathrm{delayed}}, \Delta R')},\end{aligned} \tag{5.11}$$

where in the numerator, $k_{\mathrm{on}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)$ and $k_{\mathrm{off}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)$ are the number of IBD candidates passing a set of cuts on $\{\Delta T, E_{\mathrm{delayed}}, \Delta R\}$ in order to select neutron captures on the nucleus chosen or match the selection cuts used in the oscillation analysis in the standard (also known as on-time) and off-time coincidence samples, respectively. In the denominator, $N_{\mathrm{on}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)$ and $N_{\mathrm{off}}(\Delta T, E_{\mathrm{delayed}}, \Delta R)$ are the number of IBD candidates passing a looser set of cuts on $\{\Delta T', E'_{\mathrm{delayed}}, \Delta R'\}$ in the corresponding on-time and off-time coincidence samples. Hence, the events selected by the cuts $\{\Delta T, E_{\mathrm{delayed}}, \Delta R\}$ are a subset of the events selected by $\{\Delta T', E'_{\mathrm{delayed}}, \Delta R'\}$. Because of the multiple coincidence windows used in the off-time method, the off-time samples must be scaled down by the total number of windows used, w_{acc}. Furthermore, as it is explained in sections 4.1.3.3 and 4.2.3.3, there are some cuts that do not affect equally the on-time and off-time samples, and this difference must be corrected using the accidental correction factor, f_{acc} which is a function of the ΔT, E_{delayed}, ΔR cuts used, as equations 4.8−4.13 and 4.23−4.24 show.

In what follows, the lengthy expressions in the numerator and the denominator will be abbreviated as in the second line of equation 5.11, indicating explicitly the limits on the ΔT, E_{delayed}, ΔR cuts used. In this case, $N(\Delta T', E'_{\mathrm{delayed}}, \Delta R')$ denotes the number of accidental-background-subtracted IBD candidates passing the loose cuts, and $N(\Delta T, E_{\mathrm{delayed}}, \Delta R)$ the corresponding number passing the tight cuts. Since passing the tight cuts can be viewed as a pure binomial process, the Neyman confidence interval [186] of the estimated efficiency

[2] In the H selection, an earlier version of the F_V cut is used, $E_{\mathrm{delayed}} > 0.2208 \cdot \exp(F_V/1.818)$, instead of the official cut, $E_{\mathrm{delayed}} > 0.2755 \cdot \exp(F_V/2.0125)$ which was fixed after the conclusion of this study. The impact on the neutron detection efficiency results will be negligible since the two cuts are very similar indeed.

can be computed using the exact Clopper-Pearson method [187] [3]. Whereas the Clopper-Pearson computation gives asymmetrical uncertainties on the efficiency, the correction factor in equation 5.2 is preferred to have a symmetric uncertainty for easier implementation. In most of the cases, the symmetrization is achieved simply by taking the significant figures. When this does not happen, the largest uncertainty is used to compute the symmetric correction factor uncertainty so as not to change the central value.

5.3.2 Californium-252 source

The ^{252}Cf calibration source was introduced in section 3.1.1. It is used as a prolific neutron emitter, yielding $\gtrsim 12.9$ neutrons per second. The neutron kinetic energy is Maxwellian distributed, peaked at $\sim 0.7\,\mathrm{MeV}$ [189].

The ^{252}Cf source has its own MC simulation as a part of the Double Chooz simulation, including the double encapsulation in stainless steel and the devices for deployment.

As in the case of the antineutrinos, the valid trigger selection follows the one from section 4.1.1. The prompt trigger is caused by the fission γ rays, which are required to satisfy $4\,\mathrm{MeV} < E_{\mathrm{prompt}} < 20\,\mathrm{MeV}$. In order to avoid overlapping with previous events, a veto time of $1.5\,\mu\mathrm{s}$ is opened before the prompt trigger. A clear evidence of a fission is obtained by requiring the neutron multiplicity to be > 1 (^{252}Cf typically emits 3.77 neutrons per fission [153]). These delayed triggers must pass the F_V veto from the oscillation analysis. As in the antineutrino source, the values of the ΔT, E_{delayed}, ΔR are chosen according to the efficiency or nucleus fraction being measured, so they are given in the corresponding sections. The accidental background is measured and subtracted as in equation 5.11 using an off-time method simpler than that of the antineutrinos, in which the time coincidence is shifted $4n\,\mathrm{ms}$, where n ranges from 1 to 6 to enhance the statistics.

The uncertainties on the neutron selection efficiencies and Gd or H fractions are computed similarly to the antineutrino case, but a Bayesian treatment is followed instead of the frequentist approach [189].

5.4 Neutron detection in the gadolinium selection

In this section it is described the determination of the IBD neutron detection efficiency and its correction factor for the gadolinium-based $\overline{\nu}_e$ selection. The gadolinium fraction, neutron selection efficiency and spill-in/out components which constitute this efficiency are presented in dedicated subsections. The first two components are measured duplicately using the two neutron sources just described. The last component is studied using the MC $\overline{\nu}_e$ simulation. The section ends with a summary which includes all the detection efficiency corrections and uncertainties.

[3] The calculation of the Clopper-Pearson confidence interval is implemented in the ROOT framework [172] as a function of the TEfficiency class [188].

5.4.1 Gadolinium fraction

The gadolinium fraction, f_{Gd}, represents the proportion of radiative neutron captures that occur on Gd. Its value is determined primarily by the relative Gd concentration in the liquid scintillator (0.123% by weight). The measurement of the Gd fraction, both in DATA and MC, was done using the two neutron sources described in section 5.3.

From equation 5.2, it follows that the correction factor c_{Gd} is defined as the ratio of the gadolinium fractions measured in the DATA and in the MC:

$$c_{\mathrm{Gd}} = \frac{f_{\mathrm{Gd}}^{\mathrm{DATA}}}{f_{\mathrm{Gd}}^{\mathrm{MC}}}. \tag{5.12}$$

5.4.1.1 Californium-252 measurement

The Gd fraction is only related to intrinsic properties of the liquid scintillator, so it is uniform in the Target volume. Thus, it can be measured at a specific location without loss of information. The chosen position is the detector center, where the leakage of neutrons out of the Target is negligible, and which has the highest statistics of calibration data ($\sim$ 5 hours of data).

The gadolinium fraction measured with the ^{252}Cf source is defined as:

$$f_{\mathrm{Gd},\,^{252}\mathrm{Cf}} \equiv \frac{N(3.5 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap 0 < \Delta T < 1000\,\mu\mathrm{s})}{N(0.5 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap 0 < \Delta T < 1000\,\mu\mathrm{s})}, \tag{5.13}$$

where the numerator includes the neutron captures on Gd (and also most of the ones on C, which peak at 4.95 MeV but they can be neglected since they correspond to less than 0.1% of the total captures) and the denominator includes captures on both H and Gd. When applied to the DATA and MC samples, the Gd fraction is measured to be $(85.30 \pm 0.08)\%$ and $(87.49 \pm 0.04)\%$, respectively, where the uncertainties are statistical.

The Gd-fraction correction factor computed using the ^{252}Cf source is found to be

$$c_{\mathrm{Gd},\,^{252}\mathrm{Cf}} = 0.9750 \pm 0.0011(\mathrm{stat}) \pm 0.0041(\mathrm{syst}). \tag{5.14}$$

The delayed energy spectra used for the measurement are shown in figure 5.4. As it can be seen, there is a considerable discrepancy between DATA and MC above and below the H peak, which is attributed to a source-related background (e.g. radioactive decays of the fission fragments not accounted for in the MC source simulation) [189]. This motivated the use of different definitions of $f_{\mathrm{Gd},\,^{252}\mathrm{Cf}}$ varying the energy integration limits to estimate their impact on the correction factor. Two alternatives were tested: one in which the integration is done only over the peak regions, i.e. $1.5-3$ MeV for H and $6.5-9.5$ MeV for Gd; another in which the lower energy limit in the denominator from equation 5.13 is changed to 1.5 MeV. The latter produced the largest discrepancy in the correction factor, and this discrepancy was taken as the systematic uncertainty on the correction.

The stability of the result was checked by computing a $c_{\mathrm{Gd},\,^{252}\mathrm{Cf}}$ using the data from a previous calibration campaign with fewer statistics. The value found, $0.9784 \pm 0.0017(\mathrm{stat})$, is consistent with the one in equation 5.14.

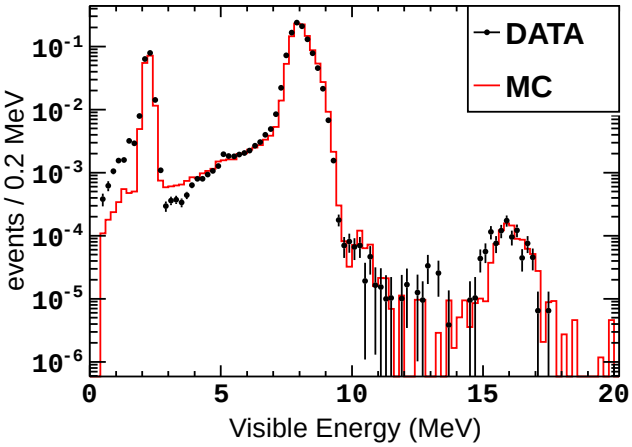


Figure 5.4: Delayed energy spectra of the ^{252}Cf source at the target center. The accidental background has been subtracted from both the DATA and MC spectra. Spectra normalized in the region 7−10 MeV. Error bars show statistical uncertainty for the DATA points. From [189].

5.4.1.2 Antineutrino measurement

Even though the Gd fraction is uniform in the Target volume, when approaching the Target boundaries border effects may induce spurious variations. Unlike the ^{252}Cf source, which can avoid the border effects by using only the data taken at the center of the detector, the $\overline{\nu}_e$ are an extended source and require a fiducial volume to be defined to isolate the captures occurred in the Target.

The dimensions of this fiducial volume must be set taking into consideration the performance of `RecoBAMA`, the reconstruction algorithm, which has a limited resolution (~ 20 cm) and a tendency to pull the vertices towards the center (for Target-edge events the shift is ~ 10 cm, cf. figure 5.5), so there may be neutron captures that occurred in the Gamma Catcher, where hydrogen is essentially the only nucleus capturing neutrons, that are erroneously reconstructed inside the Target.

In addition, it must be noted that the neutron captures are not reconstructed where they occurred but where the γ ray released interacted with the liquid scintillator. This affects especially to the single 2.2 MeV γ from the neutron capture on H, which can travel ~ 20 cm before its first interaction. On the other hand, if the prompt vertex is to be used to define the fiducial volume, it is necessary to consider the effect of the spill-out current (IBD events in which the positron is created inside the Target but the neutron escapes into the Gamma Catcher) which enriches misleadingly the number of neutron captures on H.

Because the number of $\overline{\nu}_e$ is limited, the fiducial volume must reach a compromise between purity and statistics. Several volumes were tested, and the optimal solution was found when requiring that both prompt and delayed vertices were reconstructed inside the cylinder defined by $|z| < 1062.5$ mm and $\rho < 958.3$ mm (the Target dimensions are $|z| = 1274.58$ mm and $\rho = 1150$ mm).

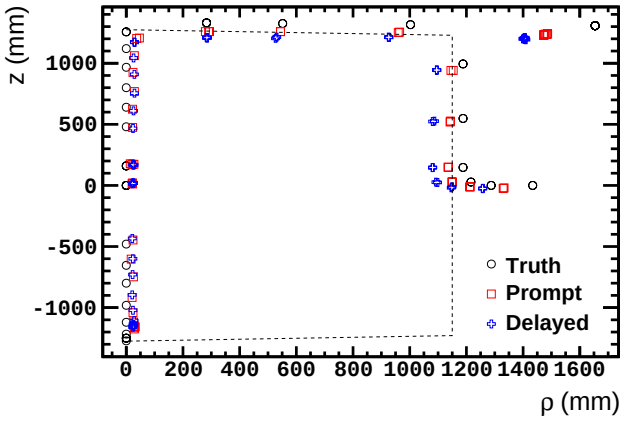


Figure 5.5: Deployment (truth) and reconstructed positions of the ^{252}Cf source DATA during the first calibration campaign. The reconstructed positions correspond to the centroid of the distributions of `RecoBAMA` vertices. Prompt triggers (red squares) are given by the fission γ rays, delayed triggers (blue crosses) are given by the neutron captures on H and Gd. The dashed line delimits the Target volume.

Since measuring the Gd fraction also requires integration over H captures, the delayed F_V cut used in the Gd selection (equation 4.4), which is tuned to remove stopping muons in the Gd energy range, must be supplemented with an additional delayed F_V cut which applies to the H range: $E_{\text{delayed}} > 0.2411 \cdot \exp(0.5223 \cdot F_V)$. In order to purify further the $\overline{\nu}_e$ from stopping muons, the DATA run list is restricted to runs which had the OV active, either in its full configuration or in the reduced one, corresponding to 84.3% of the DATA sample.

Then, the gadolinium fraction measured with the $\overline{\nu}_e$ source is defined as:

$$f_{\text{Gd, IBD}} \equiv \frac{N(3.5 < E_{\text{delayed}} < 10\,\text{MeV} \cap 0.25 < \Delta T < 200\,\mu\text{s} \cap \Delta R < 1.7\,\text{m})}{N(1.6 < E_{\text{delayed}} < 10\,\text{MeV} \cap 0.25 < \Delta T < 200\,\mu\text{s} \cap \Delta R < 1.7\,\text{m})}, \quad (5.15)$$

where the numerator and the denominator are given by the number of events satisfying the cuts between parentheses after the subtraction of the accidental background. The events passing the denominator cuts are shown in figure 5.6 binned in delayed energy. The upper values of the time and distance coincidence cuts are set so as to integrate over the full IBD distributions in the Target central volume. The lower energy limit in the denominator is established to avoid the background from the ^{212}Bi $-$ ^{212}Po decay chain, in which the delayed trigger is given by the emission of an 8.8 MeV α which results in a visible energy of about 1 MeV because of the quenching effect.

The DATA and MC Gd fractions are measured to be $(87.17^{+0.35}_{-0.36})\%$ and $(89.01 \pm 0.03)\%$ respectively, where the uncertainties are statistical. Their ratio gives the correction factor computed using the $\overline{\nu}_e$ source:

$$c_{\text{Gd, IBD}} = 0.9794 \pm 0.0040(\text{stat}) \pm 0.0044(\text{syst}), \quad (5.16)$$

where the systematic uncertainty has two components. One, amounting to

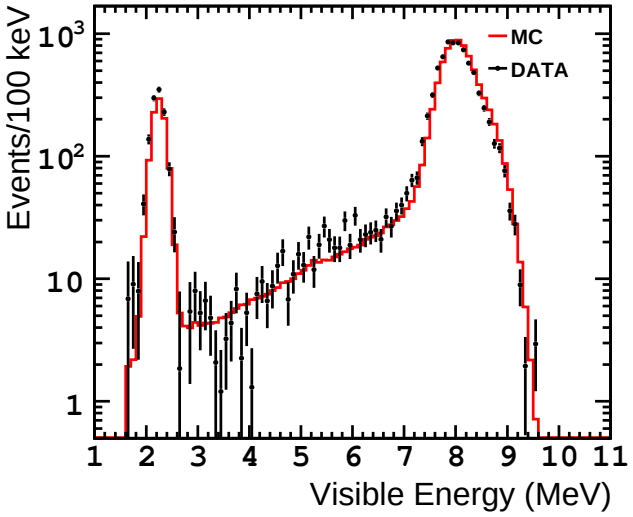


Figure 5.6: Delayed energy spectra of the antineutrinos within the fiducial volume used to compute the Gd fraction. MC normalized to the accidental background-subtracted DATA in the region from 7 to 10 MeV. Error bars show the statistical uncertainty for the DATA points.

0.0036 accounts for the uncertainty caused by the remaining stopping muon contamination in the measurement, and it is estimated as the difference between the result in equation 5.16 and the correction factor found when only events reconstructed in the bottom half of the Target are used, where this background is suppressed. The other component is analogous to the systematic uncertainty of the ^{252}Cf measurement and it refers to the uncertainty introduced by the Gd-fraction definition. An alternative correction factor is calculated in which captures on H are defined as the events between 1.6 and 3 MeV, and those on Gd are defined as the events between 6.5 and 9.5 MeV. The discrepancy with the result in equation 5.16 is 0.0025, which is added quadratically to the former component.

The two results for the gadolinium fraction correction factor obtained with the ^{252}Cf source (eq. 5.14) and the $\overline{\nu}_e$ source (eq. 5.16) are consistent with each other, although the $\overline{\nu}_e$ cannot compete with the high statistics of the ^{252}Cf source. Therefore, the ^{252}Cf value is used in the θ_{13} analysis:

$$c_{\mathrm{Gd}} = 0.9750 \pm 0.0042. \tag{5.17}$$

5.4.2 Neutron selection efficiency

The neutron selection efficiency, ϵ_{sel}, corresponds to the overall efficiency of the delayed energy, correlation time and distance cuts used to select the neutrons produced in the IBD reaction. Unlike the gadolinium fraction, this

efficiency is not uniform within the detector and it will decrease as we approach to the borders of the Target affected by the change in the composition of the scintillator between Target and Gamma Catcher, which impacts the neutrons thermalization, and the higher probability that the released γ rays from the neutron captures are not fully absorbed in the scintillator volumes.

Thus, considering the cylindrical symmetry exhibited by the detector, the selection efficiency will be a function of two variables, $\epsilon_{\text{sel}}(z, \rho)$, where z is the height and $\rho = \sqrt{x^2 + y^2}$ is the radial distance to the center of the detector, which is taken as the origin.

Nevertheless, since the MC correction factor is applied to the global MC $\overline{\nu}_e$ normalization, regardless where the event happened, the local value of the efficiency $\epsilon_{\text{sel}}(z, \rho)$ has not practical interest, and what is truly relevant is the integrated efficiency over the detector volume:

$$\epsilon_{\text{sel}} = \frac{\int_V \epsilon_{\text{sel}}(z, \rho) \mathrm{d}V}{\int_V \mathrm{d}V}. \tag{5.18}$$

Then, the neutron selection efficiency correction factor c_{sel} is written as the ratio of the volume-integrated selection efficiencies measured in the DATA and in the MC:

$$c_{\text{sel}} = \frac{\epsilon_{\text{sel}}^{\text{DATA}}}{\epsilon_{\text{sel}}^{\text{MC}}}. \tag{5.19}$$

Since the neutron selection efficiency consists of three selection cuts (delayed energy, correlation time and correlation distance), one possibility is to evaluate all the cuts at the same time. This approach is known as *inclusive*. Another possibility is to assume that ϵ_{sel} can be factorized into the individual cut efficiencies:

$$\epsilon_{\text{sel, exc}} = \epsilon_{\Delta T} \cdot \epsilon_{E_{\text{delayed}}} \cdot \epsilon_{\Delta R} \tag{5.20}$$

where $\epsilon_{\Delta T}$, $\epsilon_{E_{\text{delayed}}}$, $\epsilon_{\Delta R}$ are the correlation time, delayed energy and correlation distance efficiencies, respectively. In this approach, called *exclusive*, only the cut giving name to the efficiency is evaluated, while the others are kept fixed. As usual, the specific definitions are source dependent and they can be found below.

5.4.2.1 Antineutrino measurement

The IBD events occur in all the detector, so the neutrons produced are especially well-suited for a direct measurement of the volume-wide efficiency.

The inclusive neutron selection efficiency measured with the $\overline{\nu}_e$ source is defined as:

$$\epsilon_{\text{sel, inc, IBD}} \equiv \frac{N(0.5 < \Delta T < 150\,\mu\text{s} \cap 4 < E_{\text{delayed}} < 10\,\text{MeV} \cap \Delta R < 1\,\text{m})}{N(0.25 < \Delta T < 200\,\mu\text{s} \cap 3.5 < E_{\text{delayed}} < 10\,\text{MeV} \cap \Delta R < 1.7\,\text{m})}, \tag{5.21}$$

where the numerator represents the number of IBD events passing the selection cuts used in the gadolinium analysis (sec. 4.1.2), and the denominator corresponds to a larger number of events which satisfy a looser version of the cuts. In figure 5.7, the inclusive neutron selection efficiency can be seen as a function of

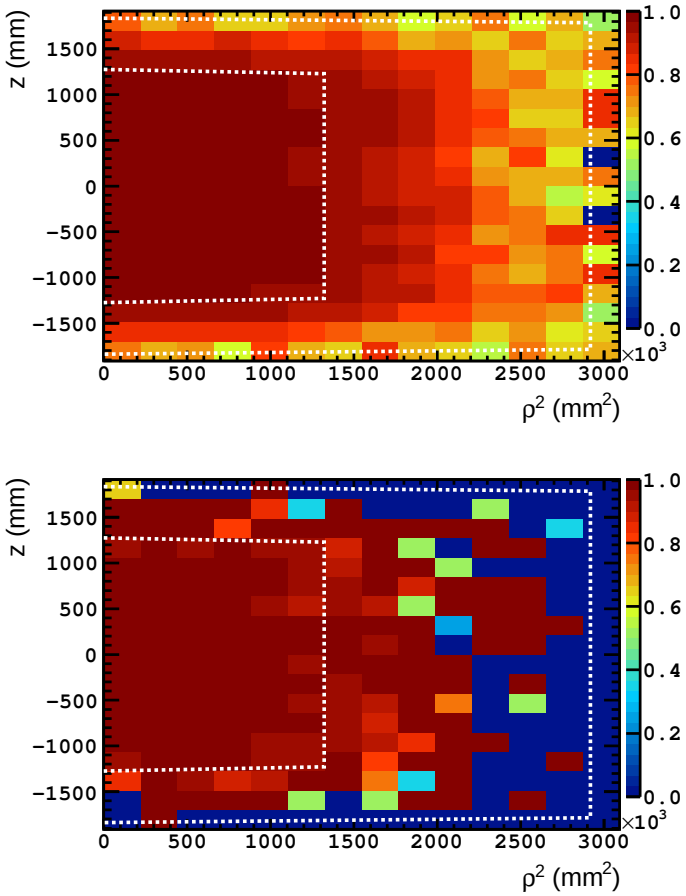


Figure 5.7: Inclusive neutron selection efficiency $\epsilon_{\text{sel, inc, IBD}}(z, \rho)$ obtained with IBD neutrons for MC (top) and accidental background-subtracted DATA (bottom). The inner and outer dashed lines demarcate the Target and Gamma Catcher vessels, respectively.

the spatial variables ρ and z. In order to compute the efficiency integrated over the Target volume, only events reconstructed within the Target ($|z| < 1250\,\text{mm}$, $\rho < 1150\,\text{mm}$) are counted.

The DATA and MC inclusive neutron selection efficiencies are measured to be $(98.58^{+0.09}_{-0.10})\%$ and $(98.618 \pm 0.009)\%$, respectively, where uncertainties are statistical. From equation 5.19, the correction factor is

$$c_{\text{sel, inc, IBD}} = 0.9996 \pm 0.0010(\text{stat}) \pm 0.0019(\text{syst}). \tag{5.22}$$

The systematic uncertainty accounts for the remaining stopping muon contamination in the sample. The correction factor was also measured counting only the IBD events reconstructed in the bottom half of the Target ($-1250 < z < 0\,\text{mm}$), where this background is suppressed. The difference found with the result using the full Target, 0.0019 ± 0.0009, was taken as the systematic uncertainty.

	$\epsilon_{\mathrm{IBD}}^{\mathrm{DATA}}$ (%)	$\epsilon_{\mathrm{IBD}}^{\mathrm{MC}}$ (%)	c_{IBD}
ΔT	$99.27^{+0.07}_{-0.07}$	$99.375^{+0.006}_{-0.006}$	0.9989 ± 0.0010 (0.0007(stat) $\pm$ 0.0007(syst))
E_{delayed}	$99.38^{+0.06}_{-0.07}$	$99.433^{+0.006}_{-0.006}$	0.9994 ± 0.0008 (0.0007(stat) $\pm$ 0.0004(syst))
ΔR	$99.94^{+0.02}_{-0.03}$	$99.813^{+0.003}_{-0.003}$	1.0013 ± 0.0003 (0.0003(stat) $\pm$ 0.00003(syst))
Exclusive	$98.58^{+0.09}_{-0.10}$	$98.627^{+0.009}_{-0.009}$	0.9996 ± 0.0013 (0.0010(stat) $\pm$ 0.0008(syst))
Inclusive	$98.58^{+0.09}_{-0.10}$	$98.618^{+0.009}_{-0.009}$	0.9996 ± 0.0021 (0.0010(stat) $\pm$ 0.0019(syst))

Table 5.1: Summary of neutron selection efficiencies and correction factors for the gadolinium analysis computed with the $\overline{\nu}_e$ source. Efficiency uncertainties are statistical only.

An alternative correction factor can be computed following the exclusive prescription. The correlation time, delayed energy and correlation distance efficiencies are defined:

$$\epsilon_{\Delta T,\,\mathrm{IBD}} \equiv \frac{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})}{N(0.25 < \Delta T < 200\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})} \quad (5.23)$$

$$\epsilon_{E_{\mathrm{delayed}},\,\mathrm{IBD}} \equiv \frac{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})}{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 3.5 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})} \quad (5.24)$$

$$\epsilon_{\Delta R,\,\mathrm{IBD}} \equiv \frac{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})}{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1.7\,\mathrm{m})}. \quad (5.25)$$

The result of computing each efficiency for the DATA and MC, together with the individual correction factor derived from their ratio can be found in table 5.1. The exclusive selection efficiencies, according to equation 5.20, are $(98.58^{+0.09}_{-0.10})\%$ for DATA and $(98.627 \pm 0.009)\%$ for MC (uncertainties are statistical). The correction factor is found to be

$$c_{\mathrm{sel,\,exc,\,IBD}} = 0.9996 \pm 0.0010(\mathrm{stat}) \pm 0.0008(\mathrm{syst}), \quad (5.26)$$

where the systematic uncertainty has been computed as in the inclusive result (eq. 5.22). The smaller systematic uncertainty in the exclusive case compared to the inclusive one is understood to come from the different denominators used in equation 5.21 and equations 5.23, 5.24, 5.25. Whereas for the inclusive efficiency all the three cuts are relaxed simultaneously, thus increasing the probability of background contamination; in the exclusive efficiency only one cut is loosened at a time, so the background acceptance is smaller.

The fact that the two results in equations 5.22 and 5.26 agree nicely demonstrates that the background is under control, and that the hypothesized factorization of eq. 5.20 can be done. However, the inclusive correction is propagated to the θ_{13} oscillation fit because the uncertainty is more conservative.

The impact of restricting the volume integration of equation 5.18 to the Target was studied by varying the integral limits, as it is shown in figure 5.8. The evolution of the correction factor is consistent with a flat trend, establishing that no bias was introduced because of the volume constraint.

Further investigations were performed to search for additional sources of systematic uncertainty. The efficiency definitions in equations 5.21, 5.23, 5.24, 5.25 were perturbed by altering the denominators, either by tightening the cuts (the upper ΔT limit was moved to $175\,\mu$s, the maximum ΔR was shortened to 1.5 m, both separately and simultaneously) or by loosening even more the cuts (the lower E_{delayed} limit was moved to 3 MeV, the maximum ΔR was enlarged to 2 m, both separately and simultaneously). All resulted in non-significant discrepancies. The reproducibility of the method was tested by repeating the calculation of the correction factors using an alternative MC simulation statistically-independent generated, finding exactly the same results. The influence of the Gd concentration on the correction factors was examined by analyzing two MC simulations produced with 960 mg/l and 1020 mg/l, which correspond to a $\pm 3\%$ variation in the standard value (989 mg/l). Both results were found to agree within statistical uncertainty with the ones obtained with the standard concentration. Finally, the hypothesis that ^{9}Li/^{8}He and fast neutron events are innocuous (since they have a neutron capture as delayed trigger) was put to the

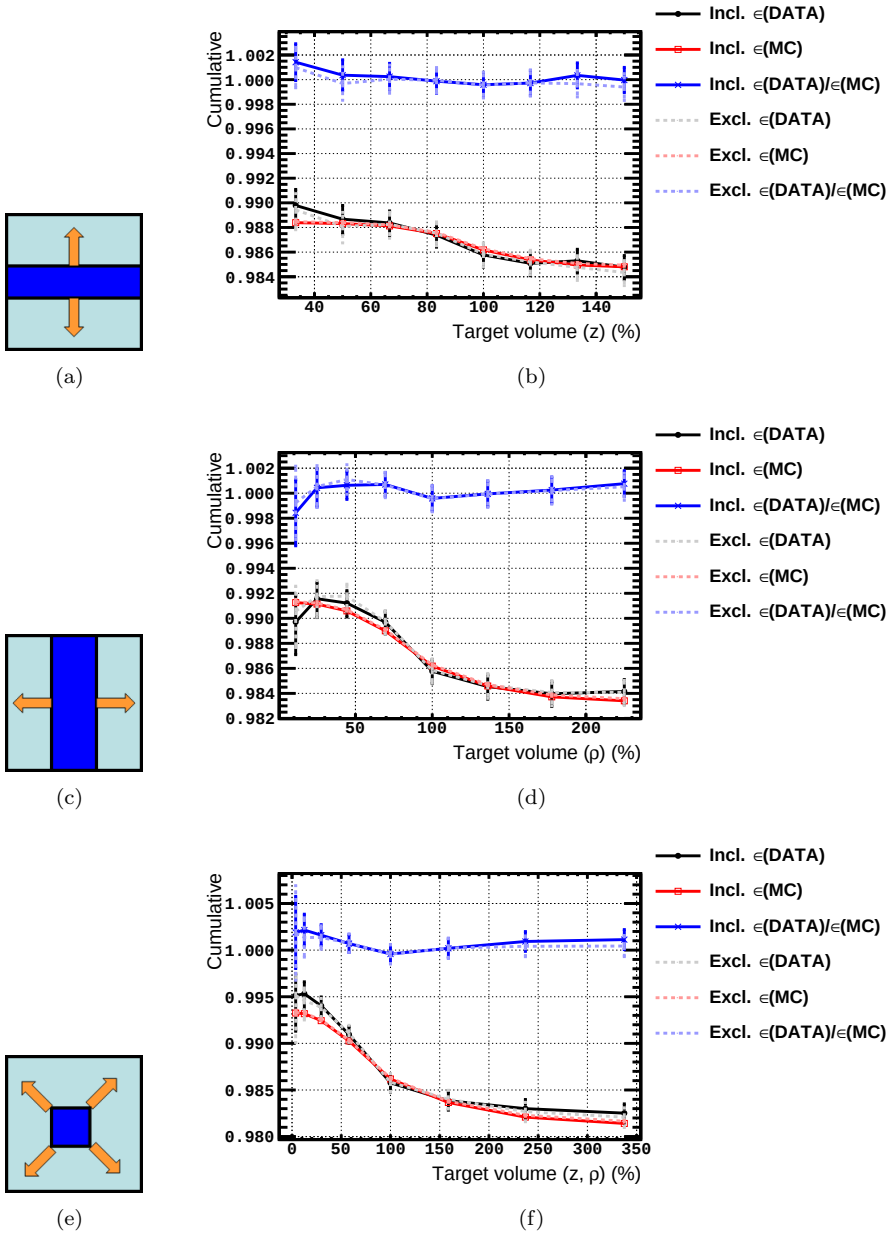


Figure 5.8: Cumulative neutron selection efficiencies and correction factors for DATA and MC as a function of the integrated volume used in their estimation. For (b), the integration limits are varied symmetrically along the z dimension as sketched in (a), keeping ρ fixed at the Target radius. For (d), the integration limit is moved along the radial dimension as sketched in (c), keeping $|z|$ fixed at the Target height. For (f), the integration limits are varied simultaneously in z and ρ dimensions as sketched in (e). The values quoted in eqs. 5.22 and 5.26 correspond to a 100% Target volume. Black (red) line shows the inclusive selection efficiency for DATA (MC). Dashed gray (dashed light red) line shows the exclusive selection efficiency for DATA (MC). Blue (dashed light blue) shows the correction factor for inclusive (exclusive) efficiency. Uncertainties are only statistical.

test by removing simultaneously the Li+He veto and the IV veto. The correction factors computed like this agree within uncertainty with the ones quoted in equations 5.22 and 5.26. For the above-mentioned reasons, any additional systematic uncertainty was deemed unnecessary.

