The NEXT experiment for neutrinoless double beta decay searches

PHD THESIS

Justo Martín-Albo

Supervisor: J.J. Gómez Cadenas
The work in this thesis has been carried out at the Instituto de Física Corpuscular (IFIC), affiliated with the Universitat de València and CSIC.
© Justo Martín-Albo, 2015
Marco Polo describes a bridge, stone by stone.
“But which is the stone that supports the bridge?”, Kublai Khan asks.
“The bridge is not supported by one stone or another”, Marco answers, “but by the line of the arch that they form”.
Kublai Khan remains silent, reflecting. Eventually, the Great Khan adds: “Why do you speak to me of the stones? It is only the arch that matters to me!”
To which Polo retorts: “Without stones there is no arch”.

— Italo Calvino, *Invisible Cities*

“It has long been an axiom of mine that the little things are infinitely the most important.”

— Arthur Conan Doyle, *The Adventures of Sherlock Holmes*

And in the dawn, armed with an ardent patience, we shall enter the splendid cities.

— Arthur Rimbaud, “Farewell”, *A Season in Hell*
Neutrinoless double beta ($0\nu\beta\beta$) decay is a postulated very slow radioactive process in which two neutrons inside a nucleus transform into two protons emitting two electrons. The discovery of this process would demonstrate that neutrinos are Majorana particles and that total lepton number is not conserved in nature, two findings with far-reaching implications in particle physics and cosmology. First, the existence of Majorana neutrinos implies a new energy scale at a level inversely proportional to the observed neutrino masses. Such a scale, besides providing a simple explanation for the striking lightness of neutrino masses, is probably connected to several open questions in particle physics, like the origin of mass or the flavour problem. Second, Majorana neutrinos violate the conservation of lepton number, and this, together with CP violation, could be responsible, through the mechanism known as leptogenesis, for the observed cosmological asymmetry between matter and antimatter.

After 75 years of experimental effort, no compelling evidence for the existence of $0\nu\beta\beta$ decay has been obtained, but a new generation of experiments that are already running or about to run promises to push forward the current limits exploring the degenerate-hierarchy region of neutrino masses. In order to do that, the experiments are using masses of $0\nu\beta\beta$ isotopes ranging from tens of kilograms to several hundreds, and will need to improve the background rates achieved by previous experiments by, at least, an order of magnitude. If no signal is found, masses of the order of thousands of kilograms and further background reduction will be required. In spite of the formidable experimental challenge (or possibly because of it), the field is brimming with new ideas. However, the requirements — which will be discussed extensively in this thesis — are often conflicting, and none of the considered technologies is capable of simultaneously optimizing all parameters.

Among the experiments of this new generation there is the Neutrino Experiment with a Xenon Tpc (NEXT), which will search for neutrinoless double beta...
decay using a high-pressure xenon gas, electroluminescent time projection chamber filled with 100 kg of xenon enriched in the $0\nu\beta\beta$ isotope $^{136}$Xe. Such a detector offers both high energy resolution and charged-particle tracking for the discrimination of signal and background. As it will be shown in this work, the combination results in excellent experimental sensitivity to $0\nu\beta\beta$ decay.

What about the longer-term future? What is the best approach for the exploration of the inverted-hierarchy region of neutrino masses? We have tried to answer this question in the last part of the dissertation. The diversity of experimental approaches we are currently witnessing will not be viable at that scale, and only two or three approaches (most likely based on different isotopes) are going to be retained. Their suitability for the task depends essentially on their scalability to large masses — including here the possible difficulties in the procurement and isotopic enrichment of tonnes of material — and on their capability to control the backgrounds to the required extremely low levels, which will have to be demonstrated in the current-generation experiments.

This thesis is organized in nine chapters. The key particle-physics concepts involving massive Majorana neutrinos and neutrinoless double beta decay are covered in Chapters 1 and 2, while Chapter 3 deals with the experimental aspects of $0\nu\beta\beta$-decay searches. These three chapters are mostly based on the following publications, of which I am one of the main authors:


Chapter 4 discusses the basic concepts behind the application of a xenon gas TPC for $0\nu\beta\beta$-decay searches, and Chapter 5 introduces the NEXT experiment, describing the NEXT-100 detector and the main results of the R&D phase. An important part of this material has been published in several papers with me as a lead author:


Chapter 6 describes the detector simulation of NEXT, essential for the estimation of the experiment’s sensitivity to $0\nu\beta\beta$ decay, which is discussed in Chapter 7. This material will become a future publication of the NEXT Collaboration. In Chapter 8, after comparing the sensitivities of the experiments of the current generation, the prospects and challenges of the future tonne-scale experiments are examined. Finally, Chapter 9 concludes.
Resumen

Introducción

La desintegración doble beta sin emisión de neutrinos \((0\nu\beta\beta)\) es un hipotético proceso radiactivo en el que un núcleo de número atómico \(Z\) y número másico \(A\) se transforma en su isóbaro de número atómico \(Z + 2\) emitiendo dos electrones:

\[
\begin{align*}
\begin{array}{c}
\text{La observación de esta desintegración probaría que el neutrino es una partícula} \\
\text{de tipo Majorana (es decir, indistinguible de su antipartícula) y que la ley de} \\
\text{conservación del número leptónico total puede violarse en las interacciones} \\
\text{físicas, dos hallazgos con profundas implicaciones en cosmología y física de} \\
\text{partículas. En primer lugar, la naturaleza Majorana de los neutrinos permitiría} \\
\text{explicar de forma natural la sorprendente ligereza de estas partículas en com-} \\
\text{paración con el resto de fermiones del Modelo Estándar. En segundo lugar, la} \\
\text{violación del número leptónico es una de las condiciones necesarias para la} \\
\text{leptogénesis primordial que podría haber generado la asimetría cósmica entre} \\
\text{materia y antimateria.}
\end{array}
\end{align*}
\]

En principio, cualquier mecanismo físico que involucre la violación del número leptónico podría inducir la desintegración \(0\nu\beta\beta\) y contribuir a su amplitud. En el caso más simple, la desintegración estaría mediada por el intercambio virtual de un neutrino ligero de naturaleza Majorana, y la semivida del proceso vendría dada por

\[
\left(T_{1/2}^{0\nu}\right)^{-1} = G_{0\nu}^2 |M_{0\nu}|^2 \left( \frac{m_{\beta\beta}}{m_e} \right)^2.
\]

En esta ecuación, \(G_{0\nu}\) es una integral de espacio de fases, \(M_{0\nu}\) es el elemento de matriz nuclear del proceso, \(m_e\) es la masa del electrón y \(m_{\beta\beta}\) es la llamada
masa Majorana efectiva del neutrino:

\[ m_{\beta\beta} \equiv \langle \sum_{i=1}^{3} U_{ei}^2 m_i \rangle, \]  

donde \( U_{ei} \) son los elementos de la primera fila de la matriz de mezcla de neutrinos y \( m_i \) son los autoestados de masa. Por tanto, una medida de la semivida de la desintegración \( 0\nu\beta\beta \) aportaría información directa sobre la escala de masas de los neutrones.

Hasta la fecha, no se ha obtenido evidencia experimental convincente de la existencia de la desintegración \( 0\nu\beta\beta \). El campo, no obstante, atraviesa en la actualidad una edad dorada, motivada posiblemente por un controvertido anuncio de descubrimiento en 2001 en el experimento Heidelberg-Moscow, y por la observación, hace ya más de 15 años, de las oscilaciones de neutrones, que implican que el neutrino tiene masa, condición necesaria para que se produzca la desintegración \( 0\nu\beta\beta \). Los experimentos de la generación actual emplean variadas técnicas de detección y masas de isótopo que oscilan entre unas pocas decenas de kilogramos hasta varios centenares. Tras unos pocos años de operación, estos experimentos alcanzarán sensibilidades a la semivida de la desintegración \( 0\nu\beta\beta \) próximas a \( 10^{26} \) años. Si no observaran ninguna señal, sería preciso expandir los experimentos a la escala de la tonelada para alcanzar sensibilidades superiores a \( 10^{27} \) años y, así, poder explorar completamente la región correspondiente a la jerarquía inversa de masas del neutrino.

**La búsqueda de la desintegración doble beta sin neutrones**

Los detectores empleados para buscar la desintegración \( 0\nu\beta\beta \) miden la energía de la radiación emitida por una fuente radiactiva apropiada. En una desintegración doble beta sin neutrones, la suma de las energías cinéticas de los dos electrones emitidos sería siempre igual al valor \( Q \) del proceso, esto es, a la diferencia de masas entre los átomos inicial y final:

\[ Q_{\beta\beta} \equiv M(Z_2X) - M(Z_{2+X}). \]  

En la práctica, la resolución energética finita de cualquier detector esparciría los posibles eventos \( 0\nu\beta\beta \) en torno a la posición \( Q_{\beta\beta} \) del espectro de energía siguiendo, típicamente, una distribución de probabilidad gaussiana. Cualquier otro proceso que deposite en el detector energía en esa región del espectro contribuye al ruido de fondo del experimento, dificultando la medida. La
resolución energética es, por tanto, un parámetro fundamental en este tipo de experimentos, pues mejora la relación señal a ruido, pero no es suficiente para garantizar un buen resultado, puesto que las fuentes de ruido potenciales son muy abundantes en la región de energías de interés. Por esta razón, todos los experimentos buscan signaturas adicionales que permitan diferenciar eficientemente los eventos de señal y ruido.