5.4.2.2 Californium-252 measurement

The ^{252}Cf source was deployed along the detector symmetry axis (z axis) and the Guide Tube that runs on the Target vessel outer surface, traverses the Gamma Catcher and passes close to the Buffer, as shown in figure 5.9. There is an evident lack of data in the radial dimension of the Target, which poses a difficulty for a volume-wide measurement of the neutron selection efficiency. Furthermore, the efficiency along the radial coordinate is the dominant contribution to the volume integral, since the number of $\overline{\nu}_e$ interactions scales with ρ^2. In previous Double Chooz results [69, 1], the absence of radial data was circumvented by fitting the z axis and Guide Tube data at $z = 0$ with a function which played the role of the unmeasured radial efficiency with the shape

$$\hat{\epsilon}(\rho) = \frac{1}{2}\tanh\left(\frac{\rho - R_{\mathrm{T}}}{a}\right), \tag{5.27}$$

where $\hat{\epsilon}$ represents a normalized efficiency, R_{T} is the Target radius and a is a constant to be determined by the fit. This approach requires to rely on a hypothetical shape which can only be constrained at the detector center, $\rho = 0$, and outside the Target, $\rho > R_{\mathrm{T}}$ where the asymptotical behavior dominates. In addition, the same shape is postulated for the ΔT and E_{delayed} efficiencies, despite the different physics behind them. In order to solve the deficiency of calibration data in the radial coordinate, an extensible articulated arm was built which would be capable of moving the calibration sources along the radial dimension and around the symmetry axis. However, this device has not been used in the detector yet. Thus, in parallel with the development of the $\overline{\nu}_e$ measurement, a new method to compute the Target-wide efficiency based on the existing ^{252}Cf data had to be conceived, with a more accurate treatment of the radial efficiency as its major request.

Prior to delving into the details of the method, in figure 5.10 the neutron selection efficiencies for the positions along the Target axis during the second calibration campaign are shown, in which the efficiency has been defined for the ^{252}Cf source as

$$\epsilon_{\mathrm{sel,\,inc,\,^{252}Cf}} \equiv \frac{N(0.5 < \Delta T < 150\,\mu\mathrm{s} \cap 4 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap \Delta R < 1\,\mathrm{m})}{N(0.25 < \Delta T < 1000\,\mu\mathrm{s} \cap 3.5 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV})} \tag{5.28}$$

which follows the inclusive prescription. The higher statistics and signal to background ratio in the ^{252}Cf source allow the use of wider cuts in the denominator compared to the $\overline{\nu}_e$ definition (eq. 5.21).

Similar to the separation of variables technique used to solve partial differential equations, and reflecting the cylindrical symmetry of the detector, the neutron selection efficiency in any position within the detector can be written as

$$\epsilon_{\mathrm{sel}}(z, \rho) = \epsilon_0 \cdot f_1(z) \cdot f_2(\rho), \tag{5.29}$$

where $\epsilon_0 = \epsilon_{\mathrm{sel}}(z = 0, \rho = 0)$ corresponds to the efficiency at the center of the detector, and $f_1(z) = \epsilon(z)/\epsilon_0$ and $f_2(\rho) = \epsilon(\rho)/\epsilon_0$ are functions of one variable

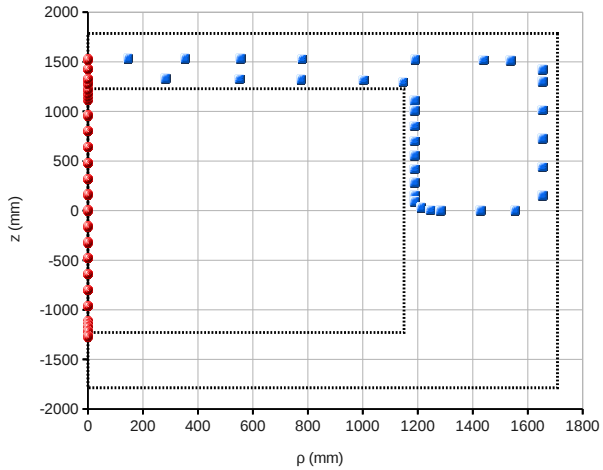


Figure 5.9: ^{252}Cf source deployment positions during the second calibration campaign. Red circles mark positions along the z axis. Blue squares mark positions in the Guide Tube. The inner and outer dashed lines delimit the Target and Gamma Catcher vessels, respectively.

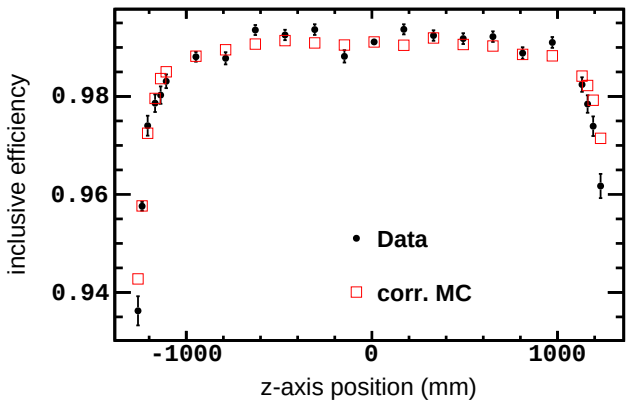


Figure 5.10: Inclusive neutron selection efficiency for the ^{252}Cf source deployment positions along the Target axis during the second calibration campaign. Black circles show DATA and empty red squares show MC after applying the correction from eq. 5.36. Error bars show statistical uncertainty. Edited from [2].

which encode the efficiency dependence along z and ρ, respectively. As stated before, we are interested in the Target-integrated efficiency, so using equation 5.29 in 5.18, we find:

$$\epsilon_{\text{sel}} = \epsilon_0 \cdot f_z \cdot f_\rho, \tag{5.30}$$

where

$$f_z = \frac{\int_{-H_{\text{T}}/2}^{+H_{\text{T}}/2} \epsilon(z)\mathrm{d}z}{\epsilon_0 \int_{-H_{\text{T}}/2}^{+H_{\text{T}}/2} \mathrm{d}z} \tag{5.31}$$

$$f_\rho = \frac{\int_0^{R_{\text{T}}} \epsilon(\rho)\rho\mathrm{d}\rho}{\epsilon_0 \int_0^{R_{\text{T}}} \rho\mathrm{d}\rho} \tag{5.32}$$

and H_{T} is the height of the Target volume.

The validity of equation 5.30 was tested using the $\overline{\nu}_e$ MC simulation, by comparing the result of a direct integration of the efficiency in equation 5.28 over the Target volume with the one obtained computing ϵ_0 from a volume defined by $|z| < 30\,\text{cm}$, $\rho < 30\,\text{cm}$; f_z from a cylinder defined by $\rho < 30\,\text{cm}$ and height equal to the Target; and f_ρ from a cylinder defined by $|z| < 30\,\text{cm}$ and radius equal to the Target. Both methods were found to agree within $(0.03 \pm 0.05)\%$ [190].

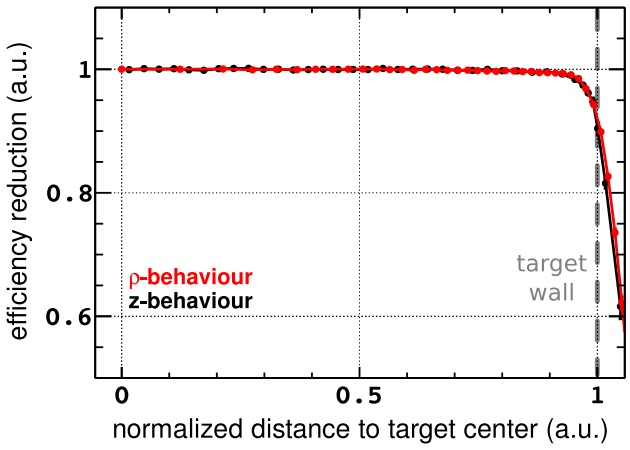
The key point of the method is to postulate that

$$f_2\left(\frac{\rho}{R_{\text{T}}}\right) = f_1\left(\frac{z}{H_{\text{T}}/2}\right) \tag{5.33}$$

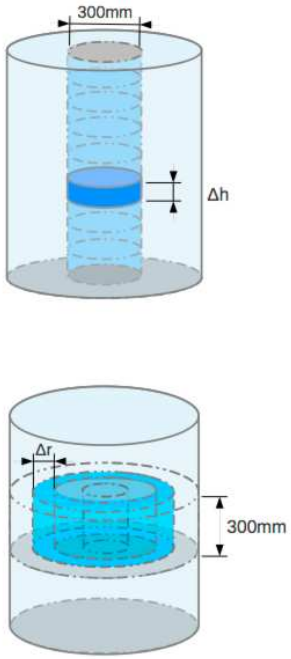
so the behavior of the efficiency along ρ, which has not been measured with calibration sources, can be inferred from the behavior along z, for which ^{252}Cf data are available. The foundation of this hypothesis lies on that, when approaching to the Target boundaries, the neutron physics cannot distinguish, to a very good approximation, between the flat surface of the Target top and bottom lids and the curved surface of the Target lateral wall. This is demonstrated in figure 5.11 using the $\overline{\nu}_e$ MC data, where the dependence of the efficiency along the radial and vertical normalized coordinates is shown to be almost equal.

Since equation 5.33, or alternatively $\epsilon(\rho/R_{\text{T}}) = \epsilon(z/H_{\text{T}}/2)$, has critical importance, an attempt to verify it with actual DATA was performed using the $\overline{\nu}_e$ source. The results can be seen in figure 5.12, which supports the hypothesis, but with the caveat that the limited statistics of the antineutrino do not make possible a precise comparison.

Performing the substitution $\rho \to z$ in equation 5.32, and taking into consideration that the ^{252}Cf source has been deployed in discrete positions along the



(a)



(b)

Figure 5.11: (a) $f_2(\rho/R_T)$ (red) and $f_1(z/H_T/2)$ (black) obtained from the $\overline{\nu}_e$ MC simulation. (b) shows how the Target volume was divided into subvolumes along z (top) and ρ (bottom) where the efficiency data points were calculated. The curves for positive and negative z values have been combined. (a) from [191], (b) from [190].

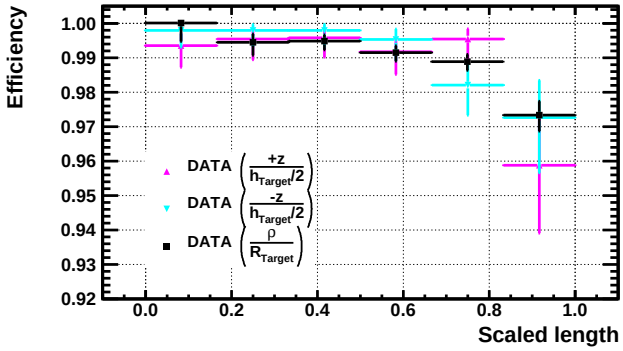


Figure 5.12: Inclusive neutron selection efficiency (eq. 5.21) measured with $\overline{\nu}_e$ DATA as a function of the scaled distance to the Target center ($z/H_{\mathrm{T}}/2$ or ρ/R_{T}). Up-pointing pink triangle and down-pointing cyan triangle show the $\epsilon(z/H_{\mathrm{T}}/2)$ efficiency for $z > 0$ and $z < 0$, respectively (both require $\rho < 575\,\mathrm{mm}$). Black squares show the $\epsilon(\rho/R_{\mathrm{T}})$ efficiency (requiring $|z| < 637.5\,\mathrm{mm}$). Horizontal error bars span the bin width and vertical error bars show the statistical uncertainty.

z axis, the equations 5.31 and 5.32 are rewritten as:

$$f_z = \frac{1}{\epsilon_0 N} \sum_{i}^{N} \epsilon(z_i) \tag{5.34}$$

$$f_\rho = \frac{1}{\epsilon_0 \frac{N'}{2} \frac{H_{\mathrm{T}}}{2}} \sum_{j}^{N'} |z|_j \cdot \epsilon(|z|_j) \tag{5.35}$$

where $\epsilon(z_i)$ corresponds to the interpolated efficiency value at z_i, and i runs over the $N = 18000$ points equidistantly distributed in the interval $-1260\,\mathrm{mm} < z < 1260\,\mathrm{mm}$ (i.e. the interpolation step is $0.14\,\mathrm{mm}$). The linear interpolation is built using the data from the deployment positions along the z axis inside the Target (figure 5.10; in the cases in which several runs have been taken at the same position, the data are combined), and it is required in order to sample the volume homogeneously, so the integrated efficiency represents faithfully the volume-wide efficiency and it is not biased by the arbitrary locations of the deployment positions [192]. Since f_ρ describes the behavior along the radial coordinate, which is non-negative, the interpolation is built after the measurements with the same $|z|$ have been combined; then $\epsilon(|z|_j)$ represents an interpolated efficiency value at z_j, where j runs over the $N' = N/2 = 9000$ points in the interval $0 < z < 1260\,\mathrm{mm}$. In order to have a statistical uncertainty associated to f_z and f_ρ, the measured data which are the seeds of the interpolation are fluctuated within their 1σ uncertainties, building a new interpolation each time. By repeating this procedure 5000 times, a distribution of results for f_z and f_ρ is obtained, which is fitted with a Gaussian, and the width is interpreted as the statistical uncertainty.

The accuracy of the interpolation was tested using the $\overline{\nu}_e$ MC simulation by comparing the results for f_z and f_ρ obtained following equations 5.31 and

	Discrete interpolation	Continuous integral
f_z	0.9976 ± 0.0006	0.9983 ± 0.0003
f_ρ	0.9951 ± 0.0008	0.9955 ± 0.0003

Table 5.2: f_z and f_ρ computed using the $\overline{\nu}_e$ MC simulation as in eqs. 5.34 and 5.35 (discrete interpolation) and as in eqs. 5.31 and 5.32 (continuous integral). Uncertainties are statistical. From [190]

5.32, where the integration is carried over cylinders such as those of fig. 5.11(b) which span completely the Target height and radius respectively, with the results obtained according to 5.34 and 5.35 (which also tests the success of the $\rho \to z$ substitution). For computing the latter pair, the $\overline{\nu}_e$ are discretized into subvolumes as in fig. 5.11(b). The agreement seen in table 5.2 confirms the goodness of the method.

Once the validity of the method has been established, it can be applied to the ^{252}Cf data. The inclusive selection efficiency, following the factorization of equation 5.30 and with f_z and f_ρ computed as in equations 5.34 and 5.35, is measured to be $(98.29 \pm 0.06)\%$ for DATA and $(98.26 \pm 0.03)\%$ for MC (uncertainties are statistical). Their ratio gives the correction factor

$$c_{\text{sel, inc, }^{252}\text{Cf}} = 1.0003 \pm 0.0007(\text{stat}) \pm 0.0031(\text{syst}) \tag{5.36}$$

The systematic uncertainty is dominated by the conversion $\rho \to z$. Although the MC $\overline{\nu}_e$ has shown that f_ρ can be obtained with great accuracy from integrating $f_1(z)$ in a radial-like way; unlike the antineutrino case where $f_1(z)$ can be continuously measured, the ^{252}Cf $f_1(z)$ is only known in discrete positions, and a linear interpolation is used where data is not present. In order to estimate the uncertainty introduced by this lack of information, the number of points in figure 5.11(a) was artificially reduced by removing randomly data points to emulate the ^{252}Cf conditions. 5000 efficiency curves were produced in this way, and the discrepancy to the *true* volume-wide efficiency obtained by direct integration of the MC $\overline{\nu}_e$ was computed for each one. The average discrepancy in the efficiency was found to be 0.22%, which corresponds to a 0.31% in the correction factor.

5.4.2.3 Combined correction factor

Ultimately, two neutron selection efficiency correction factors have been computed: one using the $\overline{\nu}_e$ neutrons (equation 5.22) and another using the ^{252}Cf neutrons (equation 5.36). Both are in very good agreement with each other. Since they were found using two independent analyses, utilizing two different sources, and the systematic uncertainties are from different origin, they can be combined as uncorrelated measurements, finding $c_{\text{sel, inc, IBD+}^{252}\text{Cf}} = 0.9998 \pm 0.0018(\text{total})$. Because the result is compatible with 1.0 (no correction), suggesting that the MC simulation reproduces accurately the neutron selection efficiency, it was decided to adopt as the correction factor to be used in the θ_{13} analysis

$$c_{\text{sel, inc}} = 1.0000 \pm 0.0019 \tag{5.37}$$

where the total uncertainty has been extended accordingly to ensure the coverage of the case in which the 0.9998 correction were applied.

5.4.3 Spill-in/out

The Target scintillator acts as a natural fiducial volume in which the neutrons from the IBD events are captured on Gd nuclei. However, some neutrons created in the Target can escape into the Gamma Catcher and be captured on a H nucleus, which makes the event not to pass the Gd selection cuts, leading to a loss of expected $\overline{\nu}_e$. This is known as *spill-out*. On the other hand, IBD neutrons created in the Gamma Catcher can enter into the Target and be captured on a Gd nucleus, increasing the number of detected $\overline{\nu}_e$. This is called *spill-in*. Even though these two effects counterbalance, they do not cancel out. Because the Gamma Catcher surrounds the Target, the Gamma Catcher has a bigger volume capable of producing spill-in events. In addition, because the Gamma Catcher scintillator is not loaded with Gd, the neutrons are able to travel a longer distance in it before they are captured. As a consequence, the number of spill-in events surpasses the spill-out one, and there is a net spill-in current.

Unfortunately, the spill currents cannot be measured isolatedly in the DATA. The resolution of the position reconstruction does not allow to classify events with enough accuracy. Hence, the strategy of comparing the DATA to the MC simulation to compute a possible correction factor as it was done for the neutron selection efficiency cannot be followed, and the comparison between different MC simulations is left as the unique option.

The number of detected events (i.e. passing all selection cuts) in the simulation can be written as

$$N_{\rm det} = \epsilon_{\rm det}(N^{\rm T} - N^{\rm SO}) + N^{\rm SI}_{\rm det} + N^{\rm SO}_{\rm det} \tag{5.38}$$

where $N^{\rm T}$ is the number of IBD interactions occurred in the Target, $N^{\rm SO}$ is the number of spill-out events, $N^{\rm SI}_{\rm det}$ is the number of spill-in events passing the Gd selection, and $N^{\rm SO}_{\rm det}$ is the number of the IBD interactions in the Target in which the neutron was captured outside on a ^{12}C nucleus and pass the Gd selection anyway. $\epsilon_{\rm det}$ is the detection efficiency of the neutrons produced and captured within the Target, and it is defined as the product of the Gd fraction and the selection efficiency in the `Geant4` simulation, which are corrected using the factors computed in the previous sections to match the DATA.

Equation 5.38 can be rewritten as

$$N_{\rm det} = \epsilon_{\rm det} \cdot N^{\rm T}(1 + \phi^{\rm spill}_{\rm det}) \tag{5.39}$$

which shows explicitly how the number of detected IBD events in the Target, $\epsilon_{\rm det} \cdot N^{\rm T}$, receives a correction from the detected fractional spill current:

$$\phi^{\rm spill}_{\rm det} \equiv \frac{1}{N^{\rm T}} \left(\frac{N^{\rm SI}_{\rm det} + N^{\rm SO}_{\rm det}}{\epsilon_{\rm det}} - N^{\rm SO} \right) \tag{5.40}$$

The $(1 + \phi^{\rm spill}_{\rm det})$ correction can be identified with the neutron mobility factor, $k_{\rm spill}$, in equation 5.6.

Because the spill current is related to the neutron transport, it is very sensitive to the low energy neutron physics. As it has been described in section 5.2, in this energy regime it is particularly relevant how the molecular bonds of H are simulated. The default MC simulation in Double Chooz uses `Geant4`, which is optimized for high-energy neutrons, together with a patch, `NeutronTH`,

for kinetic energies below 4 eV. This MC is compared against `Tripoli4` [193], a code specifically designed for the simulation of the low energy neutron transport (originally for reactor physics). Whereas `NeutronTH` uses an analytical correction to simulate the effect of the hydrogen bonds, `Tripoli4` uses exclusively experimental data. The hydrogen nuclei are considered to be bound in CH_2 groups, which is the dominant structure in dodecane, the main component of the scintillator. Including other bonds, such as aromatic rings, does not introduce a significant discrepancy. `Tripoli4` uses the same detector geometry as `Geant4`, but the generation of γ rays from radiative captures is not activated since it lacks a model for the scintillator light emission, propagation and collection in the PMTs (which makes unavailable all the variables derived from the event reconstruction). In order to simulate the detection process, the events are weighted according to probabilities of detection built from the `Geant4` simulation using the event location.

The detected fractional spill current (eq. 5.40) is found to be 2.08% using the `Geant4-NeutronTH` simulation, and 2.36% using the `Tripoli4` simulation. Since `NeutronTH` describes the neutron thermalization sufficiently well, no correction factor was considered necessary

$$c_{\text{spill}} = 1.0000 \pm 0.0027, \tag{5.41}$$

and the difference between the two simulations was taken as the normalization uncertainty induced by the spill effect.

5.4.4 Summary

In table 5.3 all the individual efficiency corrections and uncertainties that intervene in the computation of the total MC normalization correction factor for the Gd selection are summarized. Those which are not related to the neutron physics have been described in chapter 4 (see table 4.1). This section has been devoted to describe the computation of the corrections related to the neutron detection efficiency (Gd fraction, neutron selection efficiency and spill-in/out). This thorough treatment is justified since the neutron detection efficiency is the dominant source of uncertainty in the normalization of the expected $\overline{\nu}_e$, totaling 0.54%. Nevertheless, compared to the previous Double Chooz result [1], in which this uncertainty added up to 0.96%, the improvement brought by the new methods described here is evident.

Furthermore, the new result paves the way for the two-detector phase, when only the uncorrelated uncertainties between detectors will affect the precision of the $\sin^2(2\theta_{13})$ measurement. Since the two Target volumes are filled with the same scintillator, which was produced as a single batch, the portion of uncorrelated uncertainty in the Gd fraction will be simply the statistical uncertainty, which currently is 0.11% and it can be reduced even further by acquiring more ^{252}Cf calibration data at the center of the detectors. In addition, because the two detectors have been made identical, the uncertainty caused by the spill current will cancel almost completely, up to an uncorrelated residual caused by the different acceptance of the two detectors to the $\overline{\nu}_e$ flux. Hence, the major contribution to the uncertainty will be the neutron selection efficiency, which currently amounts to 0.19%. This uncertainty is composed of a statistical component, which is inherently uncorrelated, which will be reduced as more

Source	Correction factor	Uncertainty
Number of protons	1.0000	0.0030
Muon veto	0.9551	< 0.0001
Light noise	0.9999	< 0.0001
Trigger	1.0000	< 0.0001
Isolation	0.9894	< 0.0001
Background vetoes	0.9934	0.0011
Gd fraction	0.9750	0.0042
Neutron selection	1.0000	0.0019
Spill	1.0000	0.0027
Total	0.9149	0.0058

Table 5.3: Compilation of the inputs and their uncertainties used to calculate the MC normalization correction factor due to the detection efficiency for the Gd selection. The number of protons in the Target scintillator (from tab. 2.5) is also listed since it affects the normalization in the same way as the detection efficiency (see eq. 5.1).

antineutrino data and ^{252}Cf calibration data are taken with both detectors; and a systematic component which is specific to the source used. In the case of the ^{252}Cf, it is the extrapolation from the vertical to the radial dimension of the detector; and it can be brought down by conducting a dedicated calibration campaign in which the regions where the efficiency varies most (the detector edges) are surveyed with high statistics. Eventually, the articulated arm, the device capable of deploying calibration sources along the radial coordinate which is already built but has not been used yet, would make the extrapolation from the vertical to the radial dimension unnecessary. In the case of the $\overline{\nu}_e$ source, the systematic uncertainty is due to the stopping muon contamination, which can be decreased even more refining the analysis techniques exploiting the differences between these events and the IBD ones.

5.5 Neutron detection in the hydrogen selection

In this section it is described the estimation of the neutron detection efficiency and its correction factor for the hydrogen selection. Although many of the innovations developed for the gadolinium analysis can be exported, the fact that the neutron capture can occur in two volumes, the Target and the Gamma Catcher, requires to reformulate some of the techniques.

5.5.1 Hydrogen fraction

The hydrogen fraction, f_{H}, corresponds to the proportion of radiative neutron captures that occur on H nuclei. As it happens for the Gd fraction, its value is determined by intrinsic properties of the liquid scintillator. However, in the H selection, the two scintillator volumes, Target and Gamma Catcher, have different scintillator compositions. Therefore, the computation of the hydrogen fraction must be done separately for each volume.

The definition of the hydrogen fraction correction factor c_{H}^{V} for a volume V is analogous to equation 5.12:

$$c_{\mathrm{H}}^{V} = \frac{f_{\mathrm{H}}^{V,\,\mathrm{DATA}}}{f_{\mathrm{H}}^{V,\,\mathrm{MC}}}, \tag{5.42}$$

where $f_{\mathrm{H}}^{V,\,\mathrm{DATA}}$ and $f_{\mathrm{H}}^{V,\,\mathrm{MC}}$ are the hydrogen fractions measured in the DATA and in the MC, respectively, within the volume V. As in the Gd case, the measurement of the H fraction was carried out using the two neutron sources described in section 5.3.

5.5.1.1 Target hydrogen fraction

In the Target, since the neutron captures on carbon represent less than 0.1% of all captures, and indeed, most of the carbon energy range is already included in the Gd fraction, the hydrogen fraction can be taken as the complementary of the gadolinium fraction:

$$f_{\mathrm{H}}^{\mathrm{T}} = 1 - f_{\mathrm{Gd}}^{\mathrm{T}}, \tag{5.43}$$

where the neutron source used to make the measurement will set the appropriate definition of the Gd fraction to use: equation 5.13 for the ^{252}Cf source and equation 5.15 for the $\overline{\nu}_e$ source.

Californium-252 measurement

Reutilizing the Gd-fraction measurements of section 5.4.1.1, and using equation 5.43 in 5.42, the hydrogen fraction correction factor in the Target volume is found to be

$$c_{\mathrm{H},\,^{252}\mathrm{Cf}}^{\mathrm{T}} = 1.1750 \pm 0.0078(\mathrm{stat}) \pm 0.0265(\mathrm{syst}), \tag{5.44}$$

where the systematic uncertainty is computed in the same way as in the gadolinium fraction correction factor (eq. 5.14). A consistent value from an earlier calibration campaign with reduced statistics, 1.1497 ± 0.0120(stat), confirms the stability of the aforementioned result.

Antineutrino measurement

Likewise, the complementarity with the Gd fraction in the Target is invoked, so the $\overline{\nu}_e$ measurements of section 5.4.1.2 allow to compute the hydrogen correction factor:

$$c_{\mathrm{H},\,\mathrm{IBD}}^{\mathrm{T}} = 1.167 \pm 0.033(\mathrm{stat}) \pm 0.036(\mathrm{syst}). \tag{5.45}$$

Similarly to the Gd result (eq. 5.16), the systematic uncertainty can be broken into a component accounting for the stopping muon contamination, 0.033, and a component related to the uncertainty introduced by the definition chosen for the Gd fraction (eq. 5.15), 0.016, which are added in quadrature.

The two results for the hydrogen fraction correction factor found with the ^{252}Cf source and the $\overline{\nu}_e$ source are consistent with each other. Both of them see their statistical uncertainty increased compared to the gadolinium result due to the lower number of neutron captures on H within the Target, which also limits

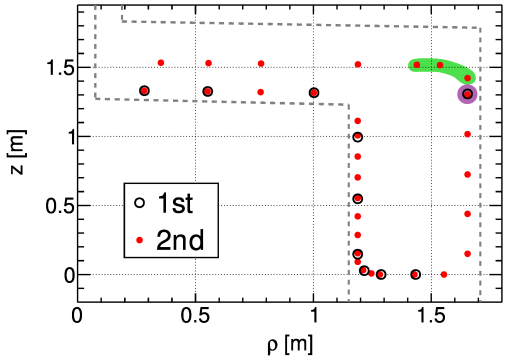


Figure 5.13: Guide Tube ^{252}Cf deployment positions in the first (empty black circles) and second (red circles) calibration campaigns. The runs used to compute the H-fraction correction are signaled with a colored background: purple for the first campaign and green for the second. From [189].

the knowledge of these events leading to poorer systematics. As it happened previously, the $\overline{\nu}_e$ cannot compete in precision with the ^{252}Cf result, which was adopted by the collaboration to correct the normalization of the MC IBD events in which the neutron is captured inside the Target:

$$c_{\mathrm{H}}^{\mathrm{T}} = 1.1750 \pm 0.0277. \tag{5.46}$$

5.5.1.2 Gamma Catcher hydrogen fraction

A dedicated measurement of the hydrogen fraction is necessary to estimate the correction factor in this volume. As in the Target, two measurements are made using the two neutron sources available.

Californium-252 measurement

The Gamma Catcher H-fraction measurement is performed with the ^{252}Cf source deployed along the Guide Tube. In order to minimize the risk of neutron captures on Gd, only the most distant deployment positions from the Target are used (see figure 5.13).

The Gamma Catcher hydrogen fraction measured with the ^{252}Cf source is defined as:

$$f_{\mathrm{H},^{252}\mathrm{Cf}}^{\mathrm{GC}} \equiv \frac{N(0.5 < E_{\mathrm{delayed}} < 3.5\,\mathrm{MeV} \cap 0 < \Delta T < 1000\,\mu\mathrm{s})}{N(0.5 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap 0 < \Delta T < 1000\,\mu\mathrm{s})}, \tag{5.47}$$

where N is the number of events passing the cuts between parentheses. When applied to the DATA and MC samples from the second calibration campaign, the H fraction is measured to be $(99.25 \pm 0.06)\%$ and $(99.05 \pm 0.02)\%$, respectively, where the uncertainties are statistical. Then, the correction factor is calculated as the ratio of the two numbers and it equals

$$c_{\mathrm{H},^{252}\mathrm{Cf}}^{\mathrm{GC}} = 1.0020 \pm 0.0007(\mathrm{stat}) \pm 0.0003(\mathrm{syst}), \tag{5.48}$$

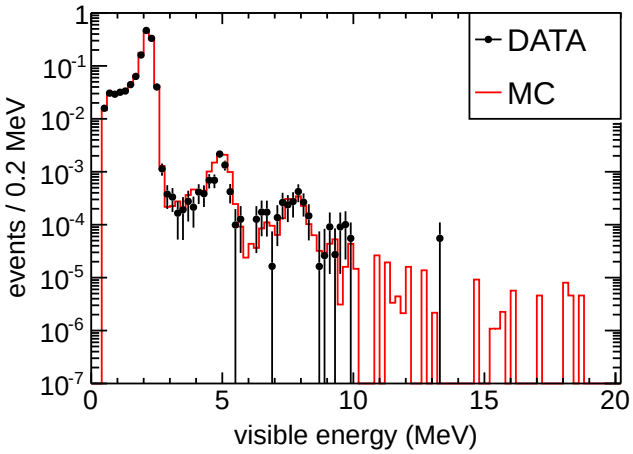


Figure 5.14: Delayed energy spectra of the second calibration campaign ^{252}Cf runs in the Guide Tube (see fig. 5.13) used to compute the Gamma Catcher hydrogen fraction. The MC (red histogram) has been normalized to the DATA (black points) integral. Error bars show the DATA statistical uncertainty. From [194].

where the systematic uncertainty corresponds to the uncertainty due to the particular definition chosen. It is estimated as the difference in the correction factor found when the delayed energy lower limit is moved from 0.5 MeV to 1.5 MeV, to account for the discrepancy between DATA and MC observed at low energies in the Target. However, in the Gamma Catcher the DATA and MC spectra exhibit good agreement in all the energy range (see figure 5.14), resulting in a much smaller uncertainty.

The stability of the result was examined by computing $c^{\mathrm{GC}}_{\mathrm{H},^{252}\mathrm{Cf}}$ using the data from the first calibration campaign, which happens to have more run time in that part of the detector (160 min vs. 65 min). The value found, 1.0023 ± 0.0005(stat), is in excellent agreement with the one in equation 5.48.

Antineutrino measurement

In order to compute the Gamma Catcher hydrogen fraction, in which all the possible delayed energies resulting from neutron captures in the scintillator must be considered, the ANN cut cannot be used since it is not designed to work beyond a delayed energy of 3.5 MeV, so an extra cut on the prompt energy $E_{\mathrm{prompt}} > 4$ MeV is used to suppress the accidental contamination. Furthermore, the F_V Gd cut, $E_{\mathrm{delayed}} > 0.068 \cdot \exp(FV/1.23)$, is added to the H cut to extend the stopping muon and light noise rejection up to high energies. Analogous to the fiducial volume defined to measure the Gd fraction in section 5.4.1.2, to maximize the probability of selecting candidates that have occurred within the Gamma Catcher; both the prompt and the delayed vertices are required to be reconstructed inside it, i.e. $|z| < 1837.24$ mm and $\rho < 1708$ mm, but out of the Target (which occupies the volume given by $|z| < 1274.54$ mm

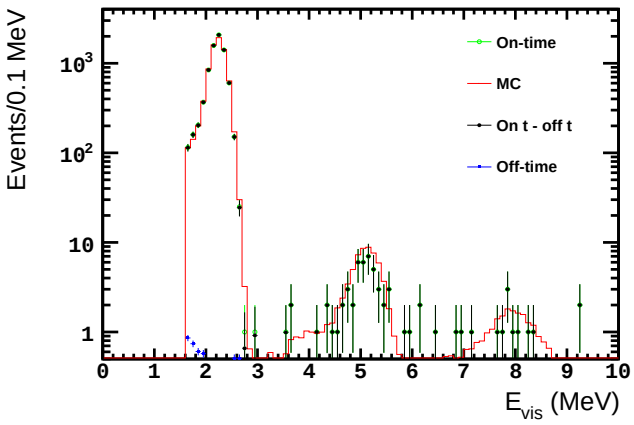


Figure 5.15: Delayed energy spectra of the IBD candidates reconstructed inside the Gamma Catcher used to compute the H fraction. MC $\overline{\nu}_e$ (red), IBD candidates (green), accidental background (blue), and background-subtracted IBD candidates (black) are shown. The MC has been normalized to the DATA integral. Error bars show the DATA statistical uncertainty.

and $\rho < 1150\,\mathrm{mm}$). Finally, the H fraction is defined as:

$$f^{\mathrm{GC}}_{\mathrm{H,\,IBD}} = \frac{N(1.6 < E_{\mathrm{delayed}} < 3\,\mathrm{MeV} \cap 0.25 < \Delta T < 600\,\mu\mathrm{s} \cap \Delta R < 0.8\,\mathrm{m})}{N(1.6 < E_{\mathrm{delayed}} < 10\,\mathrm{MeV} \cap 0.25 < \Delta T < 600\,\mu\mathrm{s} \cap \Delta R < 0.8\,\mathrm{m})}, \quad (5.49)$$

where N is the number of IBD candidates selected by the cuts shown in parentheses, after the subtraction of the remaining accidental background, which has been measured by the off-time window method. Figure 5.15 shows the delayed energy spectra of the events passing the cuts in the denominator.

The measured H fractions in DATA and MC are (99.13±0.12)% and (98.74± 0.01)%, respectively, where the uncertainties are statistical. The correction factor is found to be:

$$c^{\mathrm{GC}}_{\mathrm{H,\,IBD}} = 1.0040 \pm 0.0012(\mathrm{stat}) \pm 0.0005(\mathrm{syst}), \quad (5.50)$$

where the systematic uncertainty accounts for the possible variation of the correction factor due to a residual contamination of stopping muons, which is estimated as the difference in the correction factor found when only IBD candidates reconstructed in the bottom half of the Gamma Catcher (where this background is suppressed) are used. Compared to the result in the Target (eq. 5.45), the systematic uncertainty is remarkably low. This is due to the fact that the non-vetoed stopping muons penetrate into the detector through the chimney and decay inside the Target. In order to decay in the Gamma Catcher, the muon should cross first the Outer Veto and Inner Veto unnoticed, which is unlikely.

A further search for possible sources of systematic uncertainty was conducted:

- The minimum delayed energy was decreased to 1.3 MeV (it cannot be decreased much further because of the $^{212}\mathrm{Bi}-^{212}\mathrm{Po}$ background); and the

ΔT cut was extended to 1000 μs and the ΔR cut removed, as in the ^{252}Cf definition (eq. 5.47), to test the effect of the chosen definition.

- In figure 5.15 it can be seen a small peak at ~ 8 MeV which might come from captures on Gd which have been mistakenly reconstructed inside the Gamma Catcher [4]. To test the impact of this possible contamination, the lower ΔT cut was changed to 90 μs (about three times the neutron capture time on Gd).

- The use of two delayed F_V cuts, one for the H range and another for the Gd one, might bias the H fraction in the case they affect differently to the n-H and n-Gd captures. In order to estimate the impact to the result, these two cuts were removed (but since these cuts are required to reject stopping muons, a minimum time coincidence of 30 μs was used to compensate for the removal).

- Lastly, since stopping muons are the main background, the correction was recalculated using only runs with the OV-active to measure the repercussion of including the run time in which the OV was off.

Among all the tests listed above, none produced a significant deviation from the result in eq. 5.50; so an additional systematic uncertainty is considered unnecessary.

Both corrections, the one computed using the $\overline{\nu}_e$ (eq. 5.50) and the one found (eq. 5.48) using the ^{252}Cf source, agree within uncertainty, although the latter has slightly better statistics and systematic uncertainty, which motivated its choice as the hydrogen fraction correction factor to be applied to the normalization of the MC IBD events in which the neutron is captured inside the Gamma Catcher:

$$c_{\mathrm{H}}^{\mathrm{GC}} = 1.0020 \pm 0.0008. \tag{5.51}$$

5.5.2 Neutron selection efficiency

In the H selection, the neutron captures can occur in two volumes, the Target and the Gamma Catcher. As it was stated in the measurement of the Gd fraction with $\overline{\nu}_e$ (section 5.4.1.2), it is especially difficult to separate the events which have occurred in the Target from those of the Gamma Catcher because the reconstructed vertices tend to be shifted towards the center with respect to the true vertices, so a fraction of the neutron capture events reconstructed inside the Target volume actually occurred in the Gamma Catcher. This shift affects more to the low energy events. This observation discourages from using the prompt vertex as the classifier variable, which spans a wider energy range than the delayed, because the accuracy of the volume classification would end up being energy dependent. In addition, the delayed vertex suffers from an additional bias due to the ~ 20 cm mean free path of the single 2.2 MeV γ ray, which blurs the volume edges even more. Moreover, there are the spill-in/out currents which add additional uncertainty on the location of the interaction.

[4] However, this is not the only possibility. It could be due to neutron captures on iron and other nuclei that compose the Guide Tube which runs along the Gamma Catcher. In fact, this peak is also seen in the ^{252}Cf data (fig. 5.14), where the neutrons are emitted from the furthest positions from the Target.

For the abovesaid reasons, a separate characterization of the selection efficiency of neutron captures on H in the Target and Gamma Catcher becomes extremely challenging. This is especially problematic for the ^{252}Cf source, which requires a precise knowledge of the efficiency dependence along the vertical coordinate of the detector to infer the radial dependence. In addition, the vertical-radial symmetry which is the basis of this method is broken for H captures because of the existence of structures (the chimney, Target stiffener and feet, cf. figure 2.5) in the Gamma Catcher and close to the Target top and bottom edges.

On the other hand, the antineutrino source, because is homogeneously distributed in both Target and Gamma Catcher volumes, can be used to perform a direct measurement making no separation upon where the neutron was captured, treating the two volumes effectively as a single volume so as not to depend on the position reconstruction. An additional advantage of this approach is, since no distinction is made between the captures in the Target and the Gamma Catcher, the spill-in/out currents between Target and Gamma Catcher are integrated out and they do not contribute to the uncertainty. Therefore, this source was adopted as the reference to measure the neutron selection efficiency in the hydrogen analysis.

The inclusive neutron selection efficiency measured with the $\overline{\nu}_e$ source is defined as:

$$\epsilon_{\text{sel, inc, IBD}} \equiv \frac{N(ANN > -0.23 \cap 0.5 < \Delta T < 800\,\mu\text{s} \cap 1.3 < E_{\text{delayed}} < 3\,\text{MeV} \cap \Delta R < 1.2\,\text{m})}{N(ANN > -0.4 \cap 0.25 < \Delta T < 1000\,\mu\text{s} \cap 1.3 < E_{\text{delayed}} < 3.1\,\text{MeV} \cap \Delta R < 1.5\,\text{m})}, \tag{5.52}$$

where N is the number of IBD candidates selected by the cuts shown in parentheses, after the accidental background contamination measured by the off-time window method is subtracted. Some of the accepted combinations of the $\Delta T, E_{\text{delayed}}, \Delta R$ variables by the ANN cuts in the numerator and the denominator are shown as two-dimensional regions in figures 5.16, 5.17 and 5.18, in which the non-displayed variable is fixed at a given value. They correspond to two-dimensional sections of figure 4.21 and picture how the ANN works: the more IBD-like the $\Delta T, E_{\text{delayed}}, \Delta R$ values, the bigger the accepted region. However, from figures 5.17 and 5.18 it is observed how the ANN cuts harder on ΔR than the numerator and denominator ΔR cuts themselves, making them redundant. Also, figures 5.16 and 5.17 show that the ANN opens an accepted region for $E_{\text{delayed}} \gtrsim 3.1\,\text{MeV}$, where no H capture signal is expected as already pointed out in section 4.2.2. This motivates to restrict the denominator to $E_{\text{delayed}} < 3.1$ in order to exclude it. As a result, the ANN ends up cutting harder on E_{delayed} than the E_{delayed} cuts too.

The measured efficiencies in DATA and MC are (95.12±0.12)% and (95.164± 0.012)%, respectively, where the uncertainties are statistical. The correction factor is computed as the ratio of the DATA to the MC efficiency:

$$c_{\text{sel, inc, IBD}} = 0.9995 \pm 0.0013(\text{stat}) \pm 0.0011(\text{syst}), \tag{5.53}$$

where the systematic uncertainty includes a contribution from the accidental background subtraction (0.0001), and another one accounting for the remnant stopping muon background contamination (0.0011), which is estimated in the way described previously.

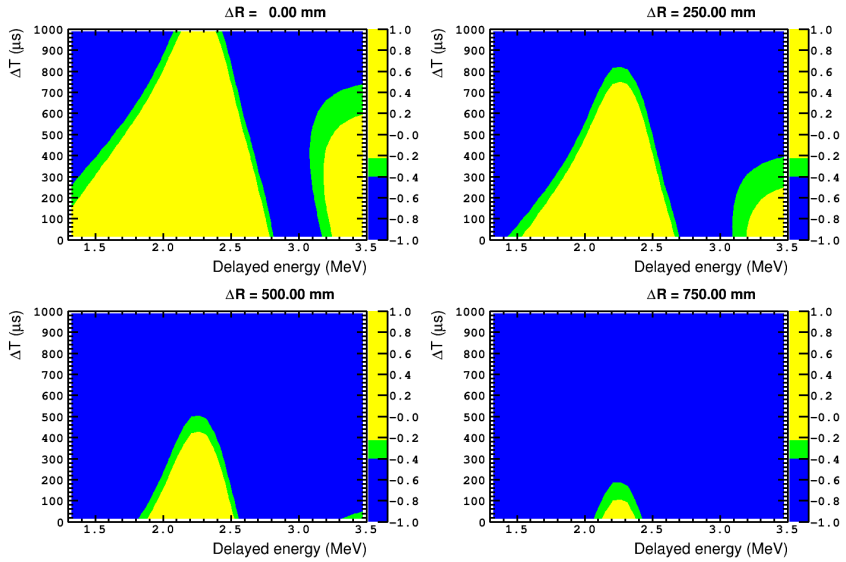


Figure 5.16: Selected regions in the $\Delta T - E_{\text{delayed}}$ plane by the ANN cuts in the numerator -0.23 (yellow), and the denominator -0.4 (green) of equation 5.52. ΔR is fixed at the value shown on top of each plot.

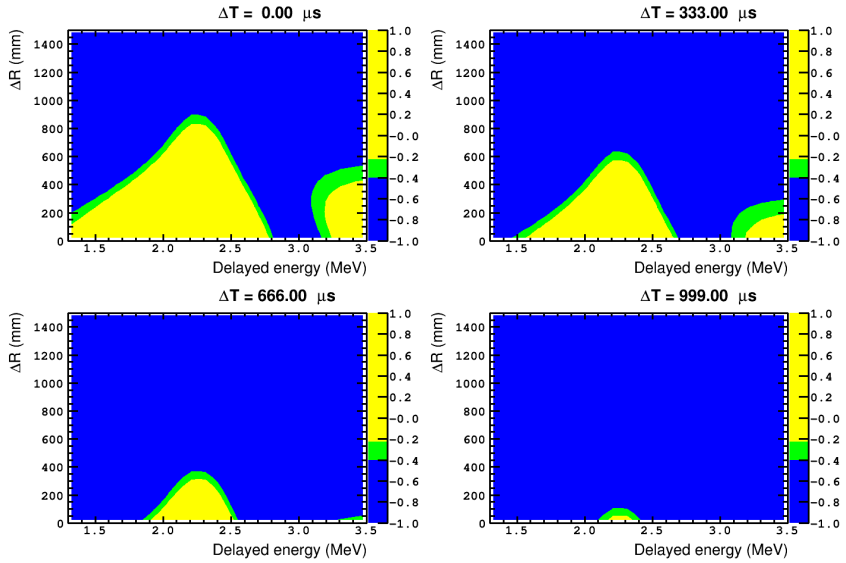


Figure 5.17: Selected regions in the $\Delta R - E_{\text{delayed}}$ plane by the ANN cuts in the numerator -0.23 (yellow), and the denominator -0.4 (green) of equation 5.52. ΔT is fixed at the value shown on top of each plot.

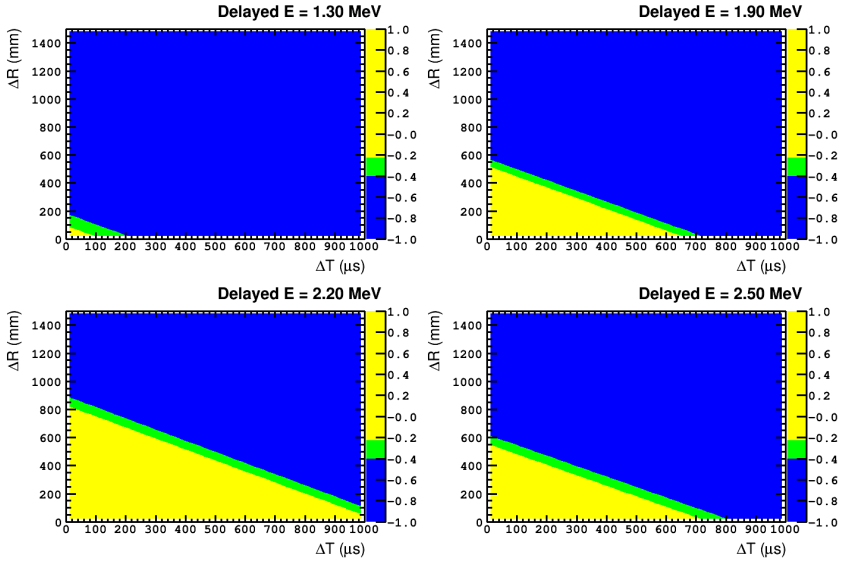


Figure 5.18: Selected regions in the $\Delta R - \Delta T$ plane by the ANN cuts in the numerator -0.23 (yellow), and the denominator -0.4 (green) of equation 5.52. E_{delayed} is fixed at the value shown on top of each plot.

The stability of the correction factor against perturbations in the efficiency definition was checked, either by loosening the cuts (only the ANN cut could be relaxed to $ANN > -0.48$ since the other cuts are already at their maximum values dictated by the range of application of the ANN) or by tightening them (the only cuts not superseded by the ANN cut, the ΔT upper limit and the ANN cut itself were both separately and simultaneously changed to 900 μs and -0.32, respectively). In both cases, no significant deviation was observed from the value in eq. 5.53.

The sufficiency of the background uncertainty was evaluated by computing the correction factor with a stopping muon-reduced sample (by using only runs with the OV active), which resulted in a consistent value with the one found. It was tested further by computing the correction with a background-enriched sample, obtained by not applying the Li+He and IV vetoes to the candidates. The innocuousness of these backgrounds for the efficiency analysis was demonstrated since the same correction factor as in eq. 5.53 could be retrieved [5].

An alternative definition to the inclusive efficiency, in which all the neutron selection cuts are loosened simultaneously, is given by the *exclusive* efficiency, which is computed as the product of the single-cut efficiencies found when only one of the cuts is relaxed at each time:

$$\epsilon_{\text{sel, exc}} = \epsilon_{ANN} \cdot \epsilon_{\Delta T} \cdot \epsilon_{E_{\text{delayed}}} \cdot \epsilon_{\Delta R}, \tag{5.54}$$

where ϵ_{ANN}, the efficiency of the ANN cut, is the only novelty introduced with respect to the definition used in the Gd measurement (equation 5.20). Since the ANN is a function of the other three variables, $ANN = f(\Delta T, E_{\text{delayed}}, \Delta R)$,

[5] This also confirmed the robustness of the analysis despite not having the MPS veto, which was adopted after the conclusion of this study.

the factorization in equation 5.54 fails to isolate the effect of each cut, for which reason it is only considered as a cross-check to the inclusive value.

For the H analysis, the single-cut efficiencies are defined as:

$$\epsilon_{ANN,\,\mathrm{IBD}} \equiv \frac{N(ANN > -0.23 \cap 0.5 < \Delta T < 800\,\mu\mathrm{s} \cap 1.3 < E_{\mathrm{delayed}} < 3\,\mathrm{MeV} \cap \Delta R < 1.2\,\mathrm{m})}{N(ANN > -0.4 \cap 0.5 < \Delta T < 800\,\mu\mathrm{s} \cap 1.3 < E_{\mathrm{delayed}} < 3\,\mathrm{MeV} \cap \Delta R < 1.2\,\mathrm{m})} \tag{5.55}$$

$$\epsilon_{\Delta T,\,\mathrm{IBD}} \equiv \frac{N(ANN > -0.23 \cap 0.5 < \Delta T < 800\,\mu\mathrm{s} \cap 1.3 < E_{\mathrm{delayed}} < 3\,\mathrm{MeV} \cap \Delta R < 1.2\,\mathrm{m})}{N(ANN > -0.23 \cap 0.25 < \Delta T < 1000\,\mu\mathrm{s} \cap 1.3 < E_{\mathrm{delayed}} < 3\,\mathrm{MeV} \cap \Delta R < 1.2\,\mathrm{m})} \tag{5.56}$$

$$\epsilon_{E_{\mathrm{delayed}},\,\mathrm{IBD}} \equiv 1 \tag{5.57}$$

$$\epsilon_{\Delta R,\,\mathrm{IBD}} \equiv 1 \tag{5.58}$$

where $\epsilon_{E_{\mathrm{delayed}},\,\mathrm{IBD}} \equiv 1$ and $\epsilon_{\Delta R,\,\mathrm{IBD}} \equiv 1$ because these cuts are superseded by the ANN cut as explained before. The results of each single-cut efficiency for the DATA and MC and the individual correction factor derived from their ratio are summarized in table 5.4.

The final exclusive selection efficiencies, from equation 5.54, are $(95.22^{+0.12}_{-0.13})\%$ for DATA and $(95.241 \pm 0.012)\%$ for MC (uncertainties are statistical). The correction factor is found to be

$$c_{\mathrm{sel,\,exc,\,IBD}} = 0.9997 \pm 0.0013(\mathrm{stat}) \pm 0.0009(\mathrm{syst}), \tag{5.59}$$

where the systematic uncertainty has the same contributions as the inclusive result (eq. 5.53), with which shows a very good agreement.

As long as the correction factor found is consistent with 1.0, that is, the MC simulation reproduces faithfully the neutron selection efficiency as it is measured in the DATA, the final neutron selection efficiency correction factor which will be used in the θ_{13} analysis can be taken as:

$$c_{\mathrm{sel,\,inc}} = 1.0000 \pm 0.0022, \tag{5.60}$$

where the uncertainty has been modified to ensure at least the same coverage as in the original result in equation 5.53.

5.5.3 Spill

When the signal is given by the neutron captures on hydrogen, the spill phenomenon becomes more complex than in the gadolinium capture case. Now the neutron can travel from the Target to the Gamma Catcher (and vice versa) and still produce a detectable signal since there are H nuclei in both scintillators, so the spill-in and spill-out cases of the Gd lose their meaning. Moreover, the neutron can also be captured on the H nuclei which make up the acrylic vessels. Finally, even though the Buffer is filled with non-scintillating mineral oil, an event produced outside the Gamma Catcher but close enough to it can be detected if part of its energy is deposited in the scintillator.

In a general way, the number of detected $\overline{\nu}_e$ events can be written as

$$N_{\mathrm{det}} = N^{\mathrm{T}}_{\mathrm{det}} + \Phi^{\mathrm{int}}_{\mathrm{det}} + N^{\mathrm{GC}}_{\mathrm{det}} + \Phi^{\mathrm{ext}}_{\mathrm{det}}, \tag{5.61}$$

	$\epsilon_{\mathrm{IBD}}^{\mathrm{DATA}}$ (%)	$\epsilon_{\mathrm{IBD}}^{\mathrm{MC}}$ (%)	c_{IBD}
ANN	$95.57^{+0.12}_{-0.12}$	$95.521^{+0.011}_{-0.011}$	1.0005 ± 0.0015 $(0.0013(\mathrm{stat}) \pm 0.0009(\mathrm{syst}))$
ΔT	$99.64^{+0.03}_{-0.04}$	$99.706^{+0.003}_{-0.003}$	0.9993 ± 0.0005 $(0.0004(\mathrm{stat}) \pm 0.0003(\mathrm{syst}))$
E_{delayed}	100.00	100.000	1.0000
ΔR	100.00	100.000	1.0000
Exclusive	$95.22^{+0.12}_{-0.13}$	$95.241^{+0.012}_{-0.012}$	0.9997 ± 0.0016 $(0.0013(\mathrm{stat}) \pm 0.0009(\mathrm{syst}))$
Inclusive	$95.12^{+0.12}_{-0.12}$	$95.164^{+0.012}_{-0.012}$	0.9995 ± 0.0017 $(0.0013(\mathrm{stat}) \pm 0.0011(\mathrm{syst}))$

Table 5.4: Neutron summary of selection efficiencies and correction factors for the hydrogen analysis computed with the $\overline{\nu}_e$ source. Efficiency uncertainties are statistical only.

where $N_{\rm det}^{\rm T}$ corresponds to events which have both the positron and the neutron contained within the Target. $\Phi_{\rm det}^{\rm int}$, the *internal* spill events, corresponds to spill events occurring in the Target-Gamma Catcher interface, or in the Target vessel. $N_{\rm det}^{\rm GC}$ represents the number of events with the two signals inside the Gamma Catcher. $\Phi_{\rm det}^{\rm ext}$, the *external* spill events, corresponds to any spill event involving the Gamma Catcher vessel or the Buffer.

The strategy to estimate the normalization uncertainty due to the neutron migration relies on the comparison between the two simulations, `Geant4` and `Tripoli4`, as it did for the Gd case (section 5.4.3). As it was stated before, `Tripoli4` does not feature a model of the detector response, so the prompt and delayed visible energies or the reconstructed vertices are not available. This represents a significant shortcoming since the reconstructed variables $E_{\rm delayed}$, ΔT, ΔR are needed to produce the ANN value which is used in the selection. In order to emulate the reconstruction with `Tripoli4`, detector response maps for the $E_{\rm delayed}$ and ΔR variables were built using the reconstructed `Geant4` simulation (the timing information can be taken directly from the truth information since it is very similar to the reconstructed one though). Those maps are binned unevenly to provide more granularity in the position dependence close to the volume boundaries, where those variables vary the most. Thus, for a `Tripoli4` neutron captured in a certain location, a $E_{\rm delayed}$ and ΔR [6] values are randomly generated for it according to the `Geant4` distributions at that spatial bin. These procedure is iterated 10000 times to minimize the statistical fluctuations.

In order to calculate the number of the detected events in equation 5.61, the $\overline{\nu}_e$ from both MC simulations are selected using the H selection cuts (sec. 4.2.2). Moreover, to ensure the likeness between the DATA and the MC, the `Geant4` normalizations are corrected using the hydrogen fraction and neutron selection efficiency corrections factors found in the previous sections. The discrepancies in the number of spill events between the `Geant4` and `Tripoli4` simulations found are 0.04% for the internal spill, and 0.13% for the external spill; both expressed relative to the total normalization of detected events in `Geant4`. While it is true that the candidate selection and the neutron selection efficiency are computed without making any distinction about where the neutron is captured, the fact that such distinction is made to apply the hydrogen correction factor makes necessary to account for the uncertainty introduced by the internal spill events, in addition to the external spill. Thus, the discrepancy between simulations in the global spill class (the sum of the internal and external spills), 0.18%, is considered as the normalization uncertainty caused by the modelization of the neutron migration.

There are additional statistical uncertainties arising from the size of the MC samples used to estimate the discrepancy (0.03%) and from the emulation of the detector response in `Tripoli4`, which is estimated from the standard deviations of the resulting distributions of the number of spill events after the 10000 iterations. The latter is found to amount to 0.02% of the `Geant4` selected $\overline{\nu}_e$.

In order to address a possible systematic bias stemming from the emulation of the detector response, the discrepancy between `Geant4` and `Tripoli4` spill numbers is recomputed correcting the `Tripoli4` simulation with the detection

[6] Specifically, for ΔR, what it is generated is a shift, δR, from the `Geant4` reconstructed ΔR - true ΔR distribution, which is then added to the true ΔR from `Tripoli4`.

Source	Relative uncertainty (%)
`Tripoli4` − `Geant4` global spill	0.18
Statistics of the simulations	0.03
Emulation of detector response in `Tripoli4` (stat.)	0.02
Emulation of detector response in `Tripoli4` (syst.)	0.22
Positron energy scale	0.07
Total	0.29

Table 5.5: Contributions to the $\overline{\nu}_e$ normalization uncertainty in the H-based selection due to the neutron migration between volumes (spill). The uncertainties are expressed relative to the total normalization of selected events in the `Geant4` simulation.

efficiencies of the spill classes measured in `Geant4`, so the reconstruction of the missing variables in `Tripoli4` is bypassed. The new discrepancies found are 0.05% in the internal spill and 0.17% in the external spill, totaling 0.22% for the global spill. Since it is not possible to disentangle how much of the discrepancy is due to the differences in the neutron modelization between the two codes and how much is caused by a bias in the emulation of the reconstruction; it is considered as a systematic uncertainty which has to be added.

Finally, the number of selected spill events in the simulation depends on the visible energy of the positron through the prompt energy cut $E_{\text{prompt}} > 1\,\text{MeV}$ (equation 4.14a). Therefore, the uncertainty on the positron energy scale affects the results. Assuming a conservative uncertainty, variations of ±5% of the energy scale at 1 MeV are allowed, which results in an additional 0.07% uncertainty on the spill normalization.