La principal fuente de ruido en la mayoría de experimentos que buscan la desintegración $0\nu\beta\beta$ es la radiactividad natural procedente de impurezas de $^{232}$Th y $^{238}$U presentes en prácticamente todos los materiales. La selección de componentes con la suficiente radiopureza es, por tanto, esencial; los detectores de la generación actual de experimentos se fabrican, en algunos casos, con materiales de purezas próximas a 1 $\mu$Bq/kg. Asimismo, los experimentos operan en instalaciones subterráneas para mitigar los ruidos generados por los rayos cósmicos.

Además de la resolución energética y del nivel de ruido en la ventana de energía alrededor de $Q_{\beta\beta}$, hay otros factores, como la eficiencia de detección y la escalabilidad a grandes masas, que deben tenerse en cuenta en el diseño de un experimento. La optimización simultánea de todos estos parámetros no es, en general, posible; consecuentemente, se han propuesto técnicas experimentales muy variadas para buscar la desintegración. Con el fin de comparar su rendimiento potencial, se suele emplear la sensibilidad experimental a $m_{\beta\beta}$:

$$S(m_{\beta\beta}) \propto \sqrt{1/\epsilon \frac{b \Delta E}{M t}}^{1/4},$$

(5)

donde $\epsilon$ es la eficiencia de detección del experimento, $M$ es la masa de isótopo $0\nu\beta\beta$ usada en el experimento, $t$ es el tiempo de medida, $\Delta E$ es la resolución energética del detector y $b$ es la tasa de ruido del experimento en la ventana de energía alrededor de $Q_{\beta\beta}$ (expresada, normalmente, en cuentas por kilo-electronvoltio, kilogramo de fuente y año).

Los primeros experimentos de la generación actual que han producido resultados, EXO-200 y KamLAND-Zen, están buscando la desintegración $0\nu\beta\beta$ del isótopo $^{136}$Xe. Ambos experimentos han obtenido resultados negativos, fijando un límite combinado a la semivida de la desintegración $0\nu\beta\beta$ del $^{136}$Xe de $3.5 \times 10^{25}$ años (90% CL). El experimento GERDA, que busca la desintegración usando detectores de germanio enriquecidos en el isótopo $^{76}$Ge, ha publicado también sus primeros resultados, en los que no hay no hay evidencia de un pico $0\nu\beta\beta$ en los datos, estableciendo un límite de $2.1 \times 10^{25}$ años (90% CL) para la semivida del proceso.
El experimento NEXT

NEXT buscará la desintegración doble beta sin neutrinos usando una cámara de proyección temporal (o TPC, por sus siglas en inglés) llena con 100 kg de xenón gaseoso (enriquecido al 91% en el isótopo $^{136}\text{Xe}$) a una presión de 15 bares. Dicho detector ofrece dos valiosas prestaciones para un experimento de desintegración $0\nu\beta\beta$: excelente resolución energética (cercana al 0.5% FWHM a 2.5 MeV) y trazado de partículas cargadas para la discriminación de señal y ruido. Además, la tecnología puede ser extrapolada sin demasiados problemas a la escala de la tonelada, permitiendo la exploración completa de la región correspondiente a la jerarquía invertida de masas de los neutrinos. El detector, conocido como NEXT-100, está actualmente en construcción, y su instalación y puesta en marcha en el Laboratorio Subterráneo de Canfranc (LSC), en España, se han previsto para el año 2017.

El gas xenón es un buen medio detector de radiación, con dos intensas señales primarias: centelleo e ionización. La primera se usa en NEXT para establecer el tiempo inicial de cada evento, mientras que la segunda se utiliza para el trazado de partículas y calorimetría. En su fase gaseosa, el xenón puede proporcionar excelente resolución energética, próxima, en principio, a 0.3% FHWM en $Q_{\beta\beta}$. La señal de ionización se amplifica en NEXT usando la electroluminescencia (EL) del xenón con el fin de optimizar la resolución energética del detector. El proceso de detección es el siguiente: Los electrones liberados por partículas ionizantes que atraviesan el gas derivan primero hacia el ánodo de la TPC bajo la acción de un campo eléctrico moderado (300–500 V cm$^{-1}$). Entran entonces en otra región del detector en la que son acelerados por un campo eléctrico de intensidad suficiente para que los electrones puedan excitar los átomos de xenón sin ionizarlos. La energía de excitación se libera en forma de luz de centelleo secundaria de intensidad proporcional al número de electrones de ionización. Un conjunto de tubos fotomultiplicadores (PMTs) situado detrás del cátodo de la TPC detecta una fracción de la luz de centelleo secundario, proporcionando una medida precisa de la energía total depositada en el gas. Estos PMTs detectan además el centelleo primario que permite establecer el tiempo inicial del evento. La luz de centelleo secundaria se detecta también usando un plano de fotomultiplicadores de silicio (SiPMs) de 1 mm$^2$ situado detrás del ánodo, muy cerca de la región de EL, y que permite la reconstrucción de las trazas de las partículas cargadas del evento.
Sensibilidad del detector NEXT-100

El experimento NEXT dispone de dos signaturas experimentales para la distinción de los eventos de señal y ruido:

- **Resolución energética**: Los eventos de señal forman un pico monocromático en el espectro de energías registrado por el detector. Por tanto, seleccionando sólo aquellos eventos en una región energética alrededor $Q_{\beta\beta}$ de tamaño proporcional a la resolución del detector se elimina la mayor parte de la actividad espuria.

- **Trazado de partículas cargadas**: Las trazas de los electrones en xenón a alta presión son tortuosas debido a la dispersión múltiple coulombiana, y tienen un patrón de deposición de energía (d$E$/dx) más o menos uniforme excepto al final, donde aumenta por un factor 4–5, dejando una gran deposición que llamamos blob. Esta signatura puede usarse para distinguir los eventos de señal (formados por dos trazas electrónicas emitidas desde un vértice común) y los eventos de ruido (fundamentalmente, electrones individuales generados en la interacción de rayos gamma de alta energía).

La relevancia de una fuente de ruido depende, por lo tanto, de su probabilidad de generar una traza similar a las de señal con energía cercana a $Q_{\beta\beta}$. El pico $0\nu\beta\beta$ del $^{136}$Xe se encuentra entre los picos fotoeléctricos de las gammas de alta energía emitidas tras las desintegraciones $\beta$ del $^{214}$Bi (isótopo de la serie de uranio) y del $^{208}$Tl (de la serie de torio). La principal fuente de ruido en el detector NEXT-100 son los fotosensores. La tasa de ruido esperada en el experimento es $6 \times 10^{-4}$ cuentas keV$^{-1}$ kg$^{-1}$ año$^{-1}$. NEXT-100 alcanzará una sensibilidad a la semivida de la desintegración $0\nu\beta\beta$ próxima a $4.6 \times 10^{25}$ años tras 3 años (efectivos) de operación.

Conclusiones

Los experimentos de búsqueda de la desintegración doble beta sin neutrinos son la única manera práctica de establecer si los neutrinos son partículas de tipo Majorana. La actual generación de experimentos, con una sensibilidad a $m_{\beta\beta}$ aproximada de 100 meV, ya ha entrado en funcionamiento, con resultados iniciales de EXO-200, KamLAND-Zen y GERDA. El detector NEXT-100, que
combina una muy buena resolución energética con la reconstrucción tipológica de los eventos, podría contribuir de manera decisiva a esta exploración, sirviendo, además, de demostración de una tecnología que podría ser clave para el futuro del campo.
I would like to thank the many individuals who have helped me through the completion of this thesis. First, I am grateful to my supervisor, J.J. Gómez Cadenas, for persuading me to study neutrinos and giving me the opportunity to participate in the NEXT experiment from its initial conception through to its development and (ongoing) realization. Juanjo has been a true mentor and friend, and both NEXT and this dissertation would not have been possible without his guidance, determination and enthusiasm.

Science is today a collaborative effort, and NEXT is no exception. I would like to acknowledge the members of the NEXT Collaboration; in particular, the nexters at the Instituto de Física Corpuscular (IFIC) and the Universidad Politécnica de Valencia. The progress of the experiment is owed to the skills and commitment of all of them.

A number of people who have been closer to me during these years deserve a special mention. The work of some of them had a direct impact on this thesis; others simply helped me stay sane (my apologies to those I may have inadvertently omitted): Michel Sorel has been a constant reference on neutrino physics and various other topics throughout these years. Pau Novella was my first collaborator and shared with me a few trips around Europe and beyond during our early days in the double beta decay community. At some point, Pau left and was substituted by Francesc Monrabal; he and Maria Cerdà were my travel companions in a couple of fun summers in Switzerland and another one in California. Javier Muñoz has been my closest collaborator all these years and his help has been invaluable; this thesis is partly his. Andrew Laing and Paola Ferrario took over many of my responsibilities in the experiment while I was writing this thesis; in addition, they were among the first to read the manuscript, providing useful comments. Moreover, both of them have been close friends outside IFIC — I want to thank Andrew for his noble nature and for being always ready for a beer, and I want to thank Paola for her sincerity
and for showing me many times the other side of the coin. During the last year, I have worked with David Lorca in the data analysis of NEXT-DEMO, and although we rarely agreed in anything, we ended up doing a good job, I think. Miquel Nebot is continuing part of the job I have done during my PhD; I thank him for his patience every time I have been too busy to discuss things. David and Miquel plus Luis Serra and Toni Mari have shared with me a noisy office at IFIC and have made much more fun my daily work. Lorena Escudero and Raphaël Schroeter were also occupants of the office at some point. Raphaël was among the first people I met in the research world, and since then he's been a loyal friend in spite of the distance. Lorena, in addition to her friendship, has given me, reluctantly, valuable advice. I would also like to thank José Pérez for his friendship and constant help with administrative tasks.

A few other close friends during this period were already with me during the undergraduate years: Alberto Martínez was my closest associate back then, and without him, graduating would have been much harder; Nacho Ruiz is the best public one's jokes can find; Dani Gosálbez shares with me a certain tendency to worry too much; and Miguel Villaplana, adoptive member of NEXT, has suffered patiently my ramblings and whining all these years. I am grateful for all of that.

Outside the world of research, I thank Diego Carrillo and Sara Pascual for their friendship and the shared kilometres.

Finally, I would like to dedicate this thesis to my parents, my sister and brother, and to the memory of my grandparents. Either by nature or nurture they are responsible for whatever I have become.

Valencia, January 2015
Contents

Preface i

Resumen v

Acknowledgments xi

1 The nature of massive neutrinos 1
  1.1 Current knowledge of neutrino mass and mixing 1
  1.2 The origin of neutrino masses 6
    1.2.1 The see-saw mechanism 9
  1.3 Dirac vs. Majorana 12
  1.4 Majorana neutrinos and the missing antimatter 14
  1.5 Lepton number violating processes 17

2 Neutrinoless double beta decay and Majorana neutrinos 21
  2.1 Double beta decay modes 21
  2.2 The black box theorem 24
  2.3 The standard mechanism 25
  2.4 Alternative mechanisms 27
  2.5 Nuclear matrix elements 29

3 The search for neutrinoless double beta decay 33
  3.1 Introduction 33
  3.2 Sensitivity of a double beta decay experiment 34
    3.2.1 Frequentist confidence intervals and upper limits 34
    3.2.2 Sensitivity to $m_{\beta\beta}$ 38
  3.3 History and status of double beta decay searches 40
  3.4 Ingredients for the ultimate double beta decay experiments 43
3.4.1 Choice of the isotope .................................................. 44
3.4.2 Energy resolution ...................................................... 45
3.4.3 Low background ...................................................... 48
3.4.4 Detection efficiency .................................................. 55
3.4.5 Exposure .............................................................. 55

3.5 Current generation of experiments ................................. 56
3.5.1 CUORE ............................................................... 56
3.5.2 EXO ................................................................. 58
3.5.3 GERDA and MAJORANA ........................................... 61
3.5.4 KamLAND-Zen ....................................................... 64
3.5.5 SNO+ ............................................................... 65
3.5.6 SuperNEMO .......................................................... 66

4 High-pressure xenon gas for double beta decay searches 69
4.1 Introduction ............................................................. 69
4.2 Gaseous xenon as detection medium ............................... 71
  4.2.1 Primary signals in xenon: scintillation and ionization ..... 71
  4.2.2 Detection of the ionization and scintillation signals ...... 76
  4.2.3 Intrinsic energy resolution in xenon ......................... 78
  4.2.4 Impact of amplification fluctuations on energy resolu-
            tion ............................................................. 80
4.3 The SOFT detector concept ........................................... 84

5 The NEXT experiment ..................................................... 89
5.1 Introduction .......................................................... 89
5.2 Results from the R&D phase ....................................... 90
  5.2.1 NEXT-DEMO ..................................................... 90
  5.2.2 NEXT-DBDM ..................................................... 98
5.3 The NEXT-100 detector .............................................. 101
  5.3.1 Electric-field cage ............................................. 102
  5.3.2 Energy plane .................................................... 103
  5.3.3 Tracking plane ................................................ 104
  5.3.4 Pressure vessel and inner copper shield .................. 105
  5.3.5 Electronics .................................................... 105
  5.3.6 Gas system ..................................................... 107
  5.3.7 External shield and supporting structures .............. 107
  5.3.8 NEW ............................................................ 109
6  NEXUS, the end-to-end detector simulation of NEXT  
6.1 Introduction ........................................... 113  
6.2 Overview of Geant4 and NEXUS ...................... 114  
  6.2.1 Event data flow .................................... 115  
  6.2.2 Fast simulation .................................... 117  
6.3 Detector geometries and materials .................. 118  
6.4 Event generators ...................................... 119  
6.5 Physics .................................................. 120  
  6.5.1 Physics lists ....................................... 120  
  6.5.2 Simulation of the primary signals ............... 121  
  6.5.3 Light transport and collection .................. 123  
6.6 Digitization ............................................ 128  
6.7 Simulation results ..................................... 130  
  6.7.1 Reflectivity of the NEXT-DEMO light tube .... 130  
  6.7.2 Position of the tracking plane of NEXT-DEMO ... 131  
  6.7.3 Energy resolution in NEXT-DEMO ............... 133  

7  Sensitivity of NEXT-100 to neutrinoless double beta decay  
7.1 Introduction ........................................... 135  
7.2 Sources of background in the NEXT experiment .... 137  
  7.2.1 High-energy gamma rays ......................... 137  
  7.2.2 Radon .............................................. 138  
  7.2.3 Muons and neutrons .............................. 141  
7.3 Signal detection efficiency and background rejection power  
  7.3.1 Simulation ......................................... 141  
  7.3.2 Voxelization and graph-search algorithms .... 142  
  7.3.3 Event selection .................................... 143  
7.4 Predicted background rate and sensitivity of NEXT-100 ...... 149  

8  Towards the tonne scale  
8.1 Reach of the current generation of experiments ....... 157  
8.2 Exploring the inverted-hierarchy region of neutrino masses 162  

9  Summary and conclusions .................................. 167  

Bibliography ................................................ 171
The nature of massive neutrinos

1.1 Current knowledge of neutrino mass and mixing

Neutrinos are the lightest known elementary fermions. They do not carry electric or colour charge, and can be observed only via weak interactions. Experimentally, we know that only three light — that is, of mass smaller than $m_Z/2$, where $m_Z$ is the mass of the $Z$ boson — active neutrino families exist [1]. Additionally, over the past fifteen years, several experiments [2–8], through the observation of neutrino oscillations [9–13], have established that neutrinos are massive particles and that the neutrino states participating in the weak interaction (flavour eigenstates) are different from the states controlling free-particle evolution (mass eigenstates). In other words, the three weak eigenstates $\nu_\ell$ ($\ell = e, \mu, \tau$) can be expressed as a linear combination of three mass eigenstates $\nu_i$ ($i = 1, 2, 3$) in the minimum joint description of all oscillation data:

$$\nu_\ell = \sum_i U^*_{\ell i} \nu_i, \quad (1.1)$$

where $U$ is a $3 \times 3$ unitary matrix different from the identity. This so-called neutrino mixing matrix is usually parametrized in terms of 3 Euler angles ($\theta_{12}, \theta_{13}, \theta_{23}$) and 3 phases ($\delta, \alpha_1, \alpha_2$):

$$U = \begin{pmatrix}
    c_{12} c_{13} & s_{12} c_{13} & s_{13} e^{-i\delta} \\
    -s_{12} c_{23} - c_{12} s_{23} s_{13} e^{i\delta} & c_{12} c_{23} - s_{12} s_{23} s_{13} e^{i\delta} & c_{23} c_{13} \\
    s_{12} s_{23} - c_{12} c_{23} s_{13} e^{i\delta} & -c_{12} s_{23} - s_{12} c_{23} s_{13} e^{i\delta} & s_{23} c_{13}
\end{pmatrix} \times \text{diag}\left(1, e^{i\alpha_1}, e^{i\alpha_2}\right), \quad (1.2)$$

where $c_{ij} \equiv \cos \theta_{ij}$ and $s_{ij} \equiv \sin \theta_{ij}$. The angles $\theta_{ij}$ can be taken without loss of generality to lie in the first quadrant, $\theta_{ij} \in [0, \pi/2]$, while the phases $\delta$ and $\alpha_{1,2}$ can take any value from 0 to $2\pi$. If massive neutrinos are Dirac particles (§1.3),
only the phase $\delta$ is physical, accounting for CP violation in the leptonic sector. If massive neutrinos are Majorana particles (§1.3), then the two additional Majorana phases $\alpha_{1,2}$ may also be different from zero.$^{1}$