If it is expressed in the correction factor notation, the final 0.29% normalization uncertainty originated by the spill (see table 5.5) is equivalent to a unity correction factor with the corresponding uncertainty:

$$c_{\text{spill}} = 1.0000 \pm 0.0029. \tag{5.62}$$

5.5.4 Summary

The individual efficiency corrections and uncertainties used to compute the total MC normalization correction factor (see equations 5.7, and 5.8, 5.9, 5.10) for the H-based $\overline{\nu}_e$ selection are listed in table 5.6. The factors not related to the neutron detection have been explained in chapter 4 (cuts and vetoes, values from table 4.3) or in the chapter 2 (number of protons, values from table 2.5). The contribution from the neutron detection efficiency, which has been computed in the previous sections, corresponds to the H fraction, neutron selection and spill entries.

Compared to the gadolinium case, the normalization uncertainty is dominated by the proton number. Among the contributions to this uncertainty, the Gamma Catcher one accounts almost for all of it. Originally, this volume was not envisioned to be used in the $\overline{\nu}_e$ selection, which made unnecessary to have a profound knowledge of its characteristics.

Concerning the neutron detection efficiency, the uncertainty is found to be almost equally distributed between the three components. In the case of the

Source	Correction factor	Uncertainty
Number of protons	1.0014[1]	0.0091
Muon veto	0.9399	< 0.0001
Light noise	0.9994	0.0001
Trigger	1.0000	< 0.0001
Isolation	0.9788	< 0.0001
Background vetoes	0.9939	0.0020
H fraction	1.0141[2]	0.0021
Neutron selection	1.0000	0.0022
Spill	1.0000	0.0029
Total	0.9280	0.0095

[1] Effective correction for the *rate* analysis. For the *Rate+Shape* analysis, the normalization of the $\overline{\nu}_e$ which interact in the Target vessel receives the correction (1.39), while for the other volumes is uncorrected (1.00) as explained in section 2.4.2.1.

[2] Effective correction for the *rate* analysis. For the *Rate+Shape* analysis, the normalization of the $\overline{\nu}_e$ with the neutron captured in the Target is corrected with 1.1750 ± 0.0277 (eq. 5.46), and for those with the neutron captured in the Gamma Catcher $\overline{\nu}_e$ is corrected with 1.0020 ± 0.0008 (eq. 5.51).

Table 5.6: Compilation of the inputs and their uncertainties used to calculate the MC normalization correction factor due to the detection efficiency for the H selection. The number of protons is also listed since it affects the normalization in the same way as the detection efficiency (see eq. 5.1).

H-fraction correction, the uncertainty comes mainly from the Target. In the Target scintillator, the H captures are the alternative to Gd captures, so the H fraction is affected by the uncertainty on the Gd fraction (which is the main uncertainty in the Gd analysis). Nevertheless, the fact that 89.5% of the captures on H nuclei occur in the Gamma Catcher, where the H fraction is precisely measured, alleviates the lack of knowledge in the Target. The neutron selection efficiency correction shows an uncertainty very similar to the Gd analysis, which is expected since the same methods using the $\overline{\nu}_e$ source have been used in both analyses. The spill has also a similar uncertainty since both H and Gd estimations rely on the comparison between the `Geant4` and `Tripoli4` simulations.

Regarding the outlook for the two-detector phase, it must be noted that since the Gamma Catcher was not planned to be used as a $\overline{\nu}_e$ detection volume, the liquid scintillator used for the Near Detector Gamma Catcher does not come from the same batch as the Far Detector one, unlike the Target scintillator. Therefore, the uncertainty due to the proton number on the Gamma Catcher is not expected to cancel in the two-detector phase. However, an improvement on this uncertainty is expected (reaching $\gtrsim 0.4\%$) because of the measurements made during the Near Detector filling and provided that only relative differences between detectors are relevant in this phase.

On the contrary, the H fraction uncertainty benefits from the two-detector phase because it is driven by the Target scintillator, which is identical for both detectors; so the only uncorrelated component is the statistical one, which can be made negligible with dedicated calibration.

Since the measurement of the neutron selection efficiency has an uncertainty very similar to the Gd one, the same considerations apply. This uncertainty is uncorrelated between detectors, so it is still relevant in the two-detector phase. The statistical component, which makes up half of the uncertainty, will decrease as more antineutrino data is acquired. It must be noted that the proximity of the Near Detector to the reactor cores allows it to gain antineutrino data faster than the Far Detector, so the dominant statistical uncertainty after a few months is the Far Detector one. The systematic component arises from the stopping muon contamination, so it can be reduced if new vetoes are devised.

Finally, in the two-detector phase the pertinent spill uncertainty is due to differences between the real detectors rather than between the simulations, and it is expected to be almost negligible.

Chapter 6

Oscillation analysis and measurement of θ_{13}

After the selection of IBD candidates in the DATA and MC has been performed according to the cuts described in chapter 4, with an estimation of the backgrounds contaminating the DATA using the techniques explained in the same chapter, and the normalization of the MC $\overline{\nu}_e$ simulation has been corrected with the factors described extensively in chapter 5; the number of observed IBD candidates can be compared to the prediction (see table 6.1). In both selections using neutron captures either on Gd or on H, a deficit with respect to the prediction is found. This deficit of IBD interactions is interpreted as the result of electron antineutrino conversion to other flavors driven by the θ_{13} mixing angle.

In order to exploit the signal information and distinguish it from the background with the goal of providing a precise measurement of θ_{13}, the $\overline{\nu}_e$ deficit is studied as a function of the reactor thermal power (*Reactor Rate Modulation* analysis, section 6.1) or as a function of the prompt trigger visible energy (*Rate+Shape* analysis, section 6.2). Finally, the chapter concludes with the prospective improvement to the measure of θ_{13} brought by the inclusion of the Near Detector (section 6.3).

6.1 Reactor Rate Modulation analysis

The Reactor Rate Modulation (RRM) analysis [74] is based on the fact that while the $\overline{\nu}_e$ interaction rate scales with the reactor thermal power, the background rate remains constant. This allows to disentangle the contribution of the background to the observed IBD candidate rate from that of the $\overline{\nu}_e$, which is the only affected by the value of θ_{13}.

The observed rate can be expressed as

$$R_{\text{obs}} = B + \left(1 - \sin^2(2\theta_{13})\eta_{\text{osc}}\right) R_\nu, \tag{6.1}$$

where B is the background rate and R_ν is the expected rate of $\overline{\nu}_e$ if there were not oscillation (given by equation 2.7). η_{osc} is the oscillatory factor of equation 1.54 averaged over the energies and baselines of the $\overline{\nu}_e$ emitted by the reactor

Gd selection	Reactor on	Reactors off
Live-time (days)	460.67	7.24
Reactor $\overline{\nu}_e$ (MC)	17530 ± 320	1.57 ± 0.47
Fast-n and stopping-μ	278 ± 23	3.83 ± 0.64
^{9}Li and ^{8}He	447^{+189}_{-74}	$7.02^{+2.97}_{-1.16}$
Accidental coincidences	32 ± 1	0.51 ± 0.02
Total prediction (events)	18287^{+372}_{-329}	$12.93^{+3.07}_{-1.41}$
IBD candidates observed	17351	7

H selection	Reactor on	Reactors off
Live-time (days)	455.57	7.15
Reactor $\overline{\nu}_e$ (MC)	30090 ± 610	2.73 ± 0.82
Fast-n and stopping-μ	706 ± 68	10.37 ± 1.43
^{9}Li and ^{8}He	433^{+260}_{-150}	$6.79^{+4.08}_{-2.36}$
Accidental coincidences	1974 ± 5	30.88 ± 0.40
Total prediction (events)	33203^{+667}_{-632}	$50.77^{+4.41}_{-2.91}$
IBD candidates observed	31835	63

Table 6.1: Comparison of the predicted and observed numbers of IBD candidates in the Far Detector for the Gd-based (top) and the H-based (bottom) $\overline{\nu}_e$ selections, separated by whether the two reactors were off or not, for the data taken between April 13$^{\text{th}}$ 2011 and January 15$^{\text{th}}$ 2013. The total prediction is broken down into the contributions by the reactor $\overline{\nu}_e$ of the MC simulation and the results of the background estimations. Neutrino oscillation effects are not assumed in the prediction.

cores,

$$\eta_{\rm osc} = \left\langle \sin^2\left(\frac{1.27\Delta m^2_{31}[{\rm eV}^2]L[{\rm m}]}{E[{\rm MeV}]}\right)\right\rangle, \tag{6.2}$$

with $\Delta m^2_{31} = (2.44^{+0.09}_{-0.10}) \cdot 10^{-3}\,{\rm eV}^2$ from the measurement by the MINOS experiment [59], assuming normal mass hierarchy.

Graphically, if the observed rate is represented as a function of the expected rate for different reactor conditions, a straight line will appear; the background rate B will be given by the intercept at $R_\nu = 0$ and the slope of the line will be proportional to θ_{13}. As a result, both θ_{13} and B can be measured simultaneously without resorting to any background model.

In order to perform the RRM analysis, the data is classified according to the total reactor thermal power. In Double Chooz, with two reactor cores, there are three evident reactor scenarios: both reactors are running, one reactor is on and the other is off, or both reactors are off (see figure 6.1). Since a reactor is not always operating at full power, the first two conditions can be split into three subdivisions each with approximately equal statistics. This results in 7 data points.

The best-fit value of θ_{13} is found minimizing a χ^2 function,

$$\chi^2 = \chi^2_{\rm on} + \chi^2_{\rm off} + \chi^2_{\rm pull}, \tag{6.3}$$

where $\chi^2_{\rm on}$ is the term corresponding to the data taken with at least one reactor on, $\chi^2_{\rm off}$ is the term for the data with the two reactors off (which is treated separately because of the low statistics) and $\chi^2_{\rm pull}$ is the term accounting for systematic uncertainties using pulls (also known as nuisance parameters).

Assuming Gaussian-distributed uncertainties, $\chi^2_{\rm on}$ is written as

$$\chi^2_{\rm on} = \sum_i \frac{\left[R_{{\rm obs},i} - \left(1 - \sin^2(2\theta_{13})\eta_{\rm osc}\right) R_{\nu,i}\left(1 + \alpha_{\rm d} + k_i\alpha_{\rm r} + w_i\alpha_{\rm res}\right) - B\right]^2}{\sigma^2_{{\rm stat},i}}, \tag{6.4}$$

where the sum is over the 6 data points in which one reactor was on and the other off, or both reactors were on. $R_{{\rm obs},i}$ and $R_{\nu,i}$ are the observed rate and the expected rate in the null-oscillation hypothesis, respectively, and $\sigma^2_{{\rm stat},i}$ is the statistical uncertainty for the i-th point. The expected rate is multiplied by several dimensionless nuisance parameters to allow for shifts due to the detection efficiency ($\alpha_{\rm d}$), the reactor-on flux prediction ($\alpha_{\rm r}$) or the residual $\overline{\nu}_e$ from the stopped reactor ($\alpha_{\rm res}$). For the latter, the weight w_i is defined as $R_{{\rm res},i}/R_{\nu,i}$; where $R_{{\rm res},i}$ is the expected rate of the residual $\overline{\nu}_e$ from the stopped reactor without oscillation, which is computed with the same simulation used to estimate the number of $\overline{\nu}_e$ when the two reactors are off [170].

Concerning the reactor-on flux prediction, all the contributions to its systematic uncertainty can be taken as independent of the thermal power, except the thermal power itself (P_{th}). In section 2.4.1, the relative uncertainty on the thermal power was said to amount to 0.5% when the reactor is running at full power; but the RRM analysis makes explicit use of the data taken at different thermal powers, for which this uncertainty is known to increase as the thermal power decreases. This motivates to define the weight k_i as $\sigma_{{\rm r},i}/\sigma_{\rm r}$, where $\sigma_{{\rm r},i}$ is the relative reactor flux systematic uncertainty at the power corresponding to the i-th point and $\sigma_{\rm r}$ is the relative uncertainty at full power. In order to

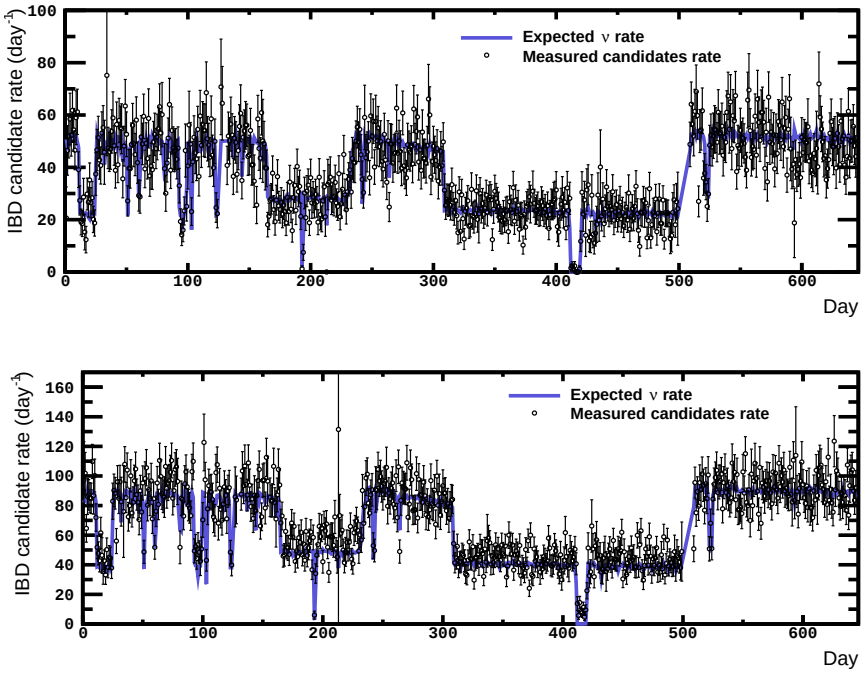


Figure 6.1: Daily rate of IBD candidates for the Gd-based (top) and H-based (bottom) $\overline{\nu}_e$ selections since the beginning of data taking. The data is represented by empty black circles (background not subtracted), superimposed to the expectation from the $\overline{\nu}_e$ MC simulation (blue line) without oscillation. The data cluster around three rate levels corresponding to the circumstances in which the two reactors are on (~ 50 events/day in Gd, ~ 90 events/day in H), one reactor is on and the other is off (~ 25 events/day in Gd, ~ 45 events/day in H), and both reactors are off (~ 0 events/day). From [164, 195].

estimate $\sigma_{\mathrm{r},i}$, an empirical model is fitted to a set of measurements made by the power plant operators. This results in a $1/P_{th}$ dependence, with a maximum $\sigma_{\mathrm{r},i} = 1.9\%$ when either one or the two reactors are not at full power and a minimum of $\sigma_{\mathrm{r},i} = \sigma_{\mathrm{r}} = 1.7\%$ when both are running at full power.

For the two reactors off period, as a result of the low number of events, χ^2_{off} is built as a binned Poisson extended likelihood:

$$\chi^2_{\mathrm{off}} = 2\left[N_{\mathrm{obs}} \ln\left(\frac{N_{\mathrm{obs}}}{N_{\mathrm{exp}}}\right) + N_{\mathrm{exp}} - N_{\mathrm{obs}}\right], \tag{6.5}$$

where N_{obs} is the number of IBD candidates observed in the reactors-off sample (see table 6.1); and the expectation is given by

$$N_{\mathrm{exp}} = \left[\left(1 - \sin^2(2\theta_{13})\eta_{\mathrm{osc}}\right) R_{\mathrm{res,off}} \left(1 + \alpha_{\mathrm{d}} + \alpha_{\mathrm{res}}\right) + B\right] \cdot T_{\mathrm{off}},$$

where $R_{\mathrm{res,off}}$ is the expected rate of residual $\overline{\nu}_e$ from the two stopped reactors in absence of oscillation and T_{off} is the total live-time of the reactors-off period.

Finally, χ^2_{pull}, the term for the systematic uncertainties is written as

$$\chi^2_{\mathrm{pull}} = \left(\frac{\alpha_{\mathrm{d}}}{\sigma_{\mathrm{d}}}\right)^2 + \left(\frac{\alpha_{\mathrm{r}}}{\sigma_{\mathrm{r}}}\right)^2 + \left(\frac{\alpha_{\mathrm{res}}}{\sigma_{\mathrm{res}}}\right)^2, \tag{6.6}$$

where σ_{d} is the relative uncertainty on the detection efficiency (0.63% for the Gd-based $\overline{\nu}_e$ selection, from table 5.3; and 1.02% for the H-based, from table 5.6) and σ_{res} is the relative uncertainty on the rate of residual $\overline{\nu}_e$ from the stopped reactor, which amounts to 30% (cf. table 6.1). The systematic uncertainty on the reactor-on flux, already introduced above ($\sigma_{\mathrm{r}} = 1.7\%$), is considered conservatively as correlated between the data points since the dominant contribution is the uncertainty on the measurement of the mean cross-section per fission by the Bugey4 experiment [144], which is independent of the thermal power.

Notice that so far, the background rate B is a free parameter. This is one of the most remarkable features of the RRM analysis, that is, it is capable of measuring the total background rate in a background-model-independent way.

Alternatively, in order to exploit all the information available to produce the most precise measurement of θ_{13}; B_{exp}, the total background rate estimation arising from the individual estimations using the IBD candidate DATA (table 4.2 for the Gd selection, table 4.4 for the H selection) can be included in the fit at the expense of losing the background-model independence. In this case, χ^2_{pull} is

$$\chi^2_{\mathrm{pull}} = \left(\frac{\alpha_{\mathrm{d}}}{\sigma_{\mathrm{d}}}\right)^2 + \left(\frac{\alpha_{\mathrm{r}}}{\sigma_{\mathrm{r}}}\right)^2 + \left(\frac{\alpha_{\mathrm{res}}}{\sigma_{\mathrm{res}}}\right)^2 + \left(\frac{B - B_{\mathrm{exp}}}{\sigma_B}\right)^2, \tag{6.7}$$

where σ_B is the uncertainty on B_{exp}.

6.1.1 Gd results

After introducing the fundamentals of the RRM analysis in the previous section, it can be applied to the particular case of the Gd-based $\overline{\nu}_e$ selection.

Firstly, in order to obtain a measurement of θ_{13} independent of the background model, B is treated as a free parameter. A global scan is executed on the $\sin^2(2\theta_{13}) - B$ plane, in which the χ^2 (see equation 6.3, with χ^2_{pull} given by equation 6.6) is minimized at each point with respect to the three nuisance parameters (α_{d}, α_{r}, α_{res}). The minimum χ^2 value is found at $\sin^2(2\theta_{13}) = 0.060 \pm 0.039$

and $B = 0.93^{+0.43}_{-0.36}$ events/day (see figure 6.2), where the confidence interval is defined by $\chi^2 - \chi^2_{\rm min} < 1.0$, and $\chi^2_{\rm min}$/d.o.f. = 1.9/5. The background rate result agrees with the expectation $B_{\rm exp} = 1.64^{+0.41}_{-0.17}$ events/day within 1.5σ, which confirms the reliability of the background model.

This RRM θ_{13} result can be compared with the previous RRM Gd result [74] (see equation 1.65), showing a 20% reduction of the uncertainty on $\sin^2(2\theta_{13})$ as a consequence of the increase in statistics (which have been doubled) and the detection efficiency uncertainty (which has decreased a 40%). The improvement is limited by the reactor-flux uncertainty, which is the dominant contribution and cannot be reduced in this single-detector phase.

Secondly, to increment the precision on θ_{13}, the background expectation is added to the χ^2 as in equation 6.7. Now a χ^2 scan is performed over the possible values of $\sin^2(2\theta_{13})$, minimizing the χ^2 with respect to the three nuisance parameters and the background rate (B). The best fit is found at $\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035}$ with $\chi^2_{\rm min}$/d.o.f. = 4.2/6. In this case, the best-fit background rate results in $B = 1.56^{+0.18}_{-0.16}$ events/day. Figure 6.3 shows the result of the fit.

6.1.2 H results

The application to the RRM analysis to the H-based $\overline{\nu}_e$ selection is straightforward, since it only requires to substitute the appropriate inputs.

As in the Gd case, θ_{13} can be measured in a background-model-independent way by treating the background rate B as a free parameter in the χ^2 (equation 6.3, with $\chi^2_{\rm pull}$ given by equation 6.6). A global scan on the $\sin^2(2\theta_{13}) - B$ plane, in which the χ^2 is minimized at each point with respect to the three nuisance parameters ($\alpha_{\rm d}$, $\alpha_{\rm r}$, $\alpha_{\rm res}$), finds the best fit at $\sin^2(2\theta_{13}) = 0.120^{+0.042}_{-0.043}$ and $B = 8.23^{+0.88}_{-0.87}$ events/day, where the quoted uncertainty corresponds to the interval for which $\chi^2 - \chi^2_{\rm min} < 1.0$, with $\chi^2_{\rm min}$/d.o.f. = 5.6/5 (see figure 6.4). The best-fit background rate is consistent with the expected rate according to the background model, $B_{\rm exp} = 6.83^{+0.59}_{-0.36}$ events/day, agreeing within 1.3σ.

This new RRM $\sin^2(2\theta_{13})$ measurement has an uncertainty 46% smaller than the previous RRM H result [74] (see equation 1.66). This is a remarkable improvement driven by the reduction of the statistical uncertainty (from 1.1% to 0.6%), not only due to the doubled statistics but also to the higher signal-to-background ratio (which has passed from 0.9 to 9.7); and the reduction of the detection efficiency uncertainty (from 1.6% to 1.02%).

However, the most precise RRM measurement of θ_{13} with the H-based $\overline{\nu}_e$ selection is obtained when the background rate is constrained by the estimation. In this case, the χ^2 is built using the $\chi^2_{\rm pull}$ from equation 6.7; and then it is minimized with respect to the three nuisance parameters and the background rate (B), for the possible values of $\sin^2(2\theta_{13})$. The best fit is found for $\sin^2(2\theta_{13}) = 0.095^{+0.038}_{-0.039}$, with $\chi^2_{\rm min}$/d.o.f. = 7.4/6 and a background rate of $B = 7.27 \pm 0.49$ events/day (see figure 6.5).

These results represent a quantitative leap forward, since for the first time the precision on $\sin^2(2\theta_{13})$ reached with a H-based $\overline{\nu}_e$ selection (for which the detector is not optimized) is comparable to the precision of the Gd-based selection.

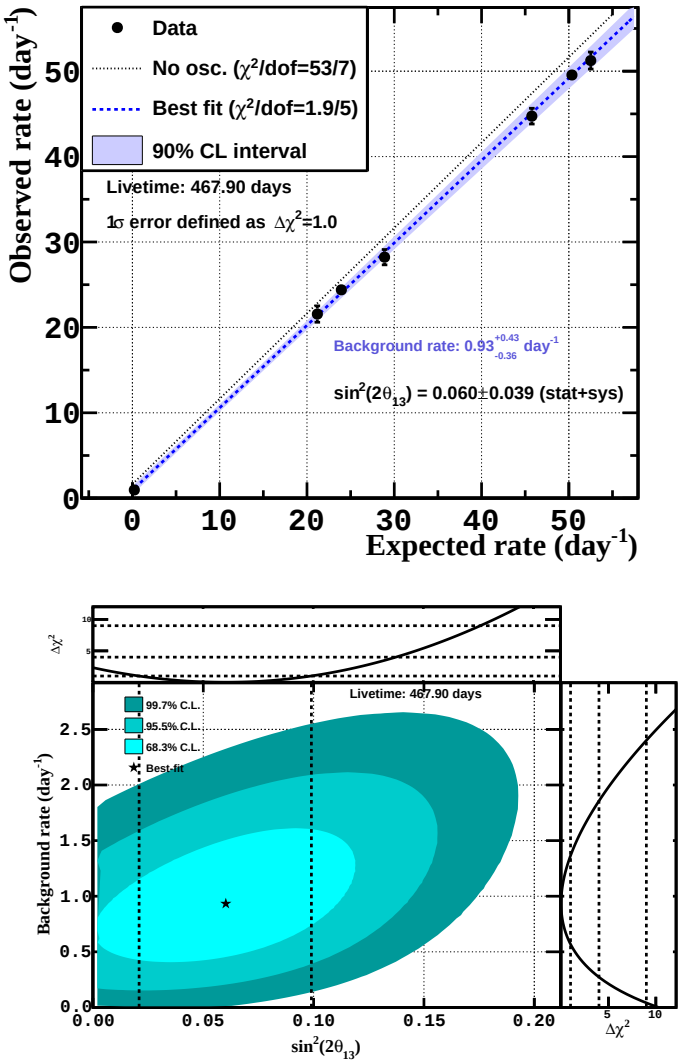


Figure 6.2: RRM Gd results treating the background rate as a free parameter. Top: Observed IBD candidate rate versus the expected $\overline{\nu}_e$ IBD rate in case of no oscillation. The best fit to the data (black circles) is shown as a blue dashed line (90% C.L. interval displayed as the shaded region). The prediction for the null-oscillation hypothesis is shown as a black dotted line. Bottom: 68.3%, 95.5% and 99.7% allowed regions on the $\sin^2(2\theta_{13}) - B$ plane. The best fit is marked by a star. The one-dimensional $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ profiles are shown in the upper and right panels. The two parallel vertical lines in the central panel show the interval of $\sin^2(2\theta_{13})$ for which $\Delta\chi^2 < 1.0$. From [196].

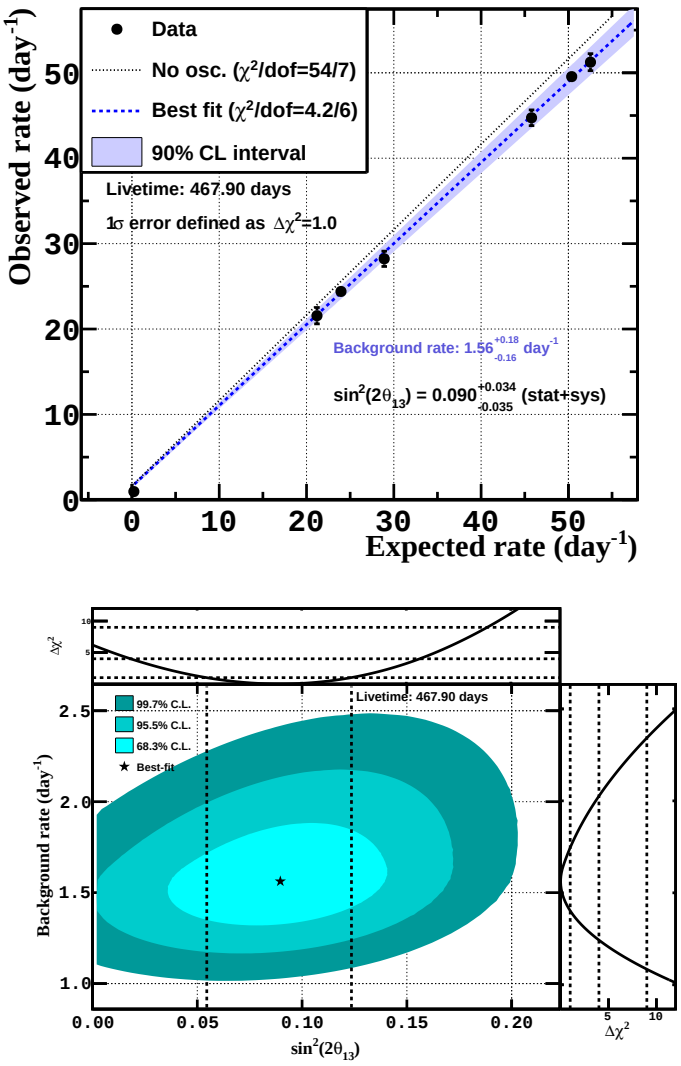


Figure 6.3: RRM Gd results constraining the background rate by its estimation. All other features of the figure are as in figure 6.2. From [196].

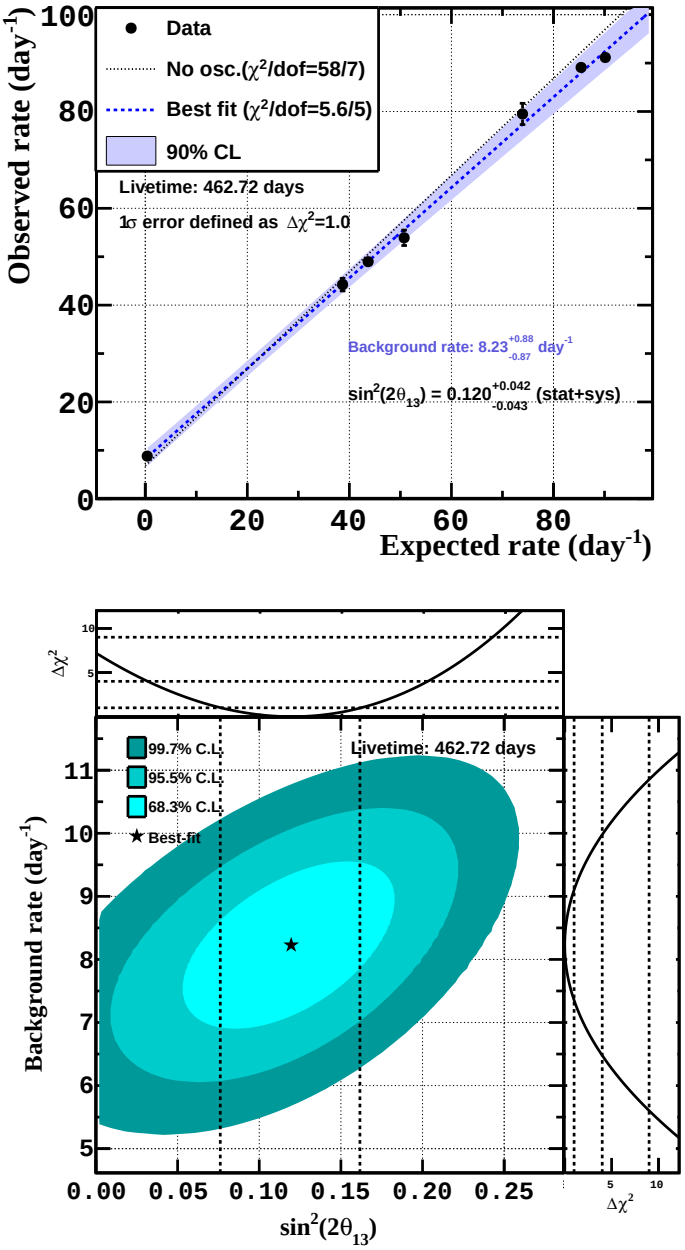


Figure 6.4: RRM H results treating the background rate as a free parameter. Top: Observed IBD candidate rate versus the expected $\overline{\nu}_e$ IBD rate in case of no oscillation. The best fit to the data (black circles) is shown as a blue dashed line (90% C.L. interval displayed as the shaded region). The prediction for the null-oscillation hypothesis is shown as a black dotted line. Bottom: 68.3%, 95.5% and 99.7% allowed regions on the $\sin^2(2\theta_{13}) - B$ plane. The best fit is marked by a star. The one-dimensional $\Delta\chi^2 = \chi^2 - \chi^2_{\rm min}$ profiles are shown in the upper and right panels. The two parallel vertical lines in the central panel show the interval of $\sin^2(2\theta_{13})$ for which $\Delta\chi^2 < 1.0$. From [197].