Global analyses$^{[16-18]}$ of the existing neutrino oscillation data have determined with relatively good precision (at the few percent level) the value of the three mixing angles: $\theta_{12} \simeq 34^\circ$, $\theta_{23} \simeq 42^\circ$ and $\theta_{13} \simeq 9^\circ$. The value of the Dirac CP-violation phase, $\delta$, also potentially measurable, is currently unknown. Neutrino oscillation experiments can only measure the differences between the squared neutrino masses and not the absolute mass scale. Experiments using solar and reactor neutrinos have measured one mass difference, the so-called solar mass splitting: $\Delta m^2_{\text{sol}} \simeq 7.5 \times 10^{-5} \text{ eV}^2$. Atmospheric and accelerator-based oscillation experiments have measured another mass difference, the atmospheric mass splitting: $\Delta m^2_{\text{atm}} \simeq 2.5 \times 10^{-3} \text{ eV}^2 \gg \Delta m^2_{\text{sol}}$. These results cannot differentiate between two possibilities for the neutrino mass ordering, usually referred to as normal and inverted orderings. In the former, $\Delta m^2_{\text{sol}}$ is the difference between the squared masses of the two lightest mass states, while in the latter it corresponds to the difference between the two heaviest states. This is illustrated in Figure 1.1. Notice that the states have been numbered in such a way that $\Delta m^2_{\text{sol}} \equiv \Delta m^2_{12}$ is always greater than zero, whereas

$^{1}$The reader interested in the topic of the physics of massive neutrinos may consult, for instance, the book by Giunti and Kim (2007)$^{[14]}$ or the review paper by González-García and Maltoni (2008)$^{[15]}$. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.1}
\caption{Current knowledge of neutrino masses and mixings from neutrino oscillation experiments. Left and right panels show, respectively, the normal and inverted mass orderings. Neutrino masses increase from bottom to top. The electron, muon and tau flavour contents of each neutrino mass eigenstate is shown via the red, green and blue fractions, respectively.}
\end{figure}
1.1. CURRENT KNOWLEDGE OF NEUTRINO MASS AND MIXING

Table 1.1. Neutrino mixing parameters according to the most recent global oscillation analysis by González-García et al. [13]. Note that $\Delta m^2_{31} \equiv \Delta m^2_{32}$ for the normal ordering and $\Delta m^2_{31} \equiv \Delta m^2_{32}$ for the inverted one.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Normal ordering</th>
<th>Inverted ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>best fit ± 1σ</td>
<td>3σ range</td>
</tr>
<tr>
<td>$\theta_{12}$ (°)</td>
<td>$33.48^{+0.78}_{-0.75}$ [31.29, 35.91]</td>
<td>$33.48^{+0.78}_{-0.75}$ [31.29, 35.91]</td>
</tr>
<tr>
<td>$\theta_{23}$ (°)</td>
<td>$42.3^{+3.0}_{-1.6}$ [38.2, 53.3]</td>
<td>$49.5^{+1.5}_{-2.2}$ [38.6, 53.3]</td>
</tr>
<tr>
<td>$\theta_{13}$ (°)</td>
<td>$8.50^{+0.20}_{-0.21}$ [7.85, 9.10]</td>
<td>$8.51^{+0.20}_{-0.21}$ [7.87, 9.11]</td>
</tr>
<tr>
<td>$\delta_{CP}$ (°)</td>
<td>$306^{+39}_{-70}$ [0, 360]</td>
<td>$254^{+63}_{-62}$ [0, 360]</td>
</tr>
<tr>
<td>$\Delta m^2_{21}$ (10$^{-5}$ eV$^2$)</td>
<td>$7.50^{+0.19}_{-0.17}$ [7.02, 8.09]</td>
<td>$7.50^{+0.19}_{-0.17}$ [7.02, 8.09]</td>
</tr>
<tr>
<td>$\Delta m^2_{31}$ (10$^{-3}$ eV$^2$)</td>
<td>$2.457^{+0.047}_{-0.047}$ [2.317, 2.607]</td>
<td>$-2.449^{+0.048}_{-0.047}$ [−2.590, −2.307]</td>
</tr>
</tbody>
</table>

$\Delta m^2_{\text{atm}}$ is a positive quantity in the normal ordering ($\Delta m^2_{\text{atm}} \equiv \Delta m^2_{31} > 0$) and a negative one in the inverted ordering ($\Delta m^2_{\text{atm}} \equiv \Delta m^2_{32} < 0$). Future oscillation experiments may allow us to establish the neutrino mass ordering thanks to the exploitation of matter effects or the interference between the two oscillation frequencies [22].

Our present knowledge of neutrino masses and mixings provided by neutrino oscillation data is summarized in Table 1.1. Besides the CP-violation phase and the neutrino mass ordering, another piece of information remains to be known: the absolute value of the lightest neutrino mass. This can be probed via cosmological observations, neutrinoless double beta decay searches and beta decay experiments. Only upper bounds on the neutrino mass scale, of the order of 1 eV, currently exist. Constraints on the lightest neutrino mass coming from neutrinoless double beta decay will be discussed in §2.3. In the following we briefly review cosmological and beta-decay constraints.

Cosmological observations can probe the sum of the neutrino masses,

$$\sum m_\nu \equiv m_1 + m_2 + m_3,$$

(1.3)

due to the impact of these on the rate of expansion of the universe and the growth of perturbations [23]. Currently-available data is still compatible with massless neutrinos, with the tightest upper limits on $\sum m_\nu$ being around 0.2 eV. For instance, a combination of cosmic microwave background (CMB) datasets — from the Planck [24] and WMAP [25] satellites, and the ACT [26] and
SPT\cite{27} telescopes— with baryon acoustic oscillation (BAO) measurements results in the limit\cite{28}

\[
\sum m_\nu < 0.23 \text{ eV (95\% CL)}.
\]  \hfill (1.4)

Since $\sum m_\nu$ must be greater than approximately 0.06 eV in the normal hierarchy and 0.1 eV in the degenerate hierarchy, the allowed neutrino mass window is already quite tight and could be closed by forthcoming observations. The bounds, nevertheless, are rather model dependent and vary strongly with the data combination adopted\cite{28,30}.

The relationship between $\sum m_\nu$ and the lightest neutrino mass $m_{\text{light}}$ ($m_1$ in the case of normal ordering, or $m_3$ in the case of inverted) is shown in Figure\hbox{1.2} (top panel). The red and green bands correspond to the normal and inverted orderings. The width of the bands is given by the $3\sigma$ ranges in the mass oscillation parameters $\Delta m^2_{\text{sol}}$ and $\Delta m^2_{\text{atm}}$ shown in Table\hbox{1.1} The horizontal blue band in the graph is the upper limit on $\sum m_\nu$ quoted above. In the quasi-degenerate regime of neutrino masses corresponding to that upper bound, $m_{\text{light}} \approx \sum m_\nu / 3 \leq 0.07 \text{ eV at 95\% CL}$, as shown by the vertical band in the graph.

The neutrino mass scale can also be probed in laboratory-based experiments. The differential electron energy spectrum in nuclear $\beta$-decay experiments is affected both by the neutrino masses and by the mixings defining the electron neutrino state in terms of mass eigenstates. In this case, the mass combination probed is given by

\[
m_\beta^2 \equiv \sum_{i=1}^3 |U_{ei}|^2 m_i^2.
\]  \hfill (1.5)

The relationship between $m_\beta$ in the above equation and $m_{\text{light}}$ is shown in Figure\hbox{1.2} (bottom). Again, the values of Table\hbox{1.1} are used to determine the $3\sigma$ bands for both the normal and inverted orderings. The most sensitive searches conducted so far are based upon the beta decay of tritium, $^3\text{H} \rightarrow ^3\text{He}^+ + e^- + \bar{\nu}_e$, but, as for cosmology, $\beta$-decay searches of neutrino mass have so far yielded negative results. The horizontal band in the bottom graph of Figure\hbox{1.2} comes from the two current best limits, from the Troitsk\hbox{32} and Mainz\hbox{33} experiments. The combined limit is\hbox{31}

\[
m_\beta < 2 \text{ eV (95\% CL)}.
\]  \hfill (1.6)
Figure 1.2. Constraints on the lightest neutrino mass $m_{\text{light}}$ coming from cosmology (top) and $\beta$-decay experiments (bottom). The red and green bands correspond to the allowed regions of the normal and inverted orderings, respectively. The $\Sigma m_\nu$ upper bound, by the Planck Collaboration [28], translates into an upper limit on $m_{\text{light}}$ shown via the vertical band in the same panel. This cosmological constraint on $m_{\text{light}}$ is also shown in the bottom graph, together with the upper limit on $m_\beta$ from tritium $\beta$-decay experiments [31].
The resulting constraint on $m_{\text{light}}$ is thus less stringent than the cosmological one. The KATRIN experiment [34] should be able to improve the $m_\beta$ sensitivity by roughly an order of magnitude in the forthcoming years thanks to its better statistics, energy resolution and background rejection.

1.2 The origin of neutrino masses

In the Standard Model (SM) of particle physics, fermion masses arise from the Yukawa couplings of the fermion fields with the Higgs doublet. For instance, in the case of the charged leptons, the interaction is described by the following Lagrangian [14]:

$$-\mathcal{L}_{\text{Yukawa}} = \sum_{\alpha,\beta=e,\mu,\tau} Y^\ell_{\alpha\beta} \overline{L^\alpha} \Phi \ell^\beta_R + \text{H.c.} \quad (1.7)$$

Here,

$$\Phi \equiv \begin{pmatrix} \Phi^{(+)} \\ \Phi^{(0)} \end{pmatrix}, \quad L_L \equiv \begin{pmatrix} \nu^L \\ \ell^L \end{pmatrix},$$

and $\ell_R$ (with $\ell = e, \mu, \tau$) are, respectively, the Higgs doublet and the leptonic weak-isospin doublets and singlets; and $Y^\ell$ is a complex $3 \times 3$ matrix of Yukawa couplings. The subscripts $L$ (for left-handed) and $R$ (for right-handed) denote, respectively, the negative and positive chirality components of a fermion field $\Psi$, which satisfy the relation $\Psi \equiv \Psi_L + \Psi_R = P_L \Psi + P_R \Psi$, with $P_L \equiv (1 - y_5)/2$ and $P_R \equiv (1 + y_5)/2$.