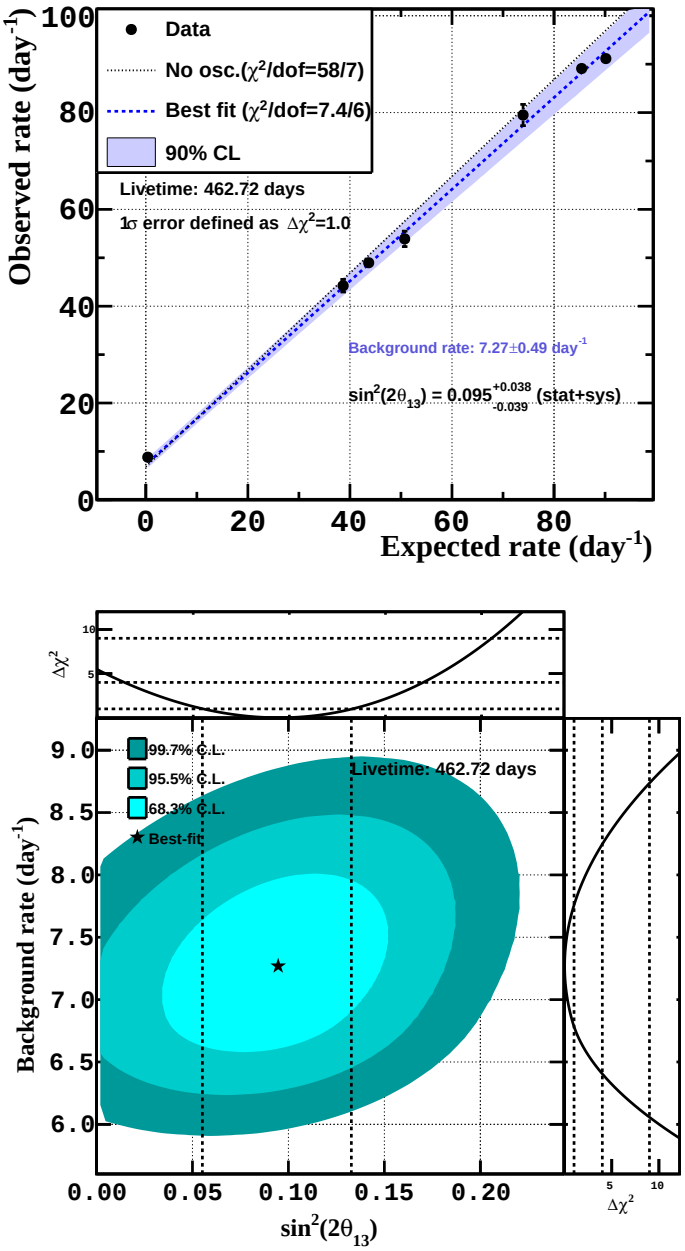


Figure 6.5: RRM H results constraining the background rate by its estimation. All other features of the figure are as in figure 6.4. From [197].

6.1.3 Combined Gd and H results

In order to boost the precision on the measurement of θ_{13}, the RRM analysis can be performed simultaneously over the Gd and H data sets. The χ^2 decomposition (equation 6.3) into three terms can be kept; with $\chi^2_{\rm on}$, $\chi^2_{\rm off}$ and $\chi^2_{\rm pull}$ corresponding to data with at least one reactor on, data with the two reactors off and a specific term for the systematic uncertainties, respectively. But each term must be modified with respect to the previous section to accommodate the two $\overline{\nu}_e$ selections.

The $\chi^2_{\rm on}$ for the combined Gd and H fit is written as

$$\begin{aligned}\chi^2_{\rm on} &= \\ &= \sum_i \frac{\left[R^{\rm Gd}_{{\rm obs},i} - \langle P_{\rm surv}\rangle R^{\rm Gd}_{\nu,i}\left(1+\alpha_{\rm d}^{\rm Gd,unc}+\alpha_{\rm d}^{\rm cor}+k_i\alpha_{\rm r}+w_i\alpha_{\rm res}\right) - B^{\rm Gd}\right]^2}{\sigma^2_{{\rm stat},i}} + \\ &+ \sum_j \frac{\left[R^{\rm H}_{{\rm obs},j} - \langle P_{\rm surv}\rangle R^{\rm H}_{\nu,j}\left(1+\alpha_{\rm d}^{\rm H,unc}+\alpha_{\rm d}^{\rm cor}+k_j\alpha_{\rm r}+w_j\alpha_{\rm res}\right) - B^{\rm H}\right]^2}{\sigma^2_{{\rm stat},j}},\end{aligned} \tag{6.8}$$

where the summation index i runs over the 6 data points of the Gd selection, and the summation index j runs over the 6 data points of the H selection. Notice that this $\chi^2_{\rm on}$ is not a mere duplication of that from equation 6.4: besides the oscillation parameters of the averaged survival probability $\langle P_{\rm surv}\rangle = (1 - \sin^2(2\theta_{13})\eta_{\rm osc})$ which must be the same regardless the nuclei chosen to capture the IBD neutron, the nuisance parameters for the reactor-on flux prediction ($\alpha_{\rm r}$) and the residual $\overline{\nu}_e$ from the stopped reactor ($\alpha_{\rm res}$) are also the same for the Gd and H terms. This is due to the fact that both parameters refer only to the reactor simulations, which are shared by the two $\overline{\nu}_e$ selections, so the systematic uncertainties associated to them are fully correlated. In addition, for each $\overline{\nu}_e$ selection, the nuisance parameter accounting for the detection efficiency in equation 6.4 has been replaced by two parameters, one which is selection-dependent to allow for uncorrelated shifts ($\alpha_{\rm d}^{X,{\rm unc}}$, where X is Gd or H) and another which is common to both selections to describe the correlated effects ($\alpha_{\rm d}^{\rm cor}$).

The $\chi^2_{\rm off}$ is simply

$$\chi^2_{\rm off} = \chi^2_{\rm off,Gd} + \chi^2_{\rm off,H}, \tag{6.9}$$

where each term is given by equation 6.5, except that the expected number of events is changed to

$$N^X_{\rm exp} = \left[\left(1-\sin^2(2\theta_{13})\eta_{\rm osc}\right) R^X_{\rm res,off}\left(1+\alpha_{\rm d}^{X,{\rm unc}}+\alpha_{\rm d}^{\rm cor}+\alpha_{\rm res}\right)+B^X\right]\cdot T^X_{\rm off},$$

where X denotes the Gd or H variables.

Finally, the term for the systematic uncertainties can be written in two different forms, depending on whether the background rates $B^{\rm Gd}$ and $B^{\rm H}$ are treated as free parameters or not. On the one hand, if they are treated as free parameters, then the analysis is background-model independent and $\chi^2_{\rm pull}$ is written

$$\chi^2_{\rm pull} = \left(\frac{\alpha_{\rm d}^{\rm Gd,unc}}{\sigma_{\rm d}^{\rm Gd,unc}}\right)^2 + \left(\frac{\alpha_{\rm d}^{\rm H,unc}}{\sigma_{\rm d}^{\rm H,unc}}\right)^2 + \left(\frac{\alpha_{\rm d}^{\rm cor}}{\sigma_{\rm d}^{\rm cor}}\right)^2 + \left(\frac{\alpha_{\rm r}}{\sigma_{\rm r}}\right)^2 + \left(\frac{\alpha_{\rm res}}{\sigma_{\rm res}}\right)^2, \tag{6.10}$$

where $\sigma_{\rm d}^{\rm Gd,unc}$ and $\sigma_{\rm d}^{\rm H,unc}$ are the uncorrelated relative uncertainties on the detection efficiency of the Gd and H selections, respectively; and $\sigma_{\rm d}^{\rm cor}$ is the correlated component. Their specific values depend on the degree of correlation estimated.

In this case, a global scan over the three-dimensional space defined by $\sin^2(2\theta_{13})$, $B^{\rm Gd}$ and $B^{\rm H}$ is carried out, in which the χ^2 is minimized with respect to the five nuisance parameters ($\alpha_{\rm d}^{\rm Gd,unc}$, $\alpha_{\rm d}^{\rm H,unc}$, $\alpha_{\rm d}^{\rm cor}$, $\alpha_{\rm r}$, $\alpha_{\rm res}$). Initially, the detection efficiency uncertainties are considered to be fully uncorrelated, which is equivalent to fixing $\alpha_{\rm d}^{\rm cor}$ to zero and $\sigma_{\rm d}^{\rm Gd,unc} = \sigma_{\rm d}^{\rm Gd}$ and $\sigma_{\rm d}^{\rm H,unc} = \sigma_{\rm d}^{\rm H}$. The minimum χ^2 is found at $\sin^2(2\theta_{13}) = 0.075 \pm 0.037$, $B^{\rm Gd} = 1.21^{+0.41}_{-0.38}$ events/day and $B^{\rm H} = 7.44^{+0.69}_{-0.68}$ events/day (see figure 6.6), where the confidence interval is defined by $\chi^2 - \chi^2_{\rm min} < 1.0$, and $\chi^2_{\rm min}/{\rm d.o.f.} = 10/11$. As in the individual fits, the resulting background rates are in good agreement (within less than 1σ) with the background-model expectations: $B^{\rm Gd}_{\rm exp} = 1.64^{+0.41}_{-0.17}$ events/day and $B^{\rm H}_{\rm exp} = 6.83^{+0.59}_{-0.36}$ events/day. To test the impact of the correlation between the detection efficiency uncertainties, the analysis was repeated assuming a 30% correlation between the two (the maximum possible because of the detection volume shared by the two $\overline{\nu}_e$ selections) and no effect was observed. This was foreseen since the systematic uncertainty is largely dominated by the reactor systematic uncertainty.

The $\sin^2(2\theta_{13}) = 0.075 \pm 0.037$ just obtained is the most precise background-model-independent measurement of θ_{13}, since this analysis is unique from Double Chooz. This result is 14% more precise than the previous combination of the Gd and H data sets [74] (see equation 1.67).

On the other hand, if the goal is to maximize the precision on θ_{13}, then the expected total background rates from the background models can be included in the χ^2. Henceforth, the background rates are treated as nuisance parameters, and $\chi^2_{\rm pull}$ is written as

$$\begin{aligned}\chi^2_{\rm pull} = &\left(\frac{\alpha_{\rm d}^{\rm Gd,unc}}{\sigma_{\rm d}^{\rm Gd,unc}}\right)^2 + \left(\frac{\alpha_{\rm d}^{\rm H,unc}}{\sigma_{\rm d}^{\rm H,unc}}\right)^2 + \left(\frac{\alpha_{\rm d}^{\rm cor}}{\sigma_{\rm d}^{\rm cor}}\right)^2 + \left(\frac{\alpha_{\rm r}}{\sigma_{\rm r}}\right)^2 + \left(\frac{\alpha_{\rm res}}{\sigma_{\rm res}}\right)^2 + \\ &+ \left(B^{\rm Gd} - B^{\rm Gd}_{\rm exp}, B^{\rm H} - B^{\rm H}_{\rm exp}\right) \cdot \\ &\cdot \begin{pmatrix} (\sigma_{\rm B}^{\rm Gd})^2 & \rho_{\rm B}\sigma_{\rm B}^{\rm Gd}\sigma_{\rm B}^{\rm H} \\ \rho_{\rm B}\sigma_{\rm B}^{\rm Gd}\sigma_{\rm B}^{\rm H} & (\sigma_{\rm B}^{\rm H})^2 \end{pmatrix}^{-1} \begin{pmatrix} B^{\rm Gd} - B^{\rm Gd}_{\rm exp} \\ B^{\rm H} - B^{\rm H}_{\rm exp} \end{pmatrix}, \end{aligned} \tag{6.11}$$

where $B^{\rm Gd}_{\rm exp}$ and $B^{\rm H}_{\rm exp}$ are the estimated total background rates for the Gd and H selections, respectively; and $\sigma_{\rm B}^{\rm Gd}$ and $\sigma_{\rm B}^{\rm H}$ their corresponding uncertainties, with $\rho_{\rm B}$ their correlation coefficient.

The total χ^2 is now minimized with respect to the seven nuisance parameters ($\alpha_{\rm d}^{\rm Gd,unc}$, $\alpha_{\rm d}^{\rm H,unc}$, $\alpha_{\rm d}^{\rm cor}$, $\alpha_{\rm r}$, $\alpha_{\rm res}$, $B^{\rm Gd}$, $B^{\rm H}$). Assuming fully uncorrelated uncertainties on the detection efficiency and the background rates (i.e. $\rho_{\rm B} = 0$), a scan over possible values of $\sin^2(2\theta_{13})$ finds the best fit at $\sin^2(2\theta_{13}) = 0.088 \pm 0.033$, with $\chi^2_{\rm min}/{\rm d.o.f.} = 11/13$ (see figure 6.7). The best-fit background rates are $B^{\rm Gd} = 1.57 \pm 0.16$ events/day and $B^{\rm H} = 7.22 \pm 0.43$ events/day. Assuming fully correlated uncertainties on the detection efficiency or the background rates instead has a negligible effect on the θ_{13} result. Again, this was expected because the reactor-flux uncertainty prevails over the rest of the uncertainties.

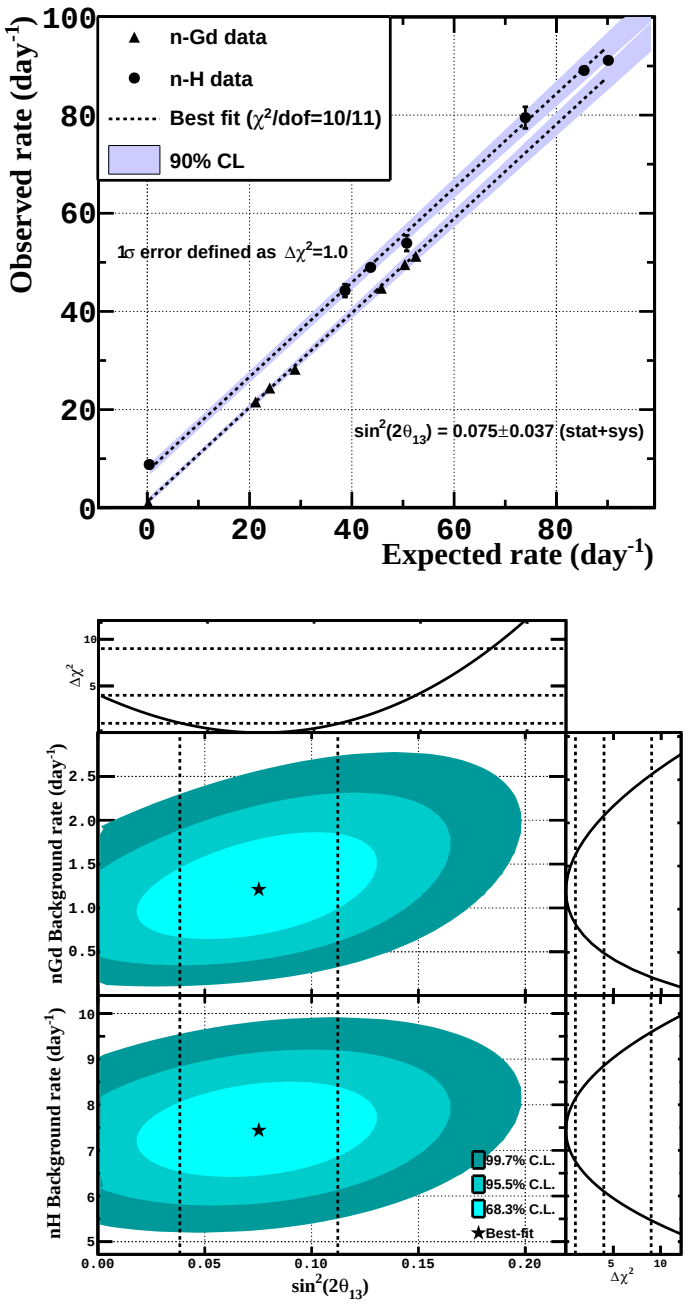


Figure 6.6: RRM Gd and H combined results treating the background rates as free parameters. Top: Observed IBD candidate rate versus the expected $\overline{\nu}_e$ IBD rate in absence of oscillation, for the Gd-based (black triangles) and the H-based (black circles) $\overline{\nu}_e$ selections. The best fit to the two samples is shown as dashed lines (with a shaded region corresponding to the 90% C.L. interval). Bottom: 68.3%, 95.5% and 99.7% allowed regions on the $\sin^2(2\theta_{13}) - B^{\mathrm{Gd}}$ and $\sin^2(2\theta_{13}) - B^{\mathrm{H}}$ planes. The best fit is marked by a star. The one-dimensional $\Delta\chi^2 = \chi^2 - \chi^2_{\mathrm{min}}$ profiles are shown in the uppermost and right panels. The two parallel vertical lines in the central panels show the interval of $\sin^2(2\theta_{13})$ for which $\Delta\chi^2 < 1.0$. From [197].

This measurement, $\sin^2(2\theta_{13}) = 0.088 \pm 0.033$, is the most precise of all the RRM measurements. The best-fit background rates are also the most precise of their kind, improving on the input rates: from $B^{\mathrm{Gd}}_{\mathrm{exp}} = 1.64^{+0.41}_{-0.17}$ events/day to $B^{\mathrm{Gd}} = 1.57 \pm 0.16$ events/day; and from $B^{\mathrm{H}}_{\mathrm{exp}} = 6.83^{+0.59}_{-0.36}$ events/day to $B^{\mathrm{H}} = 7.22 \pm 0.43$ events/day.

To conclude this section, figure 6.8 presents a summary of all the measurements of $\sin^2(2\theta_{13})$ obtained with the RRM analysis. The agreement is excellent even between the Gd and H disjoint samples.

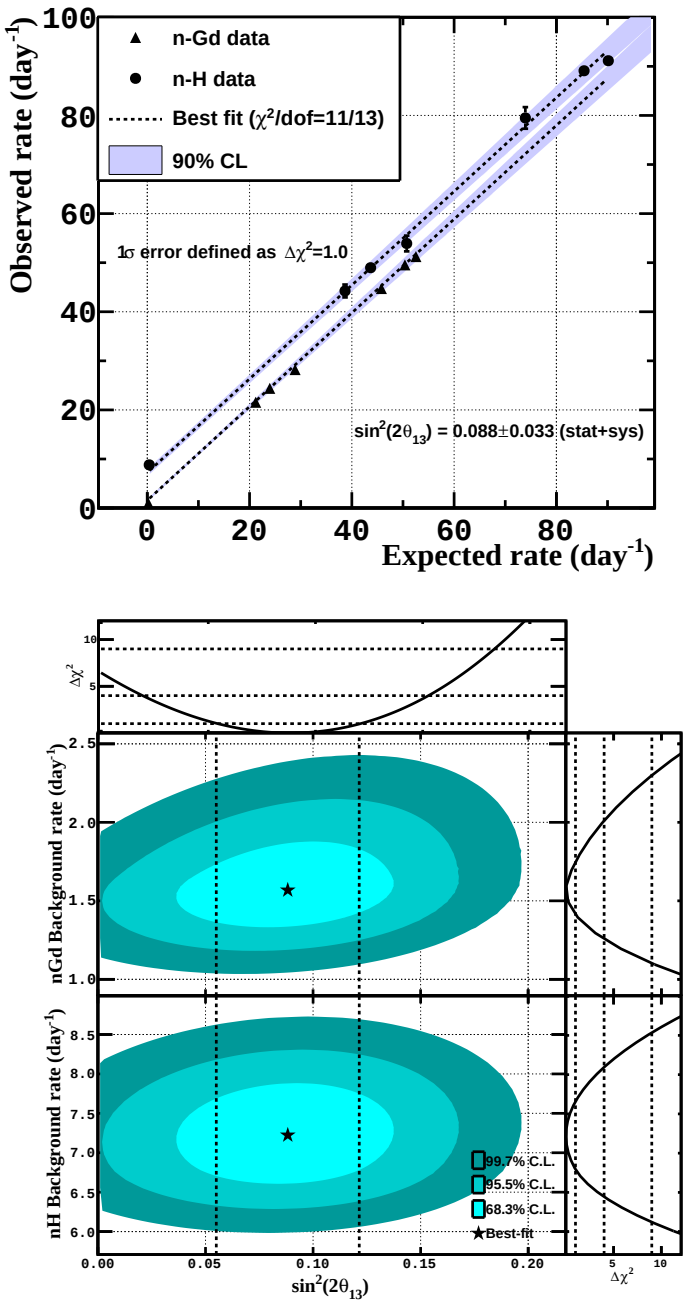


Figure 6.7: RRM Gd and H combined results constraining the background rates by their estimations. All other features of the figure are as in figure 6.6. From [197].

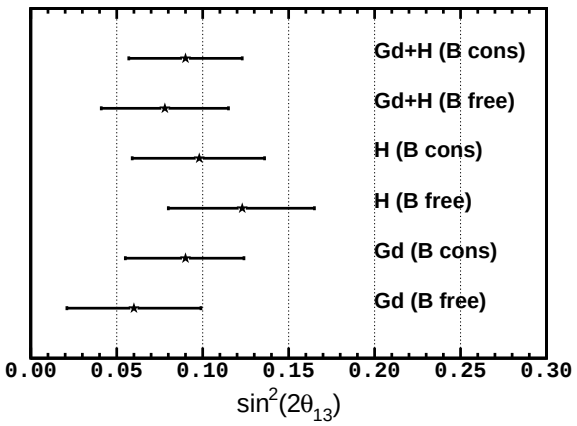


Figure 6.8: Compilation of the measurements of $\sin^2(2\theta_{13})$ obtained with the Reactor Rate Modulation analysis. Error bars show the total uncertainty, defined as the interval in which $\chi^2 - \chi^2_{\rm min} < 1.0$. A label next to each result indicates the $\overline{\nu}_e$ selection used and whether the background rate was treated as a free parameter or was constrained by the estimation.

6.2 Rate+Shape analysis

The Rate+Shape analysis studies the $\overline{\nu}_e$ deficit as a function of the prompt trigger visible energy, which is directly related to the energy of the $\overline{\nu}_e$ (see equation 2.4). The advantage of this approach is double. First, since the oscillation probability is energy-dependent (see equation 1.54), the $\overline{\nu}_e$ conversion to other flavors leaves a disappearance pattern on the energy spectrum which is specific to neutrino oscillation and whose magnitude is driven by $\sin^2(2\theta_{13})$. Second, the background events which are capable of mimicking the IBD interaction have prompt spectra which differ from the actual IBD one, so the contribution of each to the observed spectrum can be statistically determined. This feature has been enhanced by extending the range of selected prompt trigger energies up to 20 MeV, far beyond the expected end of the $\overline{\nu}_e$ spectrum with the current statistics (~ 9 MeV), in order to collect a pure sample of background events of ^{9}Li/^{8}He (with an endpoint at 13.6 MeV) and fast neutrons and stopping muons (which continue up to higher energies) to constrain their normalization. Both the energy dependence and the background-normalization constrain improve the precision on θ_{13}.

In order to find the value of θ_{13} which gives the best fit of the predicted spectrum to the observed spectrum, the following χ^2 is used:

$$\begin{aligned}
\chi^2 = & \sum_i \sum_j \left(N_i^{\text{pred}} - N_i^{\text{obs}}\right)\ M_{ij}^{-1}\ \left(N_j^{\text{pred}} - N_j^{\text{obs}}\right) + \\
& + \left(\frac{\Delta m^2_{31} - \Delta m^2_{\text{MINOS}}}{\sigma_{\text{MINOS}}}\right)^2 + \left(\frac{\alpha_{\text{res}} - 1}{\sigma_{\text{res}}}\right)^2 + \\
& + \left(\frac{\alpha_{\text{Fn+S}\mu} - 1}{\sigma_{\text{Fn+S}\mu}}\right)^2 + \left(\frac{\alpha_{\text{Li+He}} - 1}{\sigma_{\text{Li+He}}}\right)^2 + \left(\frac{\alpha_{\text{acc}} - 1}{\sigma_{\text{acc(syst)}}}\right)^2 + \\
& + (a' - a'_{\text{CV}}, b' - b'_{\text{CV}}, c' - c'_{\text{CV}}) \cdot \\
& \cdot \begin{pmatrix} \sigma^2_{a'} & \rho_{a'b'}\sigma_{a'}\sigma_{b'} & \rho_{a'c'}\sigma_{a'}\sigma_{c'} \\ \rho_{b'a'}\sigma_{b'}\sigma_{a'} & \sigma^2_{b'} & \rho_{b'c'}\sigma_{b'}\sigma_{c'} \\ \rho_{c'a'}\sigma_{c'}\sigma_{a'} & \rho_{c'b'}\sigma_{c'}\sigma_{b'} & \sigma^2_{c'} \end{pmatrix}^{-1} \begin{pmatrix} a' - a'_{\text{CV}} \\ b' - b'_{\text{CV}} \\ c' - c'_{\text{CV}} \end{pmatrix} + \\
& + \ 2 \left[N_{\text{off}}^{\text{obs}} \ln\left(\frac{N_{\text{off}}^{\text{obs}}}{N_{\text{off}}^{\text{pred}}}\right) + N_{\text{off}}^{\text{pred}} - N_{\text{off}}^{\text{obs}} \right],
\end{aligned} \tag{6.12}$$

where the sum in i and j is over the bins in which the prompt spectrum has been divided (see table 6.2). N_i^{obs} and N_i^{pred} are the observed and the predicted number of IBD candidates in the bin i-th, respectively. The latter is given by

$$\begin{aligned}
N_i^{\text{pred}} = & \sum_{R=1,2} P_{i,R}^{\text{surv}}\left(\sin^2(2\theta_{13}), \Delta m^2_{31}\right) N_{i,R}^{\nu}\left(a', b', c'\right) + \\
& + \ \alpha_{\text{Fn+S}\mu} N_i^{\text{Fn+S}\mu} + \alpha_{\text{Li+He}} N_i^{\text{Li+He}} + \alpha_{\text{acc}} N_i^{\text{acc}},
\end{aligned} \tag{6.13}$$

where $P_{i,R}^{\text{surv}}\left(\sin^2(2\theta_{13}), \Delta m^2_{31}\right)$ is the survival probability of the $\overline{\nu}_e$ originating at the reactor R and ending up in the i-th bin, which is a function of $\sin^2(2\theta_{13})$ and Δm^2_{31} (as in equation 1.54). In order to account for the uncertainty on Δm^2_{31}, Δm^2_{31} is implemented in the χ^2 as a nuisance parameter, constrained to the central value from the MINOS experiment $\Delta m^2_{\text{MINOS}}$ by the uncertainty on

Interval (MeV)	Bin width (MeV)	Number of bins
$[0.5-8)$ (Gd) ; $[1-8)$ (H)	0.25	30 (Gd) ; 28 (H)
$[8-10)$	0.50	4
$[10-12)$	1.00	2
$[12-20)$	2.00	4

Table 6.2: Binning of the prompt trigger visible energy spectrum used in the Rate+Shape analysis. The total number of bins is different in the Gd-based (40 bins) and the H-based (38 bins) $\overline{\nu}_e$ selections due to the raised minimum energy for the latter (see section 4.2.2).

the measurement, σ_{MINOS} [59]. Since the MINOS measurement has an asymmetric uncertainty, the specific value of σ_{MINOS} depends on whether Δm^2_{31} is above or below the central value. $N^{\nu}_{i,R}$ is the expected number of IBD interactions of the $\overline{\nu}_e$ emitted from the reactor R in the null-oscillation hypothesis (the result of integrating equation 2.12 over the run-time of the detector). The summation index R runs over the two reactor cores of the Chooz power plant. $N^{\nu}_{i,R}$ is also a function of the three coefficients, a', b' and c', used to modelize the visible energy of the MC simulation in the fit (equation 3.22a). a', b' and c' are treated in the χ^2 of equation 6.12 as nuisance parameters, which are constrained to their central values a'_{CV}, b'_{CV} and c'_{CV} by their respective uncertainties $\sigma_{a'}$, $\sigma_{b'}$ and $\sigma_{c'}$ (see table 3.3 for the central values and their uncertainties). In order to take into account the existing correlation between them, a 3×3 covariance matrix is used, in which the correlation coefficients $\rho_{x'y'}$ are those of the correlation matrix in equation 3.23.

The second line of equation 6.13 is the contribution of the background sources to the i-th bin of the prompt energy spectrum of the predicted number of IBD candidates; where $N_i^{\text{Fn+S}\mu}$, $N_i^{\text{Li+He}}$, and N_i^{acc} are the expected numbers of fast neutrons and stopping muons, ^{9}Li and ^{8}He cosmogenic isotopes, and accidental events, respectively, according to the spectra and rates estimated in sections 4.1.3 (Gd-based selection) and 4.2.3 (H-based selection), which are summarized in tables 4.2 and 4.4. Each contribution is modulated by a scaling factor, $\alpha_{\text{Li+He}}$, $\alpha_{\text{Fn+S}\mu}$, α_{acc}, to allow for shifts. The scaling factors are included in the χ^2 of equation 6.12 as nuisance parameters, centered at 1 and constrained by the relative uncertainties on the background rates ($\sigma_{\text{Fn+S}\mu}$, $\sigma_{\text{Li+He}}$, $\sigma_{\text{acc(syst)}}$). Since the ^{9}Li and ^{8}He background rate has an asymmetric uncertainty, the same consideration as for Δm^2_{31} applies. In addition, the accidental background rate is only constrained by the systematic uncertainty $\sigma_{\text{acc(syst)}}$, since this is the component which is fully correlated across all bins. The statistical uncertainty, which is uncorrelated between bins, is included as a diagonal covariance matrix $M_{ij}^{\text{acc(stat)}}$ in the total covariance matrix M_{ij}.