The Higgs fields have a non-zero vacuum expectation value (VEV) responsible for spontaneous electroweak symmetry breaking. In the so-called unitary gauge, the Lagrangian in Eq. (1.7) can be rewritten as follows [14]:

$$-\mathcal{L}_{\text{Yukawa}} = \left( \frac{v + H}{\sqrt{2}} \right) \sum_{\alpha=e,\mu,\tau} y^\ell_\alpha \overline{\ell^\alpha} \ell^\alpha_R + \text{H.c.}$$

$$= \sum_{\alpha=e,\mu,\tau} \frac{y^\ell_\alpha v}{\sqrt{2}} \overline{\ell^\alpha} \ell^\alpha + \sum_{\alpha=e,\mu,\tau} \frac{y^\ell_\alpha}{\sqrt{2}} \overline{\ell^\alpha} \ell^\alpha H, \quad (1.9)$$

where $y^\ell_\alpha$ are dimensionless Yukawa couplings, $v$ is the VEV of the Higgs doublet and $H$ is the Higgs boson field. The first term on the right-hand side of the equation is the mass term for the charged leptons, whose masses can then be defined as

$$m_\alpha \equiv \frac{y^\ell_\alpha v}{\sqrt{2}}. \quad (1.10)$$
The coefficients $y^e_\ell$, $y^\mu_\ell$, and $y^\tau_\ell$ are unknown parameters of the SM, and thus the masses of the charged leptons are not predicted by the theory and must be obtained from experimental measurements.

In direct analogy to the mass term described above, we can construct one for neutrinos, usually referred to as *Dirac mass term*, by introducing into the model the positive-chirality components $\nu_R$ of the neutrino fields, which are called *sterile* because they are singlets of the full SM gauge group (i.e. they do not participate in any of the fundamental interactions except gravity). The Lagrangian in Eq. (1.7) is extended in the following way:

$$\mathcal{L}_{\text{Yukawa}} = \sum_{\alpha, \beta = e, \mu, \tau} Y^\ell_{\alpha \beta} \bar{L}_\alpha \Phi \, \ell^\beta_R + \sum_{\alpha, \beta = e, \mu, \tau} Y^\nu_{\alpha \beta} \bar{L}_\alpha \tilde{\Phi} \, \nu^\beta_R + \text{H.c.}, \quad (1.11)$$

where $\tilde{\Phi} = i \sigma_2 \Phi^*$ (here $\sigma_2$ is the second Pauli matrix). The term for neutrinos analogous to Eq. (1.9) obtained diagonalizing the Yukawa Lagrangian reads

$$\mathcal{L}^D_{\text{mass}} = \sum_{i=1}^{3} \frac{y^\nu_i v_i}{\sqrt{2}} \bar{\nu}_i L \, \nu^i_R + \text{H.c.} = \sum_{i=1}^{3} m_D \nu^i_R,$$

with $v \equiv v_L + v_R$ and $m_D \equiv y^\nu v / \sqrt{2}$. The neutrino Yukawa couplings in Eq. (1.12) must be of the order of $10^{-12}$ or smaller to account for the mass scale suggested by neutrino oscillations (§1.1). The current theoretical prejudice is that such infinitesimal couplings — at least 6 orders of magnitude smaller than the coupling associated to the electron, as illustrated in Figure 1.3 — are un plausible, and hence the explanation of neutrino masses with a Dirac mass term alone is seen as unsatisfactory, if not unnatural [14, 15, 35, 36].

There is a second way in which a neutrino mass term can be added to the SM Lagrangian. Ettore Majorana was the first to realize that for neutral particles one can remove two of the four degrees of freedom in a massive spinor field by imposing the following condition [37]:

$$\nu^c = \nu,$$  \hspace{1cm} (1.13)

where $\nu^c$ is the charge-conjugate of $\nu$. This condition implies that there is only one field to describe neutrino and antineutrino states. Decomposing both sides of Eq. (1.13) into their chiral components, it can be shown that

$$\nu_R = (\nu_L)^c,$$  \hspace{1cm} (1.14)
Figure 1.3. Hierarchical structure of fermion masses. Values taken from the Review of Particle Physics [31]. The figure assumes a normal ordering for neutrino masses.

showing that the positive-chirality component of the Majorana neutrino field $\nu_R$ is not independent of, but obtained from its negative-chirality counterpart $\nu_L$. By substituting Eq. (1.14) into Eq. (1.12), we obtain a Majorana mass term:

$$- \mathcal{L}^{M,L}_{\text{mass}} = \frac{1}{2} m_L (\nu_L)^c \nu_L + \text{H.c.},$$

where $m_L$ is a free parameter with dimensions of mass. This equation represents a mass term constructed from negative-chirality neutrino fields alone. If positive-chirality fields also exist and are independent of the negative-chirality ones, we may also construct a second Majorana mass term with them:

$$- \mathcal{L}^{M,R}_{\text{mass}} = \frac{1}{2} m_R (\nu_R)^c \nu_R + \text{H.c.}$$

Since the $\nu_R$ fields are SM weak-isospin singlets, the mass parameter $m_R$, in contrast with $m_D$ or $m_L$, is not connected to a Higgs VEV and could be arbitrarily high.

As depicted in Figure 1.4, all three mass terms in Eqs. (1.12), (1.15) and (1.16) convert negative-chirality states into positive-chirality ones (the charge conjugate of a field with a given chirality always has the opposite chirality). Therefore, chirality is not a conserved quantity for massive particles, regardless of the Dirac or Majorana nature of neutrino masses. The Majorana mass terms in Eqs. (1.15) and (1.16) also convert particles into their own antiparticles, thus
1.2. THE ORIGIN OF NEUTRINO MASSES

Figure 1.4. The effect of Dirac and Majorana mass terms: Both types of mass terms turn chiral states into their opposite (the charge conjugate of a field with a given chirality always has the opposite chirality). In addition, Majorana mass terms convert particles into their antiparticles.

violating the SM total lepton number $L \equiv L_e + L_\mu + L_\tau$ by two units ($|\Delta L| = 2$). Such mass terms are therefore forbidden for all electrically charged fermions because of charge conservation.

1.2.1 The see-saw mechanism

We have seen above that in order to generate a Dirac mass for the neutrino as we do for the other fermions, we are forced to introduce in our theory the sterile neutrino field $\nu_R$. Notice that once this field exists, there is no strong reason in the SM to prevent the occurrence of a Majorana mass term like that in Eq. (1.16): such a term violates neither the conservation of weak isospin nor that of electric charge. Consequently, if nature contains a Dirac neutrino mass term, then it is very likely that it will also contain a Majorana mass term, and, of course, if no Dirac neutrino mass term exists, then, certainly, the only source of neutrino mass would be a Majorana mass term.

Therefore, the most general neutrino mass term allowed in the framework of the SM with the minimal addition of the positive-chirality field $\nu_R$ is the following:

$$L_{\text{mass}}^{D+M} = L_{\text{mass}}^D + L_{\text{mass}}^{M,R}.$$  \hspace{1cm} (1.17)

In contrast, the Majorana mass term in Eq. (1.15) is not allowed by the symmetries of the SM, but it could be generated by new physics. In order to understand the implications of the combined Dirac-Majorana mass term in
the above equation, it is useful to recast it in the following matrix form:

$$- \mathcal{L}^{D+M}_{\text{mass}} = \frac{1}{2} \left( \mathcal{N}_L \right)^c M \mathcal{N}_L + \text{H.c.},$$

(1.18)

where $\mathcal{N}_L$ is the vector of negative-chirality fields,

$$\mathcal{N}_L = \left( \begin{array}{c} \nu_L \\ (\nu_R)^c \end{array} \right),$$

(1.19)

and the matrix $M$ has the form

$$M = \left( \begin{array}{cc} 0 & m_D \\ m_D & m_R \end{array} \right).$$

(1.20)

The chiral fields $\nu_L$ and $\nu_R$ do not have a definite mass because of the off-diagonal Dirac mass elements. In order to find the fields with definite masses, it is necessary to diagonalize the mass matrix $M$ by finding a unitary mixing matrix $U$ such that

$$U^T M U = \left( \begin{array}{cc} m_1 & 0 \\ 0 & m_2 \end{array} \right),$$

(1.21)

where

$$\mathcal{N}_L = U n_L \quad \text{with} \quad n_L = \left( \begin{array}{c} \nu_L \\ N_L \end{array} \right).$$

(1.22)

That is, we obtain two fields of definite mass and chirality. The Dirac-Majorana term can then be rewritten as

$$- \mathcal{L}^{D+M}_{\text{mass}} = \frac{1}{2} \left( m_1 (\nu_L)^c \nu_L + m_2 \left( N_L \right)^c N_L \right) + \text{H.c.}$$

(1.23)

The two terms in the above Lagrangian have the same form as the pure negative-chirality Majorana mass term in Eq. (1.15). In other words, both mass eigenfields $\nu_L$ and $N_L$ are equal to their CP-conjugate fields, and thus both describe Majorana particles. The insertion of a Dirac mass term and a positive-chirality Majorana mass term in the Lagrangian for massive neutrinos has resulted in Majorana particles.