The total covariance matrix, M_{ij}, is a sum of several covariance matrices,

$$M_{ij} = M_{ij}^{\text{stat}} + M_{ij}^{\text{r}} + M_{ij}^{\text{d}} + M_{ij}^{\text{Li+He(shape)}} + M_{ij}^{\text{acc(stat)}}, \qquad (6.14)$$

where $M_{ij}^{\text{stat}} = \delta_{ij}\sqrt{N_i^{\text{pred}} N_j^{\text{pred}}}$, where δ_{ij} is the Kronecker delta, accounts for the statistical uncertainty of the prediction in each bin. M_{ij}^{r} encodes the systematic uncertainty on the reactor-flux prediction, and it is obtained propagating the reactor-related uncertainties of equation 2.12. M_{ij}^{d} accounts for the

Source	Gd selection		H selection	
	Uncertainty (%)	Gd-III/Gd-II	Uncertainty (%)	H-III/H-II
Statistics	0.8	0.7	0.6	0.5
Reactor flux	1.7	1.0	1.7	1.0
Detection efficiency	0.6	0.6	1.0	0.6
Fast-n + stopping-μ	0.1	0.2	0.2	0.3
^{9}Li and ^{8}He	$^{+1.1}_{-0.4}$	0.5	$^{+0.9}_{-0.5}$	0.4
Accidental coincidences	< 0.1	—	< 0.1	—

Table 6.3: Normalization uncertainties relative to the signal prediction for the Gd-based and H-based $\overline{\nu}_e$ selections. The columns labeled Gd-III/Gd-II and H-III/H-II show the reduction of the uncertainty with respect to the previous analyses [1, 72].

uncertainty on the detection efficiency, and is written $M^{\mathrm{d}}_{ij} = \sigma^2_{\mathrm{d}} N^{\nu}_i N^{\nu}_j$, where $N^{\nu}_i = \sum_{R=1,2} P^{\mathrm{surv}}_{i,R}\left(\sin^2(2\theta_{13}), \Delta m^2_{31}\right) N^{\nu}_{i,R}\left(a', b', c'\right)$ and σ_{d} is the relative uncertainty on the detection efficiency (0.63% for the Gd-based $\overline{\nu}_e$ selection, from table 5.3; and 1.02% for the H-based, from table 5.6). Since the uncertainty on the ^{9}Li and ^{8}He rate has already been included as a pull term, $M^{\mathrm{Li+He(shape)}}_{ij}$ accounts for the uncertainty on the shape of the spectrum extracted following reference [198]. The covariance matrix for the statistical uncertainty on the accidental rate, $M^{\mathrm{acc(stat)}}_{ij}$, has just been explained in the previous paragraph. It becomes apparent that the elements of the covariance matrix are a function of θ_{13} or the other nuisance parameters of the χ^2 of equation 6.12. In order to maintain the accuracy, the elements are recalculated at each step of the χ^2 minimization.

The normalization component of the uncertainties affecting the analyses are summarized in table 6.3.

The last line of the χ^2 of equation 6.12 deals with the reactors-off period. Because of the low statistics, there is no gain in binning the data in energy, so as in the RRM analysis (equation 6.5) a one-bin extended Poisson likelihood is used, where $N^{\mathrm{obs}}_{\mathrm{off}}$ and $N^{\mathrm{pred}}_{\mathrm{off}}$ are the observed and predicted number of IBD candidates, respectively. The prediction is given by

$$\begin{aligned} N^{\mathrm{pred}}_{\mathrm{off}} = &\left(1 - \sin^2(2\theta_{13})\eta_{\mathrm{osc}}\right) \alpha_{\mathrm{res}} N^{\mathrm{pred}}_{\mathrm{res,off}} + \\ &+ \left(\alpha_{\mathrm{Fn+S}\mu} R_{\mathrm{Fn+S}\mu} + \alpha_{\mathrm{Li+He}} R_{\mathrm{Li+He}} + \alpha_{\mathrm{acc}} R_{\mathrm{acc}}\right) \cdot T_{\mathrm{off}} \end{aligned} \tag{6.15}$$

where $\eta_{\mathrm{osc}} \approx 0.75$ is the energy-baseline-averaged oscillatory factor (equation 6.2). $N^{\mathrm{pred}}_{\mathrm{res,off}}$ is the expected number of IBD interactions of the residual $\overline{\nu}_e$ emitted by the stopped reactors according to the simulation [170] in absence of oscillation (cf. table 6.1); and it is regulated by α_{res}, a nuisance parameter which enters the χ^2 centered at 1 and is allowed to vary constrained by the relative uncertainty on the expectation. The contribution of the backgrounds is added to that of the residual $\overline{\nu}_e$, where $R_{\mathrm{Fn+S}\mu}$, $R_{\mathrm{Li+He}}$ and R_{acc} are the estimated rates of fast neutrons and stopping muons, ^{9}Li and ^{8}He cosmogenic isotopes, and accidental events in this period, respectively; which are scaled by the same factors as in the reactor-on data, and T_{off} is the live-time of the reactors-off period.

Fit parameter	Input value	Best-fit value
Δm^2_{31} ($10^{-3}\,\mathrm{eV}^2$)	$2.44^{+0.09}_{-0.10}$	$2.44^{+0.09}_{-0.10}$
Residual $\overline{\nu}_e$	1.57 ± 0.47	1.48 ± 0.47
Fast-n + stopping-μ rate (events/day)	0.604 ± 0.051	$0.568^{+0.038}_{-0.037}$
Li+He rate (events/day)	$0.97^{+0.41}_{-0.16}$	0.74 ± 0.13
Accidental rate (events/day)	0.0701 ± 0.0026	0.0703 ± 0.0026
Energy-model a' (MeV)	-0.027 ± 0.006	$-0.026^{+0.006}_{-0.005}$
Energy-model b'	1.012 ± 0.008	$1.011^{+0.004}_{-0.006}$
Energy-model c' (MeV^{-1})	-0.0001 ± 0.0006	$-0.0006^{+0.0007}_{-0.0005}$

Table 6.4: Input values and best-fit values, and their corresponding uncertainties, of the nuisance parameters of the Rate+Shape Gd analysis. The residual $\overline{\nu}_e$ and background scaling factors of equation 6.12 have been multiplied to their respective expected number and nominal background rates. Notice the uncertainty reduction resulting from the fit.

6.2.1 Gd results

In order to obtain a measurement of θ_{13} with the Rate+Shape analysis applied to the Gd-based $\overline{\nu}_e$ selection, the data corresponding to this selection are used in the χ^2 of equation 6.12; and a scan over possible values of $\sin^2(2\theta_{13})$ is performed, in which the χ^2 is minimized with respect to the eight nuisance parameters (Δm^2_{31}, α_{res}, $\alpha_{\mathrm{Fn+S}\mu}$, $\alpha_{\mathrm{Li+He}}$, α_{acc}, a', b', c') for each value of $\sin^2(2\theta_{13})$. Assuming normal mass hierarchy, so $\Delta m^2_{\mathrm{MINOS}} = (2.44^{+0.09}_{-0.10}) \cdot 10^{-3}\,\mathrm{eV}^2$ from [59], the best fit is found at $\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029}$, with $\chi^2_{\mathrm{min}}/\mathrm{d.o.f.} = 52.2/40$ (see table 6.4 for the rest of the parameters). Figure 6.9 shows the best-fit prompt spectrum (top) and the ratio of the observed to the predicted $\overline{\nu}_e$ events (bottom), where the characteristic energy-dependent deficit of the neutrino oscillation phenomenon appears.

This measurement, $\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029}$, is the most precise measurement of θ_{13} made by Double Chooz so far (the uncertainty has been reduced 22% with respect to the previous publication), and is in excellent agreement with the one obtained using the RRM analysis over the Gd and H data sets ($\sin^2(2\theta_{13}) = 0.090 \pm 0.033$, from section 6.1.3).

Table 6.4 shows how the uncertainties on the input parameters are constrained by the fit. All input parameters agree with their outputs within less than 1σ, except the cosmogenic background rate. This motivated to repeat the Rate+Shape analysis removing the constraints on the fast neutron and stopping muon, and ^{9}Li and ^{8}He background scaling factors, $\alpha_{\mathrm{Fn+S}\mu}$ and $\alpha_{\mathrm{Li+He}}$, respectively; so they are treated as free parameters in the fit (the accidental background is not released since it is measured very precisely with the off-time method in section 4.1.3.3). The best fit is found at $\sin^2(2\theta_{13}) = 0.088^{+0.030}_{-0.031}$, ^{9}Li and ^{8}He rate of $0.49^{+0.16}_{-0.14}$ events/day, and fast neutron and stopping muon rate of $0.541^{+0.052}_{-0.048}$ events/day, where the confidence intervals are defined by $\chi^2 - \chi^2_{\mathrm{min}} < 1.0$, with $\chi^2_{\mathrm{min}}/\mathrm{d.o.f.} = 46.9/38$. Whereas the fast neutron and stopping muon rate barely changes with respect to the constrained version, the ^{9}Li and ^{8}He rate is decreased further. In any case, the value of $\sin^2(2\theta_{13})$ is in excellent agreement with the background-constrained one and the precision is

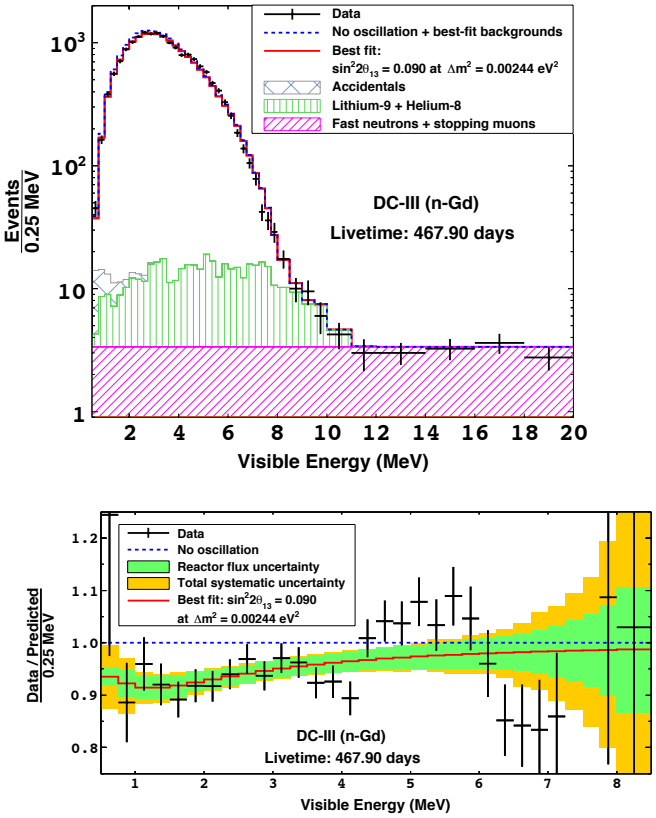


Figure 6.9: Top: Observed IBD candidate prompt trigger visible energy spectrum (black points with statistical error bars), superimposed on the $\overline{\nu}_e$ no-oscillation prediction (blue dashed line) and on the best fit (red solid line), with the stacked best-fit backgrounds added: fast neutrons and stopping muons (magenta, slant-hatched), ^{9}Li and ^{8}He (green, vertical-hatched), and accidental (gray, cross-hatched). Bottom: Ratio of the background-subtracted DATA (black points with statistical error bars) and the best fit (red solid line) to the no-oscillation MC prediction. The gold band represents the total systematic uncertainty on the best-fit prediction, with the reactor flux systematic uncertainty contribution colored in green. From [199].

similar, confirming the background constraint granted by the extended prompt energy window devoid of $\overline{\nu}_e$ signal.

A direct comparison between the Reactor Rate Modulation and the Rate+Shape analyses is made by rebinning the Rate+Shape prompt energy spectrum into a single bin, so it becomes a rate-only analysis. The best fit gives $\sin^2(2\theta_{13}) = 0.090^{+0.036}_{-0.037}$, in excellent agreement with the RRM $\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035}$ from section 6.1.1.

On the other hand, if inverse mass hierarchy is assumed, then $\left|\Delta m^2_{\text{MINOS}}\right| = (2.38^{+0.09}_{-0.10}) \cdot 10^{-3}\,\text{eV}^2$ from [59]. In this case, the best fit is found at $\sin^2(2\theta_{13}) = 0.092^{+0.033}_{-0.029}$, with $\chi^2_{\text{min}}/\text{d.o.f.} = 52.2/40$. Given the minuscule difference in the value of $\sin^2(2\theta_{13})$ between the two mass hierarchies, the normal hierarchy result can be adopted as the standard result.

Finally, figure 6.9 shows also an unexpected distortion in the prompt visible energy spectrum for energies larger than 4 MeV. The discussion of this distortion is deferred to section 6.2.3 in order to examine first the results of the H-based $\overline{\nu}_e$ selection.

6.2.2 H results

In the H-based $\overline{\nu}_e$ selection, the χ^2 of equation 6.12 requires a modification as a result of the different shape of the fast neutron and stopping muon background. Unlike the Gd sample, in which the fast neutron and stopping muon background is best modeled by a flat spectrum, so the rate and its uncertainty are sufficient to describe it fully; in the H case, the spectral shape is best modeled by the function of equation 4.20, which depends on three parameters, p_0, p_1 and p_2. In order to account for the uncertainty on them, they are added to the χ^2 as nuisance parameters constrained by the following term

$$\begin{aligned}&(p_0 - p_{0,\text{CV}}, p_1 - p_{1,\text{CV}}, p_2 - p_{2,\text{CV}}) \cdot \\ &\cdot \begin{pmatrix} \sigma^2_{p_0} & \rho_{p_0p_1}\sigma_{p_0}\sigma_{p_1} & \rho_{p_0p_2}\sigma_{p_0}\sigma_{p_2} \\ \rho_{p_1p_0}\sigma_{p_1}\sigma_{p_0} & \sigma^2_{p_1} & \rho_{p_1p_2}\sigma_{p_1}\sigma_{p_2} \\ \rho_{p_2p_0}\sigma_{p_2}\sigma_{p_0} & \rho_{p_2p_1}\sigma_{p_2}\sigma_{p_1} & \sigma^2_{p_2} \end{pmatrix}^{-1} \begin{pmatrix} p_0 - p_{0,\text{CV}} \\ p_1 - p_{1,\text{CV}} \\ p_2 - p_{2,\text{CV}} \end{pmatrix}\end{aligned} \tag{6.16}$$

where the central values $p_{n,\text{CV}}$ are given after equation 4.20, and the 3×3 covariance matrix accounts for the correlations, with the correlation coefficients being $\rho_{p_0p_1} = 0.17$, $\rho_{p_1p_2} = 0.95$, $\rho_{p_0p_2} = -0.09$.

A scan over possible values of $\sin^2(2\theta_{13})$ is carried out, in which the new χ^2 is minimized with respect to the eleven nuisance parameters. The best fit is found at $\sin^2(2\theta_{13}) = 0.124^{+0.030}_{-0.039}$, with $\chi^2_{\text{min}}/\text{d.o.f.} = 69.4/38$. The best-fit values of the nuisance parameters are displayed in table 6.5. This result is consistent with the other $\sin^2(2\theta_{13})$ measurements obtained so far. Nevertheless, a large value of χ^2_{min} shows up, which is mainly attributed to the $4.25 - 5.75\,\text{MeV}$ region (see figure 6.10), where a similar distortion as the Gd one is exhibited. The investigation of the spectral distortion is reported in the next section.

6.2.3 Spectral distortion

To sum up the situation, a distortion in the prompt trigger visible energy spectrum of the Gd and H samples is observed. From the ratio of the observed data to the prediction, it appears to be an excess in the region $\sim 4 - 6\,\text{MeV}$

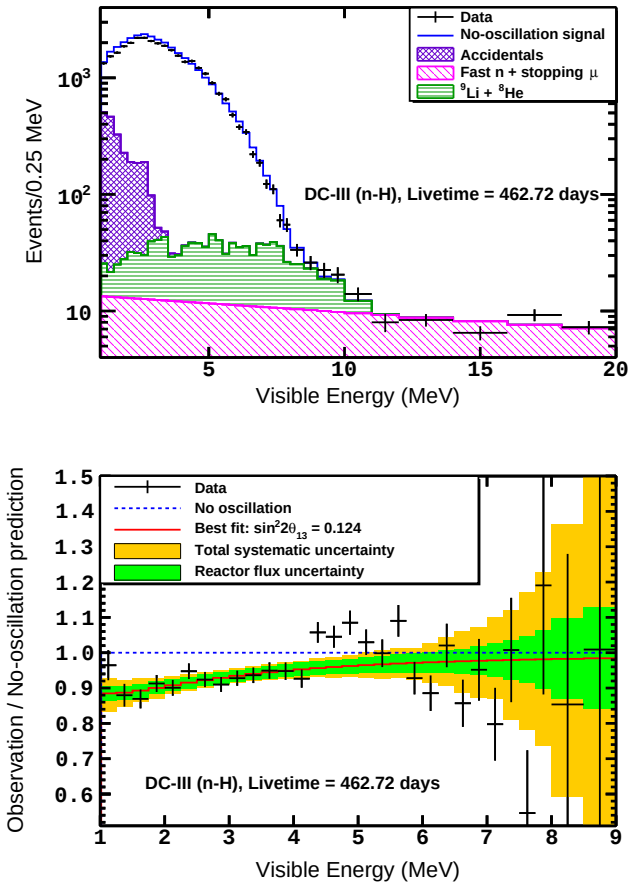


Figure 6.10: Top: Observed IBD candidate prompt trigger visible energy spectrum (black points with statistical error bars), superimposed on the $\overline{\nu}_e$ no-oscillation prediction (blue line) with the stacked best-fit backgrounds added: fast neutrons and stopping muons (magenta, slant-hatched), ^{9}Li and ^{8}He (green, horizontal-hatched), and accidental (purple, cross-hatched). Bottom: Ratio of the background-subtracted DATA (black points with statistical error bars) and the best fit (red solid line) to the no-oscillation MC prediction. The gold band represents the total systematic uncertainty on the best-fit prediction, with the reactor flux systematic uncertainty contribution colored in green. From [200].

Fit parameter	Input value	Best-fit value
Δm^2_{31} ($10^{-3}\,\mathrm{eV}^2$)	$2.44^{+0.09}_{-0.10}$	$2.44^{+0.09}_{-0.10}$
Residual $\overline{\nu}_e$	2.73 ± 0.82	2.81 ± 0.82
Fast-n + stopping-μ rate (events/day)	1.55 ± 0.15	1.62 ± 0.10
Li+He rate (events/day)	$0.95^{+0.57}_{-0.33}$	$1.60^{+0.21}_{-0.24}$
Accidental rate (events/day)	$4.334^{\pm 0.007(\mathrm{stat})}_{\pm 0.008(\mathrm{syst})}$	$4.334^{\pm 0.007(\mathrm{stat})}_{\pm 0.008(\mathrm{syst})}$
Energy-model a' (MeV)	0.000 ± 0.067	$-0.008^{+0.028}_{-0.020}$
Energy-model b'	1.004 ± 0.022	$0.997^{+0.007}_{-0.009}$
Energy-model c' (MeV^{-1})	-0.0001 ± 0.0006	$-0.0006^{+0.0006}_{-0.0005}$
Fast-n + stopping-μ p_0 (events/MeV)	12.52 ± 1.36	12.33 ± 1.34
Fast-n + stopping-μ p_1 (MeV^{-1})	0.042 ± 0.015	$0.037^{+0.015}_{-0.013}$
Fast-n + stopping-μ p_2 (events/MeV)	0.79 ± 1.39	$0.39^{+1.48}_{-1.30}$

Table 6.5: Input values and best-fit values, and their corresponding uncertainties, of the nuisance parameters of the Rate+Shape H analysis. The residual $\overline{\nu}_e$ and background scaling factors of equation 6.12 have been multiplied to their respective expected number and nominal background rates. Notice the uncertainty reduction resulting from the fit.

with respect to the IBD candidate prediction. This suggests that either an unexpected background or an unaccounted $\overline{\nu}_e$ contribution is entering both Gd-based and H-based IBD selections.

To test the nature of the distortion, a simple sideband analysis on the Gd sample is performed, in which the prompt energy spectrum in the vicinity of the distortion is modelized using a second order polynomial fitted to the observed data below 4.25 MeV and above 6 MeV, as depicted in figure 6.11. The magnitude of the excess is defined as the integral of the observed spectrum over the $4.25 - 6$ MeV interval, after subtraction of the interpolation given by the second order polynomial. When the size of the excess is studied as a function of the number of reactors running, a clear correlation is visible (see inset of figure 6.11), which points to a reactor $\overline{\nu}_e$-related origin. The correlation becomes more evident when the H sample is added [1].

To confirm further the reactor-flux nature of the distortion, the Gd data are divided into five subsamples according to their prompt energy, in which one of them corresponds to the $4 - 6$ MeV interval, and the Reactor Rate Modulation analysis is applied to each of them. Since the Reactor Rate Modulation studies the observed rate as a function of the total reactor thermal power, it is capable of discerning whether the excess correlates with reactor power or stays constant, as it would be characteristic of a background. In order to perform this analysis, the RRM is modified by removing the constraints on the reactor-on flux prediction (α_r) and the background rate (B) parameters of equation 6.7, so they are left free in the fit; and $\sin^2(2\theta_{13})$ is treated as a nuisance parameter constrained to $0.090^{+0.009}_{-0.008}$ from the measurement of the Daya Bay experiment [201] (Daya Bay is already a multi-detector experiment with near detectors monitoring the reactors, so it does not rely on a reactor-flux prediction). The best-fit background rate and shift in the relative normalization of the reactor flux for each prompt

[1] The H sample in figure 6.11 corresponds to the same data set used in this thesis, but the IBD selection was performed according to the previous H selection [72].

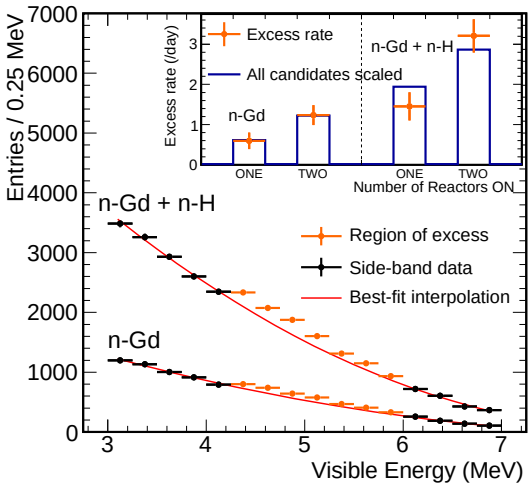


Figure 6.11: Prompt trigger visible energy spectrum in the vicinity of the distortion for the Gd sample (labeled n-Gd), and the addition of the H-based sample (labeled n-H). Points show the sidebands (black) and the excess region (orange). The best-fit polynomial to the sidebands is shown as a red line. Inset: Excess rate (orange points, after subtraction of the interpolation) as a function of the number of reactors running. For reference, the full integral over the IBD spectrum is shown in blue, normalized to the same area as the excess histogram. The H sample has been selected according to a previous H selection [72], which is accidental-background enriched. From [2].

energy subsample are shown in figure 6.12. Whereas the best-fit background rates agree with the background estimation and with the reactors-off data, thus disfavoring an unknown additional background; the shifts in the reactor flux exceed the prediction by 2.0σ in the $4.25-6$ MeV interval and are lower by 1.5σ in the $6-8$ MeV interval. Therefore, it is concluded that the distortion is most likely to be caused by the reactor flux. If the analysis is repeated constraining the background rate to the background estimation, which has been proved to be reliable, the significances of the deviations raise to 3.0σ for the $4.25-6$ MeV interval and to 1.6σ for the $6-8$ MeV one.

The spectral distortion is unlikely to be caused by the detector response. As figure 3.10 showed, a very good matching of the energy scale and resolution between DATA and MC are achieved over the whole IBD spectrum range. Moreover, there is specifically a calibration point in the region of the distortion given by the neutron captures on carbon, which results in an energy peak at 4.95 MeV [38]. The good agreement between DATA and MC at this point is within 0.5%. An additional handle is provided by the β decay spectrum of the cosmogenic ^{12}B, which spans 13.4 MeV. No distortion of the ^{12}B DATA with respect to a MC simulation is observed across the spectrum.

In addition, the other two reactor antineutrino oscillation experiments, RENO [202] and Daya Bay [203] have reported a similar excess in both near and far detectors.

The investigations to find the cause of the distortion are still ongoing. Several ideas have been put forward after the announcement of the distortion. In reference [204], Dwyer *et al.* build a reactor-flux prediction based on an *ab initio* method [2] using the `ENDF/B-VII.1` libraries [205] which features an excess at $4-6$ MeV with respect to the prediction from Huber [98] (used by the three experiments, Double Chooz, Daya Bay and RENO, for the flux due to ^{235}U, ^{239}Pu and ^{241}Pu), which derives the reactor flux from the cumulated electron spectra resulting from exposition of nuclear fuel films to the thermal neutron flux of the research reactor at the Institut Laue-Langevin (ILL) [139, 140, 141] (see section 2.4.1 for more details). However, Dwyer *et al.* do not assign any uncertainty or correction due to the incompleteness of the database, the bias in the branching fractions (which tend to favor high-energy β decays, also known as the Pandemonium effect [206]), or the spectral shape of the decays (which are all assumed to be allowed Gamow-Teller decays, despite the fact that $\sim 25\%$ are known to be of forbidden type).

In reference [207], Hayes *et al.* do correct for the bias in the branching ratios, and still find an excess in their *ab initio* prediction using the `ENDF/B-VII.1` libraries [205] with respect to the Huber [98] and Mueller *et al.* [98, 114] model (Mueller *et al.* is used for ^{238}U by the Daya Bay and RENO experiments). However, that excess is not seen when other libraries (`JEFF-3.1.1` [208]) are used instead. For Double Chooz, which uses the Haag *et al.* model for ^{238}U [142], both `ENDF/B-VII.1` and `JEFF-3.1.1` libraries predict an excess as the one observed, for the latter arising almost completely from ^{238}U.

Concerning the origin of the distortion, Hayes *et al.* [207] discuss some possibilities which are quoted next. Non-fission neutron-induced sources of $\overline{\nu}_e$ within the reactor cores are not capable of reproducing the excess. The shape cor-

[2] Summation of the individual $\overline{\nu}_e$ spectra produced by the numerous branches of β^- decays of the nuclear fuel fission products, which are obtained from the corresponding electron spectroscopy experimental data.

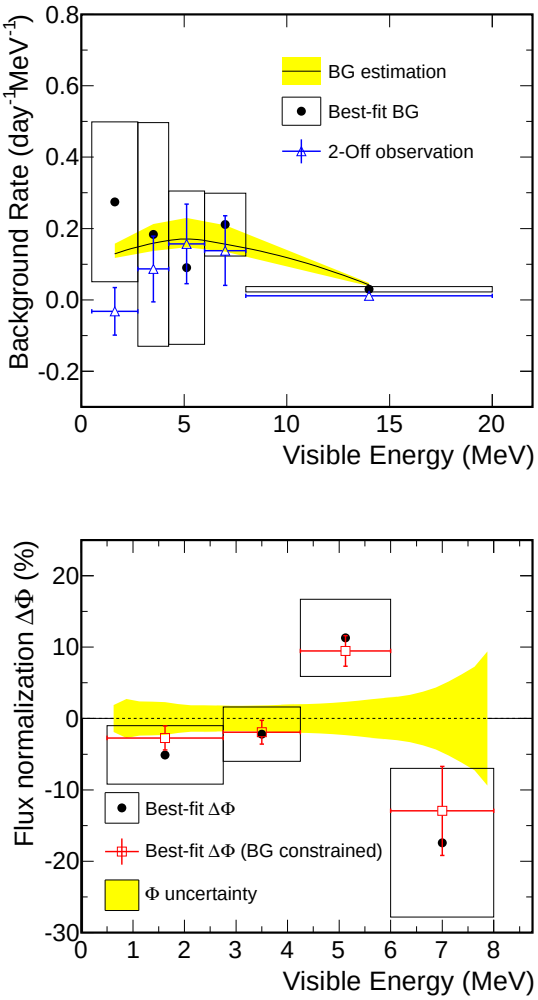


Figure 6.12: Modified Reactor Rate Modulation analysis (see text for the details) applied to subsamples of the Gd-based $\overline{\nu}_e$ selection according to their prompt energy. Top: Best-fit background rates (black circles with error boxes) for each subsample. For comparison purposes, the background rate estimation is shown as a line and its uncertainty as a yellow band, and the reactors-off observed data are represented by blue empty triangles with statistical error bars. Bottom: Best-fit shifts of the reactor-flux normalization with respect to the prediction, with the background rate left free (black circles with error boxes) and with the background constrained to the estimation (red empty squares with error bars). The uncertainty on the reactor-flux prediction is shown as a yellow band. The analysis is restricted to the $\overline{\nu}_e$-dominated region (below 8 MeV). From [2].

rections from a proper treatment of the forbidden decays cannot explain the magnitude of the excess either. The role of ^{238}U is examined, since its contribution in the $4-6$ MeV region is twice that of ^{235}U. Both `ENDF/B-VII.1` and `JEFF-3.1.1` libraries predict a shoulder in the ^{238}U spectrum with respect to Mueller *et al.* [114], but to reproduce the size exclusively with ^{238}U would require fission yields a factor four and two larger than those of the libraries, respectively. For Haag *et al.* [142], the shoulder from ^{238}U in `JEFF-3.1.1` may explain the excess, but more data is needed before such conclusion can be drawn. Another possibility is that pressurized water reactors (PWR), like those in the Double Chooz, Daya Bay and RENO experiments, are known to have a harder neutron spectrum, with a more sizable epithermal component, than the research reactor at the ILL where the reference cumulated electron spectra for ^{235}U, ^{239}Pu and ^{241}Pu used by the three experiments were measured. This difference in the neutron energies may alter the fission yields, enhancing certain nuclei and building the excess. Finally, a possible error in the original ILL measurements cannot be excluded.

Regarding the impact of the spectral distortion to the Rate+Shape measurements of θ_{13}, it must be noted first that almost all the oscillation signature is concentrated at prompt energies below 4 MeV, where the observed data exhibits the characteristic energy-dependent deficit (cf. the ratio plots in figures 6.9 and 6.10), so an unaccounted structure above 4 MeV has little effect on the determination of θ_{13}. This was tested explicitly by including an artificial excess in the predicted IBD candidate spectrum in the region of the distortion, whose normalization was left free in the fit. After trying multiple peak energies and widths, none resulted in a deviation of θ_{13} larger than 30% of its uncertainty interval, demonstrating that the distortion does not affect significantly the θ_{13} measurement. Another proof of the robustness of the Double Chooz measurements is the agreement found between the Rate+Shape results and the Reactor Rate Modulation ones, the latter not being sensitive to the distortion.

However, the appearance of a distortion in the observed spectrum not anticipated by the reactor-flux model casts doubts on the accuracy of the prediction and the quoted uncertainty. Consequently, there is nothing to be gained from a simultaneous Rate+Shape analysis of the Gd-based and H-based $\overline{\nu}_e$ samples to decrease the uncertainty on θ_{13}. Notice that this conclusion does not affect so severely the Reactor Rate Modulation analysis, which does not use the spectrum shape but the total integral, which is anchored to a short-baseline measurement by the Bugey4 experiment (see equation 2.11). In any case, this reinforces the necessity of having a near detector close to the reactor cores, in order to reduce the dependence on a reactor model and to improve the precision of the measurement of θ_{13}. This is precisely the subject of the next section.