We have mentioned already that because the positive-chirality fields, $\nu_R$, are electroweak singlets in the Standard Model, the Majorana mass of the neutrino described by such fields, $m_R$, may be orders of magnitude larger than the electroweak scale. If we assume as well that the neutrino Yukawa couplings
are of the order of those of the charged leptons, then \( m_R \gg m_D \). In this limit, it can be proved \cite{14} that
\[
m_1 \simeq \frac{m_D^2}{m_R} \tag{1.24}
\]
and
\[
m_2 \simeq m_R \tag{1.25}
\]
That is, one of the states is as heavy as \( m_R \) and the other is very light because its mass is suppressed with respect to \( m_D \) by the small ratio \( m_D/m_R \). This is the so-called see-saw mechanism \cite{38,41}: a very heavy neutrino state \( N \) is connected to (and explains) a very light \( \nu \).

The see-saw mechanism presented above can easily be generalized from the one-family case that we have discussed to three neutrino species, yielding the three light neutrinos \( \nu_i \) we are familiar with, and three heavy neutrinos \( N_i \) \cite{14}. In this case, the neutrino mass matrix in Eq. (1.20) is a \( 6 \times 6 \) mass matrix of the form
\[
M = \begin{pmatrix}
0 & (M_D^T) \times \\
M_D & M_R
\end{pmatrix}, \tag{1.26}
\]
where \( M_D \) and \( M_R \) are now \( 3 \times 3 \) complex matrices. The 6-component vector of negative-chirality fields now has the following form:
\[
N_L = \begin{pmatrix}
\nu_L \\
(\nu_R)^c
\end{pmatrix}, \text{ with } \nu_L = \begin{pmatrix}
\nu_e L \\
\nu_\mu L \\
\nu_\tau L
\end{pmatrix} \text{ and } (\nu_R)^c = \begin{pmatrix}
(\nu_{s1} R)^c \\
(\nu_{s2} R)^c \\
(\nu_{s3} R)^c
\end{pmatrix}. \tag{1.27}
\]
Here, the subscripts \( e, \mu \) and \( \tau \) label the active neutrino flavours, whereas the subscripts \( s_1, s_2 \) and \( s_3 \) indicate sterile states that do not participate in the weak interactions. The mass matrix is diagonalized via a \( 6 \times 6 \) mixing matrix analogous to \( U \) in Eq. (1.21), where the three negative-chirality fields and the three positive-chirality fields are now expressed in terms of the negative-chirality components of 6 massive neutrino fields \( \nu_{iL} \).

In the see-saw limit where the eigenvalues of \( M_R \) are much larger than those of \( M_D \), the \( 6 \times 6 \) mass matrix in Eq. (1.26) can be written in the block-diagonal form
\[
M \simeq \text{diag}(M_{\text{light}}, M_{\text{heavy}}),
\]
where the two \( 3 \times 3 \) mass matrices of the light and heavy neutrino sectors are practically decoupled, and given by
\[
M_{\text{light}} \simeq -(M_D^T(M_R)^{-1})M_D \quad \text{and} \quad M_{\text{heavy}} \simeq M_R^R,
\]
respectively. For the low-energy phenomenology, it is sufficient to consider only \( M_{\text{light}} \), sometimes
called the *neutrino mass matrix* $m_{\nu}$, that is the $3 \times 3$ matrix in the flavour basis which is diagonalized by the matrix $U$:

$$U^T M_{\text{light}} U = \text{diag}(m_1, m_2, m_3),$$

where the *neutrino mixing matrix* $U$ is the same matrix defined in Eq. (1.1), and $m_1$, $m_2$ and $m_3$ are the three light neutrino mass eigenvalues discussed in §1.1.

An important assumption in the simplest realization of the see-saw mechanism described above, known as *type I see-saw*, is that $m_L = 0$. This assumption is not arbitrary, but follows directly from enforcing the gauge symmetries of the Standard Model. In models with a left-right symmetric particle content, the see-saw mechanism is often generalized to a *type II see-saw*, where an additional direct mass term $m_L$ for the light neutrinos is present \[14\].

### 1.3 Dirac vs. Majorana

Which of the neutrino mass terms allowed in theory do exist in nature? This question could in principle be answered experimentally since Dirac and Majorana massive neutrinos should have, in fact, different Standard Model interactions. Let us consider, for example, a scattering experiment in which a muon-neutrino beam — generated from the decay of positive pions, like in the classic experiment by Lederman, Schwartz and Steinberger \[42\] — is sent in the direction of a large magnetized detector. Neutrinos can interact with the nucleons of the detector material producing muons whose charge and energy are then measured.

Table 1.2 lists all possible neutrino states in our experiment. Note that chirality is not a good quantum number for massive particles, and therefore cannot be used to describe our muon neutrinos. Instead, we will use their *helicity*. The helicity operator $\hat{h}$ is defined as the projection of the particle’s spin $\vec{s}$ onto the direction of its momentum $\vec{p}$:

$$\hat{h} \equiv \frac{\vec{s} \cdot \vec{p}}{|\vec{p}|}.$$

It has eigenvalues $h = \pm 1/2$ for all elementary fermions, and hence we speak of particles with positive and negative helicity. For massless particles, chirality and helicity are equivalent. In the case of massive particles, a given helicity state is a superposition of the negative and positive chirality states. Neutrino masses
Table 1.2. The difference between Dirac (top four states) and Majorana (last two states) neutrinos in a scattering experiment: Dirac neutrinos (as opposed to antineutrinos) only produce negatively charged leptons ($\ell^-$), conserving the total lepton number $L$; Majorana neutrinos, in contrast, can generate both positive and negative leptons, although the production rate for the wrong charge is exceptionally small, $\mathcal{O}(m_\nu/E)$.

<table>
<thead>
<tr>
<th>helicity</th>
<th>$L$</th>
<th>$\ell^-$</th>
<th>$\ell^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu$</td>
<td>+1</td>
<td>$\approx 1$</td>
<td>0</td>
</tr>
<tr>
<td>$\nu$</td>
<td>+1</td>
<td>$(m_\nu/E)^2$</td>
<td>0</td>
</tr>
<tr>
<td>$\bar{\nu}$</td>
<td>−1</td>
<td>0</td>
<td>$(m_\nu/E)^2$</td>
</tr>
<tr>
<td>$\bar{\nu}$</td>
<td>−1</td>
<td>0</td>
<td>$\approx 1$</td>
</tr>
</tbody>
</table>

are so small that, for instance, a negative helicity neutrino can be considered mostly a negative chirality state with only a small contribution of order $m_\nu/E$ (where $E$ is the energy of the neutrino) from the positive chirality state.

In the case of Dirac neutrinos (top four rows of Table 1.2), there are four mass-degenerate states: for each of the two available helicity states, two distinct particle/antiparticle states characterized by a different lepton number $L$ are available, with $L(\nu) = L(\ell^-) = -L(\bar{\nu}) = -L(\ell^+) = 1$, where $\ell^\pm$ indicate charged leptons. Total lepton number is conserved in SM interactions involving Dirac neutrinos. These can then be identified as neutrinos or antineutrinos according to the charge of the leptons produced in their interactions. However, neutrino states of positive helicity and antineutrino states of negative helicity would have much weaker lepton-producing interactions with matter compared to their helicity-reversed companions — as indicated by the coefficients in Table 1.2 — since they are predominantly the chiral states that do not participate in the charged-current weak interaction.

Majorana neutrinos, on the other hand, do not conserve total lepton number, and so the Dirac states with equal helicity become indistinguishable. Now there are only two independent neutrino states (the last two rows of Table 1.2):
one with negative helicity (the one that we usually call $\text{neutrino}$) that produces mostly negatively-charged leptons, and another one (the $\text{antineutrino}$) with positive helicity that produces predominantly positively-charged leptons.

Consequently, the detection of wrong-sign muons ($\mu^+$ in the case of a muon neutrino beam, as shown in Figure 1.5) in our scattering experiment would be a clear signature of Majorana neutrinos. Unfortunately, as we have mentioned already, the production rate of wrong-sign muons is suppressed by $m_\nu/E$ in amplitude, making it virtually impossible to detect. For example, for $m_\nu \sim 1 \text{ eV}$ and $E \sim 1 \text{ GeV}$, the cross-section for the process would be, roughly, $10^{-18}$ times the usual charged-current neutrino cross-section [36]. As we will discuss in Chapter 2, the best hope of establishing the neutrino nature seems to be the search for neutrinoless double beta decay.