6.3 Two-detector outlook

To conclude the chapter, the expected precision on θ_{13} resulting from the analysis of the data from the Near and Far Detectors is discussed. The Near Detector became operative during the writing of this thesis, and no oscillation analysis has been performed yet using the Near Detector DATA. Therefore, the following study utilizes only the $\overline{\nu}_e$ MC simulation developed for the Far Detector as a proxy for the observed and expected $\overline{\nu}_e$ interactions in both detectors.

An oscillation corresponding to $\sin^2(2\theta_{13}) = 0.10$ is artificially induced in the observed sample.

The oscillation analysis chosen is the Rate+Shape analysis, since it has provided the most precise θ_{13} measurements so far. The χ^2 statistic used is similar to that of equation 6.12, although it has been extended to accommodate two detectors:

$$\begin{aligned}\chi^2 =& \sum_i \sum_j \left(N_i^{\text{pred}} - N_i^{\text{obs}}\right)\ M_{ij}^{-1}\ \left(N_j^{\text{pred}} - N_j^{\text{obs}}\right) + \\ &+ \left(\frac{\Delta m^2_{31} - \Delta m^2_{\text{MINOS}}}{\sigma_{\text{MINOS}}}\right)^2 + \\ &+ \sum_{D=\text{ND, FD}} \left[\left(\frac{\alpha^D_{\text{Fn+S}\mu} - 1}{\sigma^D_{\text{Fn+S}\mu}}\right)^2 + \left(\frac{\alpha^D_{\text{Li+He}} - 1}{\sigma^D_{\text{Li+He}}}\right)^2 + \left(\frac{\alpha^D_{\text{acc}} - 1}{\sigma^D_{\text{acc(syst)}}}\right)^2\right]\end{aligned} \tag{6.17}$$

where the sum in i and j is over the bins in which the prompt spectra of the two detectors have been divided (following the same convention as in table 6.2). The Far and Near Detector prompt spectra can be regarded as two vectors, $\vec{N}^{\text{FD}}$ and $\vec{N}^{\text{ND}}$, respectively; which are concatenated to form a single vector, so

$$\begin{aligned}&\sum_i \sum_j \left(N_i^{\text{pred}} - N_i^{\text{obs}}\right)\ M_{ij}^{-1}\ \left(N_j^{\text{pred}} - N_j^{\text{obs}}\right) = \\ &= \left(\vec{N}^{\text{FD,pred}} - \vec{N}^{\text{FD,obs}}, \vec{N}^{\text{ND,pred}} - \vec{N}^{\text{ND,obs}}\right) \cdot \\ &\cdot \begin{pmatrix} M^{\text{FD}} & M^{\text{corr}} \\ M^{\text{corr}} & M^{\text{ND}} \end{pmatrix}^{-1} \begin{pmatrix} \vec{N}^{\text{FD,pred}} - \vec{N}^{\text{FD,obs}} \\ \vec{N}^{\text{ND,pred}} - \vec{N}^{\text{ND,obs}} \end{pmatrix}\end{aligned} \tag{6.18}$$

where the covariance matrix, M, which accounts for the uncertainties described later, has been divided into four blocks. The diagonal blocks account for the uncertainties affecting a single detector, M^{FD} for the Far Detector and M^{ND} for the Near Detector. The off-diagonal blocks account for the correlated uncertainties between detectors. The internal structure of each block depends on the correlations between the energy bins which are relating.

The observed (N_i^{obs}) and predicted (N_i^{pred}) number of IBD candidates in the i-th bin are obtained using the Gd selection (section 4.1). In this case, the predicted number of IBD candidates in the i-th bin of the D-th detector is expressed as

$$\begin{aligned}N_i^{D,\text{pred}} =& \sum_{R=1,2} P^{\text{surv,D}}_{i,R}\left(\sin^2(2\theta_{13}), \Delta m^2_{31}\right) N^{\nu,D}_{i,R}\left(b'\right) + \\ &+\ \alpha^D_{\text{Fn+S}\mu} N_i^{\text{Fn+S}\mu} + \alpha^D_{\text{Li+He}} N_i^{\text{Li+He}} + \alpha^D_{\text{acc}} N_i^{\text{acc}},\end{aligned} \tag{6.19}$$

where $P^{\text{surv,D}}_{i,R}\left(\sin^2(2\theta_{13}), \Delta m^2_{31}\right)$ is the survival probability of the $\overline{\nu}_e$ emitted from the reactor R which interacted in the detector D and ended up in the i-th energy bin, which is a function of $\sin^2(2\theta_{13})$ and Δm^2_{31} (as in equation 1.54). The treatment of Δm^2_{31} is identical to the one-detector fit, and normal mass hierarchy is assumed, so $\Delta m^2_{\text{MINOS}} = (2.44^{+0.09}_{-0.10}) \cdot 10^{-3}\,\text{eV}^2$ from reference [59]. $N^{\nu,D}_{i,R}$ is the expected number of $\overline{\nu}_e$ IBD interactions in the detector D coming from the reactor R in the null-oscillation hypothesis. Since a MC prediction has

only been generated for the Far Detector, in order to emulate a Near Detector the Far Detector sample is scaled by $(L_R^{\rm FD})^2/(L_R^{\rm ND})^2$ to correct for the different baselines between the Far ($L_R^{\rm FD}$) and Near ($L_R^{\rm ND}$) detectors to the reactor core R (the difference in the baselines is also taken into account to compute the survival probability). $N_{i,R}^{\nu}$ is also a function of the energy scale parameter b' which is described later. In order to compute $N_{i,R}^{\nu;D}$, the reactor-flux predicted rate (equation 2.12) must be integrated over the run-time of the detector. To obtain the future precision on θ_{13}, the reactors are assumed to continue operating as they have done up to this point in terms of thermal power and nuclear fuel composition, and the detectors are assumed to keep the same duty cycle as the Far Detector so far.

For the Far Detector, the background rates and spectra are those estimated for the Gd-based selection, so the second line of equation 6.19 is equal to that of equation 6.13. For the Near Detector, the same spectral shapes as in the Far Detector are assumed (although some differences are expected in the fast neutron and stopping muon, and accidental backgrounds because of the smaller overburden, the different shielding [3] and detector surroundings). The background rates in the Near Detector are scaled up as a result of the increased muon flux due to the shallower location, which has been measured to be 6.7 times higher using an Outer Veto prototype [209, 210]. To scale the fast neutron and stopping muon background rate, a distinction is made between the two sources, which account for $\sim 70\%$ and $\sim 30\%$ of the inclusive rate, respectively (estimation based on the prompt-delayed trigger time interval distribution, where the fast neutron distribution decreases with a $\sim 30\,\mu$s time constant characteristic of the neutron capture, and the stopping muon one decreases with that of the the muon lifetime, $2.2\,\mu$s). The fast neutron and cosmogenic background rates are scaled following the power-law prescription of reference [170], which results in a 4.7 times higher fast neutron rate and a 4.5 times higher cosmogenic background rate; while the stopping muons are scaled directly with the muon flux. The accidental background comes from the random coincidence of two triggers, the prompt one being dominated by the natural radioactivity, and the delayed one by cosmogenic ^{12}B decays and spallation neutrons. Hence, it must be proportional to the muon flux somehow, but not as straightforward as the other muon-induced backgrounds. A factor 3 is considered as an approximation. Regarding the relative uncertainties on the background rates, they are assumed to be the same as in the Far Detector. Since part of the uncertainty is statistical, this a conservative approach; the higher background rates in the Near Detector should improve the statistical uncertainty. The background rates in each detector are summarized in table 6.6.

As in the one-detector case, the B background rate contribution in the D-th detector is tuned during the fit by a scaling factor, α_B^D, which is included in the χ^2 of equation 6.17 as a nuisance parameter, centered at 1 and constrained by its relative uncertainty σ_B^D. An independent nuisance parameter is used for each background in each detector.

The reactors-off data are not considered, since tests revealed negligible gain because of the foreseeable reactors-off periods.

The total covariance matrix of equation 6.17, M, consists of several covari-

[3] The Near Detector is shielded by 1 m of water instead of 15 cm of steel, except on top of the detector where steel is still used.

Background source	FD rate (events/day)	ND rate (events/day)
Fast-n and stopping-μ	0.60 ± 0.05	3.18 ± 0.27
^{9}Li and ^{8}He	0.97 ± 0.29	4.37 ± 1.29
Accidental	0.070 ± 0.005	0.21 ± 0.02

Table 6.6: Summary of the background rates in the Gd-based $\overline{\nu}_e$ selection estimated in the Far Detector (FD) and the projected ones for the Near Detector (ND). The ^{9}Li and ^{8}He uncertainty has been symmetrized for easier implementation.

ance matrices, each one accounting for an uncertainty source:

$$M = M^{\mathrm{stat}} + M^{\mathrm{r}} + M^{\mathrm{d}} + M^{\mathrm{Li+He(shape)}} + M^{\mathrm{acc(stat)}} + M^{\mathrm{E-scale}} \qquad (6.20)$$

where M^{stat} accounts for the statistical uncertainty on the prediction, so it has the same diagonal form as in the one-detector case.

As it was discussed in section 2.1, from the privileged configuration of the Double Chooz experiment, only 0.1% of the 1.7% reactor-flux systematic uncertainty is estimated to be uncorrelated between the Near and Far detectors [122], which implies a correlation coefficient of 0.997 between detectors. Therefore, to build the reactor-flux covariance matrix for the two-detector analysis, M^{r}, the Far detector covariance matrix is used in each of the four blocks which compose the two-detector matrix (as shown in equation 6.18) as an approximation, with the off-diagonal blocks multiplied by the correlation coefficient. Notice that, even though the χ^2 of equation 6.17 utilizes a prediction based on an $\overline{\nu}_e$ MC simulation, as a result of this high degree of correlation between Near and Far detectors, the θ_{13} measurement becomes almost insensitive to the reactor-flux model.

According to table 5.3, the relative uncertainty on the $\overline{\nu}_e$ normalization because of the IBD detection is 0.63%. Leaving aside the contributions due to the muon veto, the light noise cuts, the trigger efficiency, the isolation cut and the background vetoes, which together account for less than 0.13% of the detection relative uncertainty; the main contributions to the remaining 0.62% uncertainty, which is assumed to apply to both detectors, are: the number of protons in the detector (0.30%), the Gd-fraction (0.43%), the neutron selection efficiency (0.19%) and the spill-in/out (0.27%).

Concerning the number of protons and the Gd-fraction, both Near and Far Inner Detectors (where the $\overline{\nu}_e$ are detected) have been built identical, and filled with the Gd-doped liquid scintillator produced in a single batch for the two. Consequently, both uncertainties are fundamentally correlated between detectors.

The uncertainty on the neutron selection efficiency arises from a combination of the IBD and ^{252}Cf measurements (sections 5.4.2.1 and 5.4.2.2, respectively), but it is driven by the former, which is more precise. In the IBD measurement, the relative statistical uncertainty is 0.10%, which is inherently uncorrelated between detectors; and the systematic uncertainty is 0.19%, which comes from the possible contamination of stopping muons in the estimation, and therefore it can be taken as uncorrelated between detectors too.

Uncertainty	Correlation coefficient
Background rates	0.0
Statistics	0.0
Reactor flux	0.997
Detection efficiency	0.88
^{9}Li and ^{8}He shape	1.0
Energy scale	0.0

Table 6.7: Summary of the estimated correlations between the Near and Far detectors uncertainties. A zero correlation coefficient means that the uncertainty is uncorrelated between detectors; a unity correlation coefficient implies that the uncertainties are fully correlated between detectors.

As a first approximation, the uncertainty due to the spill-in and spill-out currents can be considered fully correlated between detectors (minor differences between detectors, such as the solid angle subtended by the reactor cores from each detector, should not change this conclusion significantly).

Summing up, only 0.2% of the 0.62% relative detection uncertainty is uncorrelated between the Near and Far detectors, resulting in a 0.88 correlation coefficient. Hence, the detection covariance matrix M^{d} is constructed in the manner of equation 6.18 as four blocks with the form of the one-detector detection covariance matrix, in which the off-diagonal blocks are multiplied by the correlation coefficient. Although it has no effect on the uncertainty at all, because of the higher muon flux at the Near Detector, the dead-time induced by the muon veto and OV veto increases from 4.4% at the Far Detector to 29% at the Near Detector, so the $\overline{\nu}_e$ prediction in the Near Detector must be scaled down to reflect this.

$M^{\mathrm{Li+He(shape)}}$ encodes uncertainty on the spectral shape of the cosmogenic ^{9}Li and ^{8}He. Since it affects both detectors equally, it is taken as fully correlated across them.

As in the one-detector case, the systematic uncertainty on the accidental background rate has been taken into account using nuisance parameters, so the remaining statistical uncertainty is accounted for using the covariance matrix $M^{\mathrm{acc(stat)}}$. Since the spectra in each detector will be measured directly using the off-time method (section 4.1.3.3), it is uncorrelated between detectors, and between energy bins too, resulting in a diagonal matrix.

The elaborated non-linear implementation of the energy scale uncertainty using three nuisance parameters in equation 6.12, which was based on equation 3.22a; has been replaced by a covariance matrix $M^{\mathrm{E-scale}}$ arising from a simpler linear model,

$$E^{\mathrm{MC}}_{\mathrm{vis,fit}} = b' \cdot E^{\mathrm{MC}}_{\mathrm{vis}}, \tag{6.21}$$

where the uncertainty on b' is the equivalent uncertainty of the non-linear model weighted by the IBD prompt spectrum, 0.77%. The energy scale uncertainties on both detectors can be regarded as uncorrelated, which results in a block-diagonal covariance matrix.

The uncertainty correlations between detectors are summarized in table 6.7.

Finally, figure 6.13 shows the expected precision on $\sin^2(2\theta_{13})$ as a function of the elapsed time since the Far Detector started the data taking (April 2011).

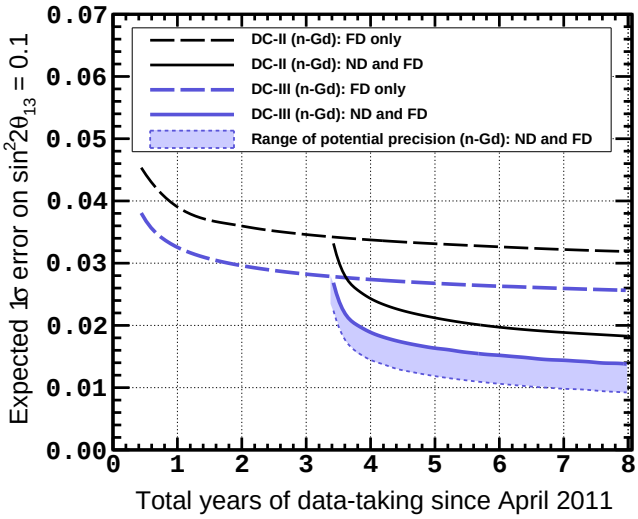


Figure 6.13: Expected precision of the Double Chooz Rate+Shape measurement of $\sin^2(2\theta_{13})$ considering only IBD neutrons captured on Gd, as a function of the elapsed time since the Far Detector (FD) started the data taking (April 2011), assuming $\sin^2(2\theta_{13}) = 0.10$ (see text for further assumptions). The projections using only the Far Detector (blue dashed line) and both Near Detector (ND) and Far Detector (blue solid line) are shown. The shaded region represents the potential improvement brought by the reduction of systematic uncertainties, delimited from below by the case in which all systematic uncertainties but the reactor-flux one are negligible. For comparison purposes, black curves representing the expected precision from the previous Gd-based $\overline{\nu}_e$ selection [1] are included. From [2].

In order to quantify the precision achieved, one half of the 68.3% C.L. interval on the best-fit $\sin^2(2\theta_{13})$ is used. For comparison purposes, it includes the expectation from the previous Gd-based $\overline{\nu}_e$ analysis [1] so as to reveal the improvement due to the new analysis [2] described in this thesis. With the assumptions listed in this section, a 15% precision on $\sin^2(2\theta_{13})$ will be achieved in three years of data taking with both detectors. However, there is room for improvement (e.g. the uncertainty on the background rates has a statistical component which will decrease as more data comes in), and a 10% precision is within reach. From figure 6.13, it is concluded that the Double Chooz measurement of θ_{13} will be still dominated by statistics even after three years running with two detectors. Therefore, the data from the H-based $\overline{\nu}_e$ selection, which has not been considered in figure 6.13, plays a crucial role, since it has the potential to shorten the timescale to achieve the precision goals.

Chapter 7

Conclusions

The measurement of the $\overline{\nu}_e$ flux at the Double Chooz Far Detector has provided evidence of $\overline{\nu}_e$ disappearance as a result of the oscillation into other flavors driven by the neutrino mixing angle θ_{13}. Two independent $\overline{\nu}_e$ samples, selected according to which nucleus (Gd or H) captures the neutrons produced by the $\overline{\nu}_e + p \rightarrow e^+ + n$ reaction, have been collected in 489.51 days of run-time. Two oscillation analyses, one studying the disappearance as a function of the expected $\overline{\nu}_e$ rate (RRM) and another studying the disappearance as a function of the $\overline{\nu}_e$ energy (R+S), have measured $\sin^2(2\theta_{13})$ to be:

$$\begin{aligned}
&\sin^2(2\theta_{13}) = 0.090^{+0.034}_{-0.035} && \text{RRM (Gd)} \\
&\sin^2(2\theta_{13}) = 0.095^{+0.038}_{-0.039} && \text{RRM (H)} \\
&\sin^2(2\theta_{13}) = 0.088^{+0.033}_{-0.033} && \text{RRM (Gd+H)} \\
&\sin^2(2\theta_{13}) = 0.090^{+0.032}_{-0.029} && \text{R+S (Gd)} \\
&\sin^2(2\theta_{13}) = 0.124^{+0.030}_{-0.039} && \text{R+S (H)}
\end{aligned}$$

The agreement exhibited by the several measurements (see figure 7.1) tells about the robustness of the measurement. Furthermore, tests freeing the contribution of the background processes mimicking the $\overline{\nu}_e$ signature yielded consistent results. Even though these measurements are dominated by the systematic uncertainty on the reactor prediction, and therefore cannot compete in precision with the ones from the multi-detector Daya Bay experiment, $\sin^2(2\theta_{13}) = 0.084 \pm 0.005$ [211], or the RENO experiment, $\sin^2(2\theta_{13}) = 0.101 \pm 0.013$ [202]; they demonstrate the level of precision which can be attained with a single detector. Since the Double Chooz Near Detector has already started the data taking, these measurements will remain as the last ones made by Double Chooz in its one-detector phase. In the next years, profiting from the cancellation of correlated uncertainties and the improved $\overline{\nu}_e$ selection described in this thesis, Double Chooz is expected to reach a $10-15\%$ precision on $\sin^2(2\theta_{13})$, providing a competitive measurement of θ_{13}.

Despite having only the Far Detector data, Double Chooz has been able to reveal a spectrum distortion with respect to the reactor-flux prediction which has been confirmed by the RENO and Daya Bay experiments, which have data from the detectors close by the reactor cores too. Although this distortion has been shown not to affect the Double Chooz θ_{13} measurement in a significant

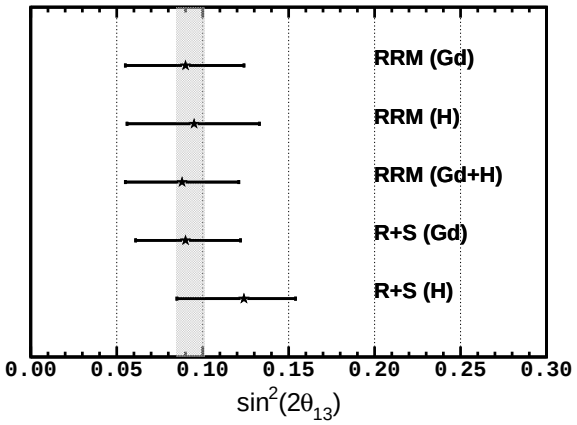


Figure 7.1: Summary of the most precise measurements of $\sin^2(2\theta_{13})$ by the Double Chooz experiment using the Reactor Rate Modulation (RRM) and the Rate+Shape (R+S) analyses of the Gd-based and H-based $\overline{\nu}_e$ selections. For reference, the current PDG value $\sin^2(2\theta_{13}) = 0.093 \pm 0.008$ [10] is shown.

way, and by no means affects the ultimate θ_{13} measurement because the multi-detector configuration of the experiments alleviates the dependence on a reactor-flux model; it calls for a revision of the significance of the reactor antineutrino anomaly [115] which pointed to the existence of sterile antineutrinos, and the sensitivity of future single-detector experiments such as JUNO [106], which aims to measure the neutrino mass hierarchy based on the fine structure of the $\overline{\nu}_e$ spectrum.

The analyses which have enabled to achieve such precision in the Double Chooz one-detector phase have been discussed extensively in the thesis, specially the parts which the author contributed to their development.

In particular, the continuous improvements on the energy scale uncertainty have made possible a reduction of the $\overline{\nu}_e$ normalization uncertainty due to the selection cuts on the prompt and delayed trigger visible energies, and a more precise Rate+Shape analysis. One of these improvements is on the efficiency of the reconstruction of low energy signals, which has increased by 4% with respect to the first result on θ_{13} [69]. A good control of the energy scale will remain of central importance to ensure the identical response of the Near and Far detectors, and to further the Near Detector physics program: search for sterile neutrinos, discriminate between models explaining the origin of the spectral distortion, and eventually produce a model-independent measurement of the reactor $\overline{\nu}_e$ flux and spectrum.

The neutron detection efficiency has been shown to be the dominant contribution to the $\overline{\nu}_e$ detection uncertainty in the Gd-based selection, and the next-to-dominant in the H-based selection (in which the proton number uncertainty prevails). Nevertheless, the relative uncertainties have been reduced with respect to the previous θ_{13} measurements: from 0.96% of the previous Gd selection [1] to the current 0.54%, and from 1.25% of the previous H [72] selection to the current 0.42%. Moreover, for the first time in Double Chooz, a direct

measurement of the neutron selection efficiency in the full volume of the detector has been performed using the neutrons from the inverse β-decay (IBD). In the future two-detector measurement in the Gd channel, as a consequence of the identicalness of the detectors, the contribution of the neutron detection efficiency uncertainty is expected to reach 0.2%, from the uncorrelated component between detectors. It will be dominated by the neutron selection efficiency, which is measured most precisely using the IBD neutron source as it has been described in this thesis. The uncertainty on the IBD neutron measurement consists of approximately equal statistic and systematic parts. While the statistics in the Far Detector will continue improving slowly; the systematic uncertainty arising from the stopping-muon contamination could be brought down if new veto techniques against the stopping-muon background are implemented. Another advantage of the IBD source to estimate the neutron detection efficiency compared to the ^{252}Cf is that it does not require to take dedicated calibration runs, so the detectors can operate with a higher duty cycle.

Given the good performance demonstrated by the Artificial Neural Network-based H selection, which has reached a precision comparable to that of the Gd selection, an extension to include both H and Gd captures at selection level can be envisioned. This would not only boost the statistics, but also would remove the dominant systematic uncertainty on the neutron detection with a single detector, the Gd/H fraction, since the two neutron capture possibilities would be treated in an inclusive way. Furthermore, as the Target–Gamma Catcher distinction would disappear effectively, this would also reduce the uncertainty due to the migration of neutrons. Consequently, the single-detector physics capabilities would be enhanced.

The level of precision on θ_{13} that will be achieved by the three reactor experiments (Double Chooz, Daya Bay and RENO) will not be surpassed by other experiments in the near future; and unlike the accelerator-based experiments, their measurement does not require assumptions on the values of the other mixing angles or the CP violation phase. An eventual combination of the three reactor measurements should provide the most precise θ_{13} measurement ever, to be used as an input parameter by all the subsequent experiments searching for the neutrino mass hierarchy or the CP violation.

References

[1] **Double Chooz** Collaboration, Y. Abe et al., *Reactor electron antineutrino disappearance in the Double Chooz experiment*, *Phys.Rev.* **D86** (2012) 052008, [arXiv:1207.6632].

[2] **Double Chooz** Collaboration, Y. Abe et al., *Improved measurements of the neutrino mixing angle θ_{13} with the Double Chooz detector*, *JHEP* **1410** (2014) 86 [Erratum ibid. 1502 (2015) 74], [arXiv:1406.7763].

[3] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984.

[4] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.

[5] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[6] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc.* **C680519** (1968) 367–377.

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.

[8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.

[10] **Particle Data Group** Collaboration, K. A. Olive et al., *Review of Particle Physics*, *Chin.Phys.* **C38** (2014) 090001.

[11] **ATLAS** Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214].

[12] **CMS** Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [arXiv:1207.7235].

[13] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of Neutrinos*, *Phys. Rev.* **109** (1958) 1015–1017.

[14] L. M. Brown, *The idea of the neutrino*, *Phys. Today* **31N9** (1978) 23–28.

[15] J. Chadwick, *Possible Existence of a Neutron*, *Nature* **129** (1932) 312.

[16] E. Fermi, *Trends to a Theory of beta Radiation. (In Italian)*, *Nuovo Cim.* **11** (1934) 1–19.

[17] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the free neutrino: A Confirmation*, *Science* **124** (1956) 103–104.

[18] G. Feinberg, *Decays of the mu Meson in the Intermediate-Meson Theory*, *Phys. Rev.* **110** (1958) 1482–1483.

[19] B. Pontecorvo, *Electron and Muon Neutrinos*, *Sov. Phys. JETP* **10** (1960) 1236–1240. [Zh. Eksp. Teor. Fiz.37,1751(1959)].

[20] G. Danby, J. M. Gaillard, K. A. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz, and J. Steinberger, *Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*, *Phys. Rev. Lett.* **9** (1962) 36–44.

[21] **DONUT** Collaboration, K. Kodama et al., *Observation of tau neutrino interactions*, *Phys. Lett.* **B504** (2001) 218–224, [hep-ex/0012035].

[22] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e+ - e- Annihilation*, *Phys. Rev. Lett.* **35** (1975) 1489–1492.

[23] **SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3** Collaboration, S. Schael et al., *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257–454, [hep-ex/0509008].

[24] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533.

[25] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[26] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, *Sov. Phys. JETP* **26** (1968) 984–988. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[27] V. N. Gribov and B. Pontecorvo, *Neutrino astronomy and lepton charge*, *Phys. Lett.* **B28** (1969) 493.

[28] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870–880.

[29] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics*. 2007.

[30] B. Kayser, *On the Quantum Mechanics of Neutrino Oscillation*, *Phys. Rev.* **D24** (1981) 110.

[31] W. Grimus and P. Stockinger, *Real oscillations of virtual neutrinos*, *Phys. Rev.* **D54** (1996) 3414–3419, [hep-ph/9603430].

[32] L. Wolfenstein, *Neutrino Oscillations in Matter*, *Phys. Rev.* **D17** (1978) 2369–2374.

[33] S. P. Mikheev and A. Yu. Smirnov, *Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos*, *Sov. J. Nucl. Phys.* **42** (1985) 913–917. [Yad. Fiz.42,1441(1985)].

[34] S. P. Mikheev and A. Yu. Smirnov, *Resonant amplification of neutrino oscillations in matter and solar neutrino spectroscopy*, *Nuovo Cim.* **C9** (1986) 17–26.

[35] X. Qian and P. Vogel, *Neutrino Mass Hierarchy*, *Prog. Part. Nucl. Phys.* **83** (2015) 1–30, [arXiv:1505.0189].

[36] J. N. Bahcall, *Solar neutrinos. I: Theoretical*, *Phys. Rev. Lett.* **12** (1964) 300–302.

[37] R. Davis, *Solar neutrinos. II: Experimental*, *Phys. Rev. Lett.* **12** (1964) 303–305.

[38] National Nuclear Data Center, "Information extracted from the NuDat 2 database." http://www.nndc.bnl.gov/nudat2/, 2015.

[39] C. Giunti and M. Laveder, *Essential solar neutrinos*, hep-ph/0301276.

[40] J. N. Bahcall, A. M. Serenelli, and S. Basu, *New solar opacities, abundances, helioseismology, and neutrino fluxes*, *Astrophys. J.* **621** (2005) L85–L88, [astro-ph/0412440].

[41] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, *Astrophys. J.* **496** (1998) 505–526.

[42] J. N. Bahcall and M. H. Pinsonneault, *What do we (not) know theoretically about solar neutrino fluxes?*, *Phys. Rev. Lett.* **92** (2004) 121301, [astro-ph/0402114].

[43] **GNO** Collaboration, M. Altmann et al., *Complete results for five years of GNO solar neutrino observations*, *Phys. Lett.* **B616** (2005) 174–190, [hep-ex/0504037].

[44] **SAGE** Collaboration, J. N. Abdurashitov et al., *Solar neutrino flux measurements by the Soviet-American Gallium Experiment (SAGE) for half the 22 year solar cycle*, *J. Exp. Theor. Phys.* **95** (2002) 181–193, [astro-ph/0204245]. [Zh. Eksp. Teor. Fiz.122,211(2002)].

[45] **Kamiokande** Collaboration, Y. Fukuda et al., *Solar neutrino data covering solar cycle 22*, *Phys. Rev. Lett.* **77** (1996) 1683–1686.

[46] **Super-Kamiokande** Collaboration, J. Hosaka et al., *Solar neutrino measurements in Super-Kamiokande-I*, *Phys. Rev.* **D73** (2006) 112001, [hep-ex/0508053].

[47] **SNO** Collaboration, Q. R. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301, [nucl-ex/0204008].

[48] **SNO** Collaboration, B. Aharmim et al., *Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory*, *Phys. Rev.* **C72** (2005) 055502, [nucl-ex/0502021].

[49] **SNO** Collaboration, B. Aharmim et al., *An Independent Measurement of the Total Active B-8 Solar Neutrino Flux Using an Array of He-3 Proportional Counters at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **101** (2008) 111301, [arXiv:0806.0989].

[50] **Super-Kamiokande** Collaboration, S. Fukuda et al., *Determination of solar neutrino oscillation parameters using 1496 days of Super-Kamiokande I data*, *Phys. Lett.* **B539** (2002) 179–187, [hep-ex/0205075].

[51] **KamLAND** Collaboration, A. Gando et al., *Reactor On-Off Antineutrino Measurement with KamLAND*, *Phys. Rev.* **D88** (2013), no. 3 033001, [arXiv:1303.4667].

[52] **Borexino** Collaboration, G. Bellini et al., *Final results of Borexino Phase-I on low energy solar neutrino spectroscopy*, *Phys. Rev.* **D89** (2014), no. 11 112007, [arXiv:1308.0443].

[53] **Kamiokande-II** Collaboration, K. S. Hirata et al., *Experimental Study of the Atmospheric Neutrino Flux*, *Phys. Lett.* **B205** (1988) 416.

[54] R. Becker-Szendy et al., *The Electron-neutrino and muon-neutrino content of the atmospheric flux*, *Phys. Rev.* **D46** (1992) 3720–3724.

[55] **Super-Kamiokande** Collaboration, R. Wendell et al., *Atmospheric neutrino oscillation analysis with sub-leading effects in Super-Kamiokande I, II, and III*, *Phys. Rev.* **D81** (2010) 092004, [arXiv:1002.3471].

[56] **Soudan-2** Collaboration, W. W. M. Allison et al., *Neutrino oscillation effects in Soudan-2 upward-stopping muons*, *Phys. Rev.* **D72** (2005) 052005, [hep-ex/0507068].

[57] **MACRO** Collaboration, M. Ambrosio et al., *Measurements of atmospheric muon neutrino oscillations, global analysis of the data collected with MACRO detector*, *Eur. Phys. J.* **C36** (2004) 323–339.

[58] **K2K** Collaboration, M. H. Ahn et al., *Measurement of Neutrino Oscillation by the K2K Experiment*, *Phys. Rev.* **D74** (2006) 072003, [hep-ex/0606032].

[59] **MINOS** Collaboration, P. Adamson et al., *Combined analysis of ν_μ disappearance and $\nu_\mu \to \nu_e$ appearance in MINOS using accelerator and atmospheric neutrinos*, *Phys.Rev.Lett.* **112** (2014) 191801, [arXiv:1403.0867].

[60] **T2K** Collaboration, K. Abe et al., *Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam*, *Phys. Rev. Lett.* **112** (2014), no. 18 181801, [arXiv:1403.1532].

[61] **OPERA** Collaboration, N. Agafonova et al., *Discovery of τ Neutrino Appearance in the CNGS Neutrino Beam with the OPERA Experiment*, *Phys. Rev. Lett.* **115** (2015), no. 12 121802, [arXiv:1507.0141].

[62] P. Vogel, L. Wen, and C. Zhang, *Neutrino Oscillation Studies with Reactors*, arXiv:1503.0105.

[63] **CHOOZ** Collaboration, M. Apollonio et al., *Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station*, *Eur.Phys.J.* **C27** (2003) 331–374, [hep-ex/0301017].

[64] **CHOOZ** Collaboration, M. Apollonio et al., *Initial results from the CHOOZ long baseline reactor neutrino oscillation experiment*, *Phys. Lett.* **B420** (1998) 397–404, [hep-ex/9711002].

[65] F. Boehm et al., *Final results from the Palo Verde neutrino oscillation experiment*, *Phys. Rev.* **D64** (2001) 112001, [hep-ex/0107009].

[66] **Double Chooz** Collaboration, F. Ardellier et al., *Double Chooz: A Search for the neutrino mixing angle theta(13)*, hep-ex/0606025.

[67] **Daya Bay** Collaboration, X. Guo et al., *A Precision measurement of the neutrino mixing angle θ_{13} using reactor antineutrinos at Daya-Bay*, hep-ex/0701029.

[68] **RENO** Collaboration, J. K. Ahn et al., *RENO: An Experiment for Neutrino Oscillation Parameter θ_{13} Using Reactor Neutrinos at Yonggwang*, arXiv:1003.1391.

[69] **Double Chooz** Collaboration, Y. Abe et al., *Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment*, *Phys.Rev.Lett.* **108** (2012) 131801, [arXiv:1112.6353].

[70] **Daya Bay** Collaboration, F. P. An et al., *Observation of electron-antineutrino disappearance at Daya Bay*, *Phys.Rev.Lett.* **108** (2012) 171803, [arXiv:1203.1669].

[71] **RENO** Collaboration, J. K. Ahn et al., *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, *Phys. Rev. Lett.* **108** (2012) 191802, [arXiv:1204.0626].

[72] **Double Chooz** Collaboration, Y. Abe et al., *First Measurement of θ_{13} from Delayed Neutron Capture on Hydrogen in the Double Chooz Experiment*, *Phys.Lett.* **B723** (2013) 66–70, [arXiv:1301.2948].

[73] **Double Chooz** Collaboration, C. Buck, *The Double Chooz experiment*, *PoS* **EPS-HEP2013** (2013) 514.

[74] **Double Chooz** Collaboration, Y. Abe et al., *Background-independent measurement of θ_{13} in Double Chooz*, *Phys. Lett.* **B735** (2014) 51–56, [arXiv:1401.5981].

[75] **T2K** Collaboration, K. Abe et al., *Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam*, *Phys. Rev. Lett.* **107** (2011) 041801, [arXiv:1106.2822].

[76] **MINOS** Collaboration, P. Adamson et al., *Improved search for muon-neutrino to electron-neutrino oscillations in MINOS*, *Phys. Rev. Lett.* **107** (2011) 181802, [arXiv:1108.0015].

[77] M. Freund, *Analytic approximations for three neutrino oscillation parameters and probabilities in matter*, *Phys. Rev.* **D64** (2001) 053003, [hep-ph/0103300].

[78] **T2K** Collaboration, K. Abe et al., *Observation of Electron Neutrino Appearance in a Muon Neutrino Beam*, *Phys. Rev. Lett.* **112** (2014) 061802, [arXiv:1311.4750].

[79] **Troitsk** Collaboration, V. N. Aseev et al., *An upper limit on electron antineutrino mass from Troitsk experiment*, *Phys. Rev.* **D84** (2011) 112003, [arXiv:1108.5034].

[80] C. Kraus et al., *Final results from phase II of the Mainz neutrino mass search in tritium beta decay*, *Eur. Phys. J.* **C40** (2005) 447–468, [hep-ex/0412056].

[81] **KATRIN** Collaboration, G. Drexlin, *KATRIN: Direct measurement of a sub-eV neutrino mass*, *Nucl. Phys. Proc. Suppl.* **145** (2005) 263–267. [,263(2005)].

[82] 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** (1996) 6065–6077.

[83] **ALEPH** Collaboration, R. Barate et al., *An Upper limit on the tau-neutrino mass from three-prong and five-prong tau decays*, *Eur. Phys. J.* **C2** (1998) 395–406.

[84] P. Minkowski, $\mu \to e\gamma$ *at a Rate of One Out of* 10^9 *Muon Decays?*, *Phys. Lett.* **B67** (1977) 421–428.

[85] P. Ramond, *The Family Group in Grand Unified Theories*, in *International Symposium on Fundamentals of Quantum Theory and Quantum Field Theory Palm Coast, Florida, February 25-March 2, 1979*, pp. 265–280, 1979. hep-ph/9809459.

[86] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*, *Conf. Proc.* **C790927** (1979) 315–321, [arXiv:1306.4669].

[87] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Violation*, *Phys. Rev. Lett.* **44** (1980) 912.

[88] T. Yanagida, *Horizontal Symmetry and Masses of Neutrinos*, *Conf. Proc.* **C7902131** (1979) 95–99.

[89] J. Schechter and J. W. F. Valle, *Neutrinoless Double beta Decay in SU(2) x U(1) Theories*, *Phys. Rev.* **D25** (1982) 2951.

[90] S. Dell'Oro, S. Marcocci, and F. Vissani, *New expectations and uncertainties on neutrinoless double beta decay*, *Phys. Rev.* **D90** (2014), no. 3 033005, [arXiv:1404.2616].

[91] H. V. Klapdor-Kleingrothaus and I. V. Krivosheina, *The evidence for the observation of 0nu beta beta decay: The identification of 0nu beta beta events from the full spectra*, *Mod. Phys. Lett.* **A21** (2006) 1547–1566.

[92] **GERDA** Collaboration, M. Agostini et al., *Results on Neutrinoless Double-β Decay of ^{76}Ge from Phase I of the GERDA Experiment*, *Phys. Rev. Lett.* **111** (2013), no. 12 122503, [arXiv:1307.4720].

[93] **EXO-200** Collaboration, M. Auger et al., *Search for Neutrinoless Double-Beta Decay in ^{136}Xe with EXO-200*, *Phys. Rev. Lett.* **109** (2012) 032505, [arXiv:1205.5608].

[94] **EXO-200** Collaboration, J. B. Albert et al., *Search for Majorana neutrinos with the first two years of EXO-200 data*, *Nature* **510** (2014) 229–234, [arXiv:1402.6956].

[95] **KamLAND-Zen** Collaboration, A. Gando et al., *Limit on Neutrinoless $\beta\beta$ Decay of ^{136}Xe from the First Phase of KamLAND-Zen and Comparison with the Positive Claim in ^{76}Ge*, *Phys. Rev. Lett.* **110** (2013), no. 6 062502, [arXiv:1211.3863].

[96] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.0158.

[97] G. Mangano, G. Miele, S. Pastor, and M. Peloso, *A Precision calculation of the effective number of cosmological neutrinos*, *Phys. Lett.* **B534** (2002) 8–16, [astro-ph/0111408].

[98] P. Huber, *On the determination of anti-neutrino spectra from nuclear reactors*, *Phys.Rev.* **C84** (2011) 024617, [arXiv:1106.0687].

[99] M. C. Gonzalez-Garcia, M. Maltoni, and T. Schwetz, *Updated fit to three neutrino mixing: status of leptonic CP violation*, *JHEP* **11** (2014) 052, [arXiv:1409.5439].

[100] **NOvA** Collaboration, D. S. Ayres et al., *NOvA: Proposal to build a 30 kiloton off-axis detector to study $\nu_\mu \to \nu_e$ oscillations in the NuMI beamline*, hep-ex/0503053.

[101] K. Abe et al., *Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —*, arXiv:1109.3262.

[102] **LBNE** Collaboration, C. Adams et al., *The Long-Baseline Neutrino Experiment: Exploring Fundamental Symmetries of the Universe*, 2013. arXiv:1307.7335.

[103] **IceCube PINGU** Collaboration, M. G. Aartsen et al., *Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU)*, arXiv:1401.2046.

[104] **KM3NeT** Collaboration, U. F. Katz, *The ORCA Option for KM3NeT*, in *Proceedings of the 15th International Workshop on Neutrino Telescopes (Neutel 2013)*, 2014. arXiv:1402.1022.

[105] **ICAL** Collaboration, S. Ahmed et al., *Physics Potential of the ICAL detector at the India-based Neutrino Observatory (INO)*, arXiv:1505.0738.

[106] **JUNO** Collaboration, F. An et al., *Neutrino Physics with JUNO*, arXiv:1507.0561.

[107] **LSND** Collaboration, A. Aguilar-Arevalo et al., *Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam*, *Phys. Rev.* **D64** (2001) 112007, [hep-ex/0104049].

[108] **KARMEN** Collaboration, B. Armbruster et al., *Upper limits for neutrino oscillations $\overline{\nu}_\mu \to \text{ø}verline\nu_e$ from muon decay at rest*, *Phys. Rev.* **D65** (2002) 112001, [hep-ex/0203021].

[109] **MiniBooNE** Collaboration, A. A. Aguilar-Arevalo et al., *Improved Search for $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillations in the MiniBooNE Experiment*, *Phys. Rev. Lett.* **110** (2013) 161801, [arXiv:1207.4809].

[110] **MicroBooNE** Collaboration, H. Chen et al., *Proposal for a New Experiment Using the Booster and NuMI Neutrino Beamlines: MicroBooNE*, .

[111] **LArTPC** Collaboration, C. Adams et al., *LAr1-ND: Testing Neutrino Anomalies with Multiple LArTPC Detectors at Fermilab*, in *Community Summer Study 2013: Snowmass on the Mississippi (CSS2013) Minneapolis, MN, USA, July 29-August 6, 2013*, 2013. arXiv:1309.7987.

[112] D. Gibin and J. Kisiel, *ICARUS report to the CXVIII Meeting of SPSC, June 23-24, 2015*, Tech. Rep. CERN-SPSC-2015-024. SPSC-SR-163, CERN, Geneva, Jun, 2015.

[113] C. Giunti and M. Laveder, *Statistical Significance of the Gallium Anomaly*, *Phys. Rev.* **C83** (2011) 065504, [arXiv:1006.3244].

[114] T. A. Mueller, D. Lhuillier, M. Fallot, A. Letourneau, S. Cormon, et al., *Improved Predictions of Reactor Antineutrino Spectra*, *Phys.Rev.* **C83** (2011) 054615, [arXiv:1101.2663].

[115] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, *The Reactor Antineutrino Anomaly*, *Phys. Rev.* **D83** (2011) 073006, [arXiv:1101.2755].

[116] **Borexino** Collaboration, G. Bellini et al., *SOX: Short distance neutrino Oscillations with BoreXino*, *JHEP* **08** (2013) 038, [arXiv:1304.7721].

[117] A. Gando et al., *CeLAND: search for a 4th light neutrino state with a 3 PBq 144Ce-144Pr electron antineutrino generator in KamLAND*, arXiv:1312.0896.

[118] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, *Phys. Lett.* **B174** (1986) 45.

[119] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35. [Usp. Fiz. Nauk161,61(1991)].

[120] Y. Kozlov, L. Mikaelyan, and V. Sinev, *Two detector reactor neutrino oscillation experiment Kr2Det at Krasnoyarsk: Status report*, *Phys.Atom.Nucl.* **66** (2003) 469–471, [hep-ph/0109277].

[121] **RENO** Collaboration, J. K. Ahn et al., *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, *Phys.Rev.Lett.* **108** (2012) 191802, [arXiv:1204.0626].

[122] A. S. Cucoanes, P. Novella, A. Cabrera, M. Fallot, A. Onillon, et al., *Reactor Neutrino Flux Uncertainty Suppression on Multiple Detector Experiments*, arXiv:1501.0035.

[123] P. Vogel and J. F. Beacom, *Angular distribution of neutron inverse beta decay, $\bar{\nu}_e + p \to e^+ + n$*, *Phys.Rev.* **D60** (1999) 053003, [hep-ph/9903554].

[124] A. Pichlmaier, V. Varlamov, K. Schreckenbach, and P. Geltenbort, *Neutron lifetime measurement with the UCN trap-in-trap MAMBO II*, *Phys.Lett.* **B693** (2010) 221–226.

[125] D. H. Wilkinson, *Phase space for neutron beta-decay: An update*, *Nucl.Instrum.Meth.* **A404** (1998) 305–310.

[126] S. A. Fayans, *Radiative Corrections and Recoil Effects in the Reaction $\bar{\nu}_e + P \to N + e^+$ at Low-energies. (In Russian)*, *Sov.J.Nucl.Phys.* **42** (1985) 590.

[127] C. Aberle, C. Buck, B. Gramlich, F. X. Hartmann, M. Lindner, et al., *Large scale Gd-beta-diketonate based organic liquid scintillator production for antineutrino detection*, *JINST* **7** (2012) P06008, [arXiv:1112.5941].

[128] C. Aberle, C. Buck, F. X. Hartmann, and S. Schonert, *Light yield and energy transfer in a new Gd-loaded liquid scintillator*, *Chem.Phys.Lett.* **516** (2011) 257–262.

[129] E. Calvo, M. Cerrada, C. Fernandez-Bedoya, I. Gil-Botella, C. Palomares, et al., *Characterization of large area photomutipliers under low magnetic fields: Design and performances of the magnetic shielding for the Double Chooz neutrino experiment*, *Nucl.Instrum.Meth.* **A621** (2010) 222–230, [arXiv:0905.3246].

[130] D. Dietrich, D. Greiner, J. Jochum, T. Lachenmaier, M. Rohling, et al., *Monte Carlo aided design of the inner muon veto detectors for the Double Chooz experiment*, *JINST* **7** (2012) P08012, [arXiv:1207.1623].

[131] F. Beissel, A. Cabrera, A. Cucoanes, J. V. Dawson, D. Kryn, et al., *The Trigger and Timing System of the Double Chooz Experiment*, *JINST* **8** (2013) T01003, [arXiv:1209.5607].

[132] O. Meplan et al., *MURE: MCNP Utility for Reactor Evolution - Description of the methods, first applications and results*, in *Proceedings of the European Nuclear Conference (ENC 2005). Nuclear Power for the XXIst Century: From basic research to high-tech industry*, 2005.

[133] MURE, MCNP Utility for Reactor Evolution: couples Monte-Carlo transport with fuel burnup calculations, "NEA-1845/01." `http://www.oecd-nea.org/tools/abstract/detail/nea-1845`, 2009.

[134] X-5 Monte Carlo Team, *MCNP – A General Monte Carlo N-Particle Transport Code, Version 5*, Tech. Rep. LA-UR-03-1987, Los Alamos National Laboratory, Los Alamos (New Mexico), Apr, 2003.

[135] G. Marleau et al., *DRAGON: A Collision Probability Transport Code for Cell and Supercell Calculations*, Tech. Rep. Report IGE-157, Institut de génie nucléaire, École Polytechnique de Montréal, Montréal, Québec, 1994.

[136] C. L. Jones, A. Bernstein, J. M. Conrad, Z. Djurcic, M. Fallot, et al., *Reactor Simulation for Antineutrino Experiments using DRAGON and MURE*, *Phys.Rev.* **D86** (2012) 012001, [`arXiv:1109.5379`].

[137] V. Kopeikin, L. Mikaelyan, and V. Sinev, *Reactor as a source of antineutrinos: Thermal fission energy*, *Phys.Atom.Nucl.* **67** (2004) 1892–1899, [`hep-ph/0410100`].

[138] M. Fallot and A. Onillon, *Reactor plots*, *Double Chooz Internal* **DC-doc-3377** (2011).

[139] F. Von Feilitzsch, A. A. Hahn, and K. Schreckenbach, *Experimental beta spectra from Pu-239 and U-235 thermal neutron fission products and their correlated anti-neutrinos spectra*, *Phys.Lett.* **B118** (1982) 162–166.

[140] K. Schreckenbach, G. Colvin, W. Gelletly, and F. Von Feilitzsch, *Determination of the anti-neutrino spectrum from U-235 thermal neutron fission products up to 9.5-MeV*, *Phys.Lett.* **B160** (1985) 325–330.

[141] A. A. Hahn et al., *Anti-neutrino Spectra From ^{241}Pu and ^{239}Pu Thermal Neutron Fission Products*, *Phys.Lett.* **B218** (1989) 365–368.

[142] N. Haag et al., *Experimental Determination of the Antineutrino Spectrum of the Fission Products of ^{238}U*, *Phys.Rev.Lett.* **112** (2014), no. 12 122501, [`arXiv:1312.5601`].

[143] 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** (2012) 202504, [`arXiv:1208.3877`].

[144] Y. Declais et al., *Study of reactor anti-neutrino interaction with proton at Bugey nuclear power plant*, *Phys.Lett.* **B338** (1994) 383–389.

[145] A. J. Franke, *Searching for Reactor Antineutrino Flavor Oscillations with the Double Chooz Far Detector*. PhD thesis, Columbia University, Dec, 2012.

[146] **GEANT4** Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl.Instrum.Meth.* **A506** (2003) 250–303.

[147] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. A. Dubois, et al., *Geant4 developments and applications*, *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

[148] D. Motta and S. Schonert, *Optical properties of Bialkali photocathodes*, *Nucl.Instrum.Meth.* **A539** (2005) 217–235, [physics/0408075].

[149] C. Aberle, C. Buck, F. X. Hartmann, S. Schonert, and S. Wagner, *Light output of Double Chooz scintillators for low energy electrons*, *JINST* **6** (2011) P11006.

[150] C. Aberle, *Optimization, simulation and analysis of the scintillation signals in the Double Chooz experiment*. PhD thesis, Ruperto-Carola-University of Heidelberg, Dec, 2011.

[151] C. Buck and B. Reinhold, *Proton Number Tech Note*, *Double Chooz Internal* **DC-doc-4211** (2012).

[152] I. Ostrovskiy, *Measuring the neutrino mixing angle θ_{13} with the Double Chooz Far Detector*. PhD thesis, The University of Alabama, 2012.

[153] R. R. Spencer, R. Gwin, and R. Ingle, *A Measurement of the Average Number of Prompt Neutrons from Spontaneous Fission of Californium-252*, *Nucl. Sci. and Eng.* **80** (1982) 603–629.

[154] A. S. Cucoanes, *A Toy Model for the Gamma Spectrum Generated in n-Captures on Gd*, *Double Chooz Internal* **DC-doc-3139** (2011).

[155] Y. Abe, T. Akiri, A. Cabrera, B. Courty, J. V. Dawson, et al., *The Waveform Digitiser of the Double Chooz Experiment: Performance and Quantisation Effects on PhotoMultiplier Tube Signals*, *JINST* **8** (2013) P08015, [arXiv:1307.4917].

[156] Y. Abe, C. Buck, A. Cabrera, M. Ishitsuka, G. Pronost, J. Reichenbacher, B. Reinhold, L. F. F. Stokes, and S. Wagner, *DCIII Energy Blessing candidates (ESv5)*, *Double Chooz Internal* **DC-doc-5603** (2014).

[157] A. Chyzh, B. Baramsai, J. A. Becker, F. Becvar, T. A. Bredeweg, et al., *Measurement of the Gd-157 (n, gamma) reaction with the DANCE gamma calorimeter array*, *Phys.Rev.* **C84** (2011) 014306.

[158] B. Baramsai, G. E. Mitchell, U. Agvaanluvsan, F. Becvar, T. A. Bredeweg, et al., *Neutron resonance parameters in Gd-155 measured with the DANCE gamma-ray calorimeter array*, *Phys.Rev.* **C85** (2012) 024622.

[159] **Double Chooz** Collaboration, Y. Abe et al., *Improved measurement of θ_{13} in Double Chooz using neutron captures on hydrogen using novel background rejection techniques*, .

[160] J. I. Crespo-Anadón, *Estudio de la emisión de luz en los tubos fotomultiplicadores del experimento double chooz*, Master's thesis, Universidad Complutense de Madrid, Jun, 2011.

[161] E. Calvo et al., *Toward a better understanding of the glowing: a possible explanation of the light emission process*, *Double Chooz Internal* **DC-doc-3234** (2011).

[162] **Double Chooz** Collaboration, Y. Abe et al., *Characterization of the Spontaneous Light Emission of the PMTs used in the Double Chooz Experiment*, .

[163] R. Roncin, *From the measurement of the θ_{13} mixing angle to the search for geo-neutrinos: studying $\overline{\nu}_e$ with Double Chooz and Borexino*. PhD thesis, Université Paris Diderot (Paris 7) Sorbonne Paris Cité/Universitá degli Studi dell'Aquila, Sep, 2014.

[164] P. Novella, *DCIII nGd Candidates: Blessed Plots*, *Double Chooz Internal* **DC-doc-5608** (2014).

[165] E. Conover, *Muon-induced backgrounds in the Double Chooz neutrino oscillation experiment*. PhD thesis, The University of Chicago, Jun, 2014.

[166] **KamLAND** Collaboration, S. Abe et al., *Production of Radioactive Isotopes through Cosmic Muon Spallation in KamLAND*, *Phys.Rev.* **C81** (2010) 025807, [arXiv:0907.0066].

[167] **Double Chooz** Collaboration, Y. Abe et al., *Precision Muon Reconstruction in Double Chooz*, *Nucl.Inst.Meth.Phys.Res.* **A** (2014) [arXiv:1405.6227].

[168] J. M. López-Castaño, *Evaluación del fondo del detector en la medida de θ_{13} en el experimento Double Chooz*. PhD thesis, Universidad Complutense de Madrid, In preparation, 2015.

[169] J. M. López-Castaño, *Acc plots to be blessed*, *Double Chooz Internal* **DC-doc-5541** (2014).

[170] **Double Chooz** Collaboration, Y. Abe et al., *Direct Measurement of Backgrounds using Reactor-Off Data in Double Chooz*, *Phys.Rev.* **D87** (2013) 011102, [arXiv:1210.3748].

[171] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al., *TMVA - Toolkit for Multivariate Data Analysis*, *PoS* **ACAT** (2007) 040, [physics/0703039].

[172] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl.Instrum.Meth.* **A389** (1997) 81–86.

[173] R. Sharankova and J. I. Crespo-Anadón, *H-III: ANN plots for blessing*, *Double Chooz Internal* **DC-doc-6106** (2015).

[174] G. Pronost, *HIII IVV OVV blessing package*, *Double Chooz Internal* **DC-doc-6181** (2015).

[175] A. Hourlier, *HIII FVV blessing package*, *Double Chooz Internal* **DC-doc-6114** (2015).

[176] D. C. A. Group, *H-III analysis tech-note*, *Double Chooz Internal* **DC-doc-5918** (2015).

[177] A. Meregaglia, *MPS-Blessing*, *Double Chooz Internal* **DC-doc-6104** (2015).

[178] Y. Abe et al., *Ortho-positronium observation in the Double Chooz Experiment*, *JHEP* **10** (2014) 32, [arXiv:1407.6913].

[179] A. Cabrera, M. Ishitsuka, A. Meregaglia, and G. Pronost, *H-III Correlated Background Blessing Package*, *Double Chooz Internal* **DC-doc-6105** (2015).

[180] S. L. F. F., *HIII LiHe Plots for blessing*, *Double Chooz Internal* **DC-doc-6099** (2015).

[181] G. A. Horton-Smith, *Composition and volume of scintillator in MC*, *Double Chooz Internal* **DC-doc-3188** (2011).

[182] A. Etenko, *Slow neutrons modelling*, *Double Chooz Internal* **DC-doc-1264** (2009).

[183] A. Etenko, *Slow neutrons modelling*, *Double Chooz Internal* **DC-doc-1430** (2010).

[184] J. R. Granada, *Slow-neutron scattering by molecular gases: A synthetic scattering function*, *Phys. Rev. B* **31** (Apr, 1985) 4167–4177.

[185] *Plot produced online using the Evaluated Nuclear Data File (ENDF)*, *National Nuclear Data Center, Brookhaven National Laboratory* **ENDF/B-VII.1 and CENDL-3.1 databases** (2015).

[186] J. Neyman, *Outline of a theory of statistical estimation based on the classical theory of probability*, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **236** (1937), no. 767 pp. 333–380.

[187] C. J. Clopper and E. S. Pearson, *The use of confidence or fiducial limits illustrated in the case of the binomial*, *Biometrika* **26** (1934), no. 4 pp. 404–413.

[188] "TEfficiency Class Reference." https://root.cern.ch/doc/master/classTEfficiency.html, 2015.

[189] J. Haser, *Neutron Detection Uncertainties in the θ_{13} Analysis of the Double Chooz Experiment*. PhD thesis, Ruperto-Carola-University of Heidelberg, Feb, 2015.

[190] J. Haser, *n-Gd Neutron Detection Efficiency Volume Factorization Method using 252 Cf*, *Double Chooz Internal* **DC-doc-5293** (2013).

[191] A. P. Collin, J. I. Crespo-Anadón, J. Haser, and G. Yang, *Measurement of the detection systematic uncertainty in the Double Chooz experiment*, in *Poster in the* 37^{th} *International Conference on High Energy Physics (ICHEP 2014)*, 2014.

[192] K. Terao, B. Reinhold, K. Nakajima, J. Maeda, S. Lucht, K. Crum, and R. Carr, *nH Analysis Technote*, *Double Chooz Internal* **DC-doc-4156** (2013).

[193] TRIPOLI-4 version 8.1, 3D general purpose continuous energy Monte Carlo Transport code, "NEA-1716/07." `http://www.oecd-nea.org/tools/abstract/detail/nea-1716`, 2013.

[194] J. Haser, "Private communication." 2015.

[195] P. Novella, *H-III Candidates plots for blessing*, *Double Chooz Internal* **DC-doc-6096** (2015).

[196] P. Novella, *DCIII nGd RRM: blessed plots*, *Double Chooz Internal* **DC-doc-5622** (2014).

[197] P. Novella, *RRM H-III Blessed plots*, *Double Chooz Internal* **DC-doc-6103** (2015).

[198] M. H. Shaevitz, *Separating Components of Error Matrices*, *BooNE Technical Note* **253** (2008).

[199] R. Carr and K. Crum, *Gd-III Rate+Shape results - blessed*, *Double Chooz Internal* **DC-doc-5610** (2014).

[200] R. Carr and G. Yang, *H-III Rate+Shape fit: Candidates for blessing*, *Double Chooz Internal* **DC-doc-6146** (2015).

[201] **Daya Bay** Collaboration, F. P. An et al., *Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay*, *Phys.Rev.Lett.* **112** (2014) 061801, [`arXiv:1310.6732`].

[202] **RENO** Collaboration, S.-H. Seo, *New Results from RENO and The 5 MeV Excess*, *AIP Conf. Proc.* **1666** (2015) 080002, [`arXiv:1410.7987`].

[203] **Daya Bay** Collaboration, F. P. An et al., *Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay*, `arXiv:1508.0423`.

[204] D. A. Dwyer and T. J. Langford, *Spectral Structure of Electron Antineutrinos from Nuclear Reactors*, *Phys. Rev. Lett.* **114** (2015), no. 1 012502, [`arXiv:1407.1281`].

[205] M. Chadwick, M. Herman, P. Obložinský, et al., *ENDF/B-VII.1 nuclear data for science and technology: Cross sections, covariances, fission product yields and decay data*, *Nuclear Data Sheets* **112** (2011), no. 12 2887 – 2996. Special Issue on ENDF/B-VII.1 Library.

[206] J. Hardy, L. Carraz, B. Jonson, and P. Hansen, *The essential decay of pandemonium: A demonstration of errors in complex beta-decay schemes*, *Physics Letters B* **71** (1977), no. 2 307 – 310.

[207] A. C. Hayes, J. L. Friar, G. T. Garvey, D. Ibeling, G. Jungman, T. Kawano, and R. W. Mills, *The Possible Origin and Implications of the Shoulder in Reactor Neutrino Spectra*, arXiv:1506.0058.

[208] Kellett, M. A. and Bersillon, O. and Mills, R. W., "The JEFF3.1/-3.1.1 Radioactive Decay Data and Fission Yields Sub-Libraries. JEFF Report 20, NEA Report No. 6287.." http://www.oecd-nea.org/dbforms/data/eva/evatapes/jeff_31/index-JEFF3.1.1.html, 2009.

[209] M. Toups, *Chooz Mega-mini Data*, *Double Chooz Internal* **DC-doc-1589** (2010).

[210] L. N. Kalousis, *Muon flux measurement in the near lab*, *Double Chooz Internal* **DC-doc-5006** (2013).

[211] **Daya Bay** Collaboration, F. P. An et al., *A new measurement of antineutrino oscillation with the full detector configuration at Daya Bay*, arXiv:1505.0345.