1.4 Majorana neutrinos and the missing antimatter

Inflationary models of the universe predict that matter and antimatter had to be equally abundant at the very hot beginning, given that any potential initial asymmetry would have been diluted by inflation [43]. However, the observable universe today is almost entirely made of matter: no traces of cosmological antimatter have been observed to date and all the structures that we see — stars, galaxies, clusters — consist mostly of matter. This dominance of matter over antimatter is consistent with the small level of baryon asymmetry that is inferred from Big Bang nucleosynthesis (BBN) and cosmic microwave background (CMB) observations (see the Review of Particle Physics [31] and
1.4. MAJORANA NEUTRINOS AND THE MISSING ANTIMATTER

where \( n_B, n_{\bar{B}} \) and \( n_\gamma \) are, respectively, the number densities of baryons, anti-baryons and photons. Had the universe been perfectly symmetric, the annihilations of baryons and anti-baryons into photons (which started once the temperature went well below the nucleon mass) would have made baryons about 8 orders of magnitude less abundant than what we observe today [44]. In conclusion, the asymmetry between matter and antimatter was not an initial condition and must have been generated dynamically in the early universe through processes which are known generically as \textit{baryogenesis}.

In 1967, Andrei Sakharov proposed [45] three necessary conditions that any baryon-generating process must satisfy to produce matter and antimatter at different rates:

1. Violation of baryon number \( B \).
2. Violation of the C and CP symmetries.
3. Departure from thermal equilibrium.

All these ingredients are present in the Standard Model; however, the amount of CP violation observed in the quark sector does not appear to be sufficient to quantitatively account for the observed level of baryon asymmetry. Therefore, several mechanisms for baryogenesis involving new physics beyond the SM have been proposed in the past decades [46]. In particular, baryogenesis could have been induced by a lepton asymmetry, known as \textit{leptogenesis} [43,47], possible if neutrinos are Majorana particles. The decay of the heavy Majorana neutrinos predicted by the see-saw mechanism into leptons (\( \ell_a \)) plus Higgs particles (\( \phi \)) can provide all of the conditions enumerated above:

1. Heavy Majorana neutrinos are their own anti-particles, and as a result they can decay to both \( \ell_a \phi \) and \( \bar{\ell}_a \bar{\phi} \) final states. If there is an asymmetry in the two decay rates, a net lepton asymmetry will be produced, which via the so-called \textit{sphaleron processes} [43] can be efficiently converted into a baryon asymmetry. Figure 1.6 shows the processes that would contribute to lepton asymmetry in the presence of heavy Majorana

\[
\text{CMB: } \eta \equiv \frac{n_B - n_{\bar{B}}}{n_\gamma} = (6.05 \pm 0.07) \times 10^{-10}, \\
\text{BBN: } 5.7 \times 10^{-10} \leq \eta \leq 6.7 \times 10^{-10} \text{ (95% CL)},
\]

\[
\frac{\text{six.fitted}}{\text{zero.fitted}} + \frac{\text{five.fitted}}{\text{zero.fitted}} \pm \frac{\text{zero.fitted}}{\text{seven.fitted}} \times \frac{\text{one.fitted}}{\text{zero.fitted}} - \frac{\text{one.fitted}}{\text{zero.fitted}} \leq \eta \leq \frac{\text{six.fitted}}{\text{seven.fitted}} \times \frac{\text{one.fitted}}{\text{zero.fitted}} - \frac{\text{one.fitted}}{\text{zero.fitted}} \left( \frac{\text{nine.fitted}}{\text{five.fitted}} \text{CL} \right)
\]
neutrino decays, in the simplest case where the asymmetry is dominated by the decay of the lightest among the three heavy neutrinos, $N_1$.

2. C and CP can be violated in these decays provided that there is more than one heavy Majorana field.

3. Departure from thermal equilibrium is obtained if the decay rate is slower than the expansion rate of the universe at the time of decoupling, occurring for $T \sim M_1$, where $T$ is the temperature of the universe’s thermal bath and $M_1$ is the mass of the lightest among the three heavy neutrinos.

In order to be fully successful, any theory of baryogenesis must be able to explain the observed magnitude of baryon asymmetry given in Eqs. (1.30) and (1.31). Leptogenesis via heavy Majorana neutrino decays is in principle able to do this. In this case, the asymmetry in lepton flavour $\alpha$ produced in the decay of $N_1$, defined as

$$
\varepsilon_{\alpha\alpha} \equiv \frac{\Gamma(N_1 \rightarrow \phi \ell_\alpha) - \Gamma(N_1 \rightarrow \phi \ell_\alpha)}{\Gamma(N_1 \rightarrow \phi \ell) + \Gamma(N_1 \rightarrow \phi \ell)},
$$

should be of order $|\varepsilon_{\alpha\alpha}| > 10^{-7}$ [43], where the factors $\Gamma$ in Eq. (1.32) stand for the decay rates into the corresponding $N_1$ decay final states. In general, there is no direct connection between the high-energy CP-violating processes responsible for the asymmetry in the early universe of Eq. (1.32), and the low-energy CP-violating processes that may potentially affect laboratory-based experiments. Nonetheless, the discovery of CP violation in the lepton sector via neutrino oscillations on the one hand and the establishment of a Majorana
nature for neutrinos on the other would undoubtedly strengthen the case for leptogenesis as a source of the baryon asymmetry of the universe.

1.5 Lepton number violating processes

We have seen that a Majorana mass term induces lepton number violating processes of the type $|\Delta L| = 2$. The heavy neutrino decay needed for leptogenesis requires Majorana mass terms, and is, therefore, an example of a lepton number violating process. However, heavy neutrino decay is unobservable in a laboratory-based experiment, given the tremendous energies needed for heavy neutrino production. A number of more promising lepton number violating processes have been proposed to probe the Majorana nature of neutrinos. The best-known example is neutrinoless double beta decay ($0\nu\beta\beta$), which will be discussed in Chapter 2. However, and because of neutrino mixing, the phenomenology associated with $|\Delta L| = 2$ processes is very rich. A basic process with $|\Delta L| = 2$ is mediated by

$$W^- W^- \rightarrow \ell^-_\alpha \ell^-_\beta,$$  \hspace{1cm} (1.33)

and we can categorize such processes according to the lepton flavours ($\alpha, \beta$) involved. Assuming no lepton flavour violating contributions other than light Majorana neutrino exchange, the matrix element for the generic $|\Delta L| = 2$ process in Eq. (1.33) is proportional to the element ($\alpha, \beta$) of the neutrino mass matrix:

$$(m_\nu)_{\alpha\beta} \equiv (U^* \operatorname{diag}(m_1, m_2, m_3) U^\dagger)_{\alpha\beta} = \sum_{i=1}^{3} U^*_\alpha i U^*_\beta i m_i,$$  \hspace{1cm} (1.34)

where $m_\nu = M^{\text{light}}$ is the matrix appearing in Eq. (1.28), $U_{\alpha i}$ are the elements of the neutrino mixing matrix and $m_i$ are the three light neutrino mass eigenstates. In a sense, this effective neutrino mass definition provides a metric to compare the sensitivity of various $|\Delta L| = 2$ processes.

The processes with the most competitive constraints on $|\Delta L| = 2$ processes involving the flavours ($\alpha, \beta$) are reported in Table 1.5. The constraint on the effective Majorana mass $m_{ee}$ coming from $0\nu\beta\beta$ searches outperforms other searches involving a different flavour combination ($\alpha, \beta$) by several orders of magnitude. The most important reason behind this is of statistical nature. While it is possible to amass macroscopic quantities of a $\beta\beta$ emitter to study $0\nu\beta\beta$ decay, this is not the case for the other experimental techniques listed in
Nevertheless, it is important to keep exploring other lepton flavour violating processes for two reasons. First, it is, in principle, possible that phase cancellations are such that $m_{ee} \ll m_{\alpha\beta}$ with $(\alpha, \beta) \neq (e, e)$, making the search for $0\nu\beta\beta$ decay much less favourable than others because of the particular values of the neutrino masses and mixing parameters. Second, this effort may possibly lead to the identification of an even most promising experimental probe of lepton flavour violation in the future.
Table 1.3. Current bounds on effective neutrino masses from total lepton number violating processes, organized according to the flavours involved. Numbers taken (or derived) from Atre, Barger and Han (2005) [48] and the Review of Particle Physics [31].

<table>
<thead>
<tr>
<th>Flavours</th>
<th>Exp. technique</th>
<th>Experimental bound</th>
<th>Mass bound (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e, e)</td>
<td>0νββ</td>
<td>$T_{1/2}^{^{136}Xe} &gt; 1.6 \times 10^{25}$ yr</td>
<td>$</td>
</tr>
<tr>
<td>(e, μ)</td>
<td>$μ^- \rightarrow e^+$ conversion</td>
<td>$\frac{Γ(Ti + μ^- \rightarrow e^+ + C_{ags})}{Γ(Ti + μ^- \text{ capture}) &lt; 1.7 \times 10^{-12}}$</td>
<td>$</td>
</tr>
<tr>
<td>(e, τ)</td>
<td>Rare τ decays</td>
<td>$Γ(τ^- \rightarrow e^+ π^- π^-)/Γ_{tot} &lt; 8.8 \times 10^{-8}$</td>
<td>$</td>
</tr>
<tr>
<td>(μ, μ)</td>
<td>Rare kaon decays</td>
<td>$Γ(K^+ → π^- μ^+ μ^+)/Γ_{tot} &lt; 1.1 \times 10^{-9}$</td>
<td>$</td>
</tr>
<tr>
<td>(μ, τ)</td>
<td>Rare τ decays</td>
<td>$Γ(τ^- → μ^+ π^- π^-)/Γ_{tot} &lt; 3.7 \times 10^{-8}$</td>
<td>$</td>
</tr>
<tr>
<td>(τ, τ)</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>
Neutrinoless double beta decay & Majorana neutrinos

2.1 Double beta decay modes

Double beta ($\beta\beta$) decay is a second-order weak process that transforms a nuclide of atomic number $Z$ into its isobar with atomic number $Z + 2$. The ordinary decay mode consisting in two simultaneous beta decays ($2\nu\beta\beta$),

$$ (Z, A) \rightarrow (Z + 2, A) + 2 e^- + 2 \nu_e, \quad (2.1) $$

was first considered by Maria Goeppert-Mayer, in 1935 [49]. There has been geochemical evidence of its existence since the 1950s [50], but the first direct observation, in $^{82}\text{Se}$ and using a time projection chamber as detector, was not made until 1987 [51]. Since then, it has been repeatedly observed in several nuclides with typical lifetimes of the order of $10^{18}$–$10^{21}$ years (see Table 3.1), the longest ever measured among radioactive decay processes. With such long half-lives, for $2\nu\beta\beta$ to be a competitive decay mode, the $\beta$ decay to the $Z + 1$ nuclide must be either energetically forbidden or highly suppressed. Such a condition is fulfilled by 35 naturally-occurring isotopes thanks to the nuclear pairing force, which ensures that even-even nuclides are more bound than their odd-odd isobars, as in the case shown in Figure 2.1.

The neutrinoless decay mode ($0\nu\beta\beta$),

$$ (Z, A) \rightarrow (Z + 2, A) + 2 e^-, \quad (2.2) $$

was proposed by Wendell H. Furry in 1939 [52] as a method to test Majorana’s theory [37] applied to neutrinos. In contrast to the two-neutrino mode, the neutrinoless mode violates total lepton number conservation, and is, therefore, forbidden in the Standard Model of particle physics. Its existence is linked to that of Majorana neutrinos (§2.2). No convincing experimental evidence of the decay exists to date (§3.3).
Phase-space considerations alone would give preference to the $0\nu\beta\beta$ mode over the $2\nu\beta\beta$ one, but the decay rate of the former is suppressed by the very small neutrino masses (§ 2.3). Both transition modes involve the $0^+$ ground state of the initial nucleus and, in almost all cases, the $0^+$ ground state of the final nucleus. For some isotopes, it is also energetically possible to have a transition to an excited $0^+$ or $2^+$ final state, even though these are suppressed because of the smaller phase space available. In both decay modes the emitted leptons carry essentially all the available energy and the nuclear recoil is negligible. Therefore, in the $0\nu\beta\beta$ mode, the spectrum for the sum of the kinetic energies of the emitted electrons (see Figure 2.2) is a mono-energetic line at $Q_{\beta\beta}$, the $Q$ value of the reaction, defined as the mass difference between the parent and daughter nuclides:

$$Q_{\beta\beta} \equiv M(A, Z) - M(A, Z + 2).$$ (2.3)

In the case of the $2\nu\beta\beta$-decay mode, the spectrum is continuous, extending
2.1. DOUBLE BETA DECAY MODES

Figure 2.2. Spectra for the sum of the kinetic energies of the two emitted electrons in three different $\beta\beta$ modes: $2\nu\beta\beta$, $0\nu\beta\beta$ and $\beta\beta$ decay with Majoron emission. The amplitudes are arbitrary.

from 0 to $Q_{\beta\beta}$ and peaking below $Q_{\beta\beta}/2$.

In addition to the the two basic decay modes described above, several decay modes involving the emission of a light neutral boson, the Majoron ($\chi^0$), have been proposed in extensions of the Standard Model ($\S$ 2.4).

While in the following we will focus on the $0\nu\beta\beta$ decay as defined in Equation (2.2), there are three closely related lepton number violating processes that can be investigated:

\[
0\nu\beta^+\beta^+ : \quad (Z, A) \rightarrow (Z - 2, A) + 2 e^+ \quad (2.4)
\]

\[
0\nu\beta^+\text{EC} : \quad (Z, A) + e^- \rightarrow (Z - 2, A) + e^+ \quad (2.5)
\]

\[
0\nu\text{ECEC} : \quad (Z, A) + 2 e^- \rightarrow (Z - 2, A)^* \quad (2.6)
\]

Such processes are called respectively double positron emission, single positron emission plus electron capture (EC), and double electron capture. All three involve transitions where the nuclear charge decreases (as opposed to increasing, as in the $0\nu\beta\beta$ decay) by two units. From a theoretical point of view, the physics probed by all these processes is identical to that probed by the $0\nu\beta\beta$.
CHAPTER 2. NEUTRINOLESS DOUBLE BETA DECAY AND MAJORANA NEUTRINOS

Figure 2.3. Diagram showing how any neutrinoless double beta decay process induces a $\bar{\nu}$-to-$\nu$ transition, that is, an effective Majorana mass term. This is the so-called black box theorem \[56\].

decay. From the experimental point of view, however, they are less favorable than $0\nu\beta\beta$ decay due to the smaller phase space available \[53-55\].

2.2 The black box theorem

In general, any source of lepton number violation predicted by theories that extend the Standard model can lead to $0\nu\beta\beta$ decay. Nevertheless, as it was first pointed out by Schechter and Valle in 1982 \[56\], irrespective of the mechanism, the existence of neutrinoless double beta decay necessarily implies Majorana neutrinos, since any $\Delta L \neq 0$ diagram contributing to the decay would also contribute to the $(e, e)$ entry of the Majorana neutrino mass matrix, $(m_\nu)_{ee}$. This is called the black box theorem, and is illustrated in Figure 2.3 where a $\bar{\nu}_e \rightarrow \nu_e$ transition, i.e. a non-zero $(m_\nu)_{ee}$, is induced as a consequence of any $\Delta L \neq 0$ operator responsible for $0\nu\beta\beta$ decay.

From a quantitative point of view, however, the diagram in Figure 2.3 corresponds to a tiny mass generated at four-loop level that is far too small to explain the neutrino mass splittings observed in neutrino oscillation experiments \[57\]. Other, unknown, Majorana or Dirac mass contributions must exist. As a consequence, the black box theorem cannot tell us anything about the physics mechanism that would make the dominant contribution to a $0\nu\beta\beta$-decay rate large enough to be observable. As such, theory cannot provide direct insight into whether this dominant mechanism is related directly, indirectly or not at all related to neutrino oscillation phenomenology \[58\]. The first case is realized in the standard $0\nu\beta\beta$ mechanism of light neutrino exchange, discussed in §2.3. The last case involves alternative $0\nu\beta\beta$ mechanisms, briefly outlined in §2.4.
2.3. The standard mechanism

Neutrinoless double beta decay can arise from a diagram (see Figure 2.4) in which the parent nucleus emits a pair of virtual $W$ bosons, and then these exchange a Majorana neutrino to produce the outgoing electrons. At the vertex where it is emitted, the exchanged neutrino is created, in association with an electron, as an antineutrino with almost total positive helicity, and only its small, $O(m_\nu/E)$, negative-helicity component is absorbed at the other vertex. Considering that the amplitude ($\S/one.fitted./five.fitted$) is, in this case, a sum over the contributions of the three light neutrino mass states $\nu_i$ and is proportional to $U_{ei}^2$, we conclude that the modulus of the amplitude for the $0\nu\beta\beta$ process must be proportional in this case to the effective neutrino Majorana mass:

$$m_{\beta\beta} \equiv \left| \sum_{i=1}^{3} U_{ei}^2 \cdot m_i \right|$$  \hspace{1cm} (2.7)

In other words, the effective neutrino Majorana mass corresponds to the modulus of the $(e, e)$ element of the neutrino mass matrix of Eq. (1.34), $m_{\beta\beta} \equiv \left| (m_\nu)_{e\ell} \right|$.

In the case where light Majorana neutrino exchange is the dominant contribution to $0\nu\beta\beta$-decay, the inverse of the half-life for the process can be written as [59]

$$\left( T_{1/2}^{0\nu} \right)^{-1} = G^{0\nu} \left| M^{0\nu} \right|^2 \left( \frac{m_{\beta\beta}}{m_e} \right)^2 .$$  \hspace{1cm} (2.8)
Figure 2.5. The effective neutrino Majorana mass, $m_{\beta\beta}$, as a function of the lightest neutrino mass, $m_{\text{light}}$. The green band corresponds to the inverted ordering of neutrino masses ($m_{\text{light}} = m_3$), while the red band corresponds to the normal ordering ($m_{\text{light}} = m_1$). The vertically-excluded region comes from cosmological bounds, the horizontal one from $0\nu\beta\beta$ constraints. This graphical representation was first proposed by F. Vissani [61].

Here, $G^{0\nu}$ is a phase-space factor that depends on the transition $Q$ value and on the nuclear charge $Z$, and $M^{0\nu}$ is the nuclear matrix element (NME). The phase-space factor can be calculated analytically with sufficient accuracy (error estimates of about 1 per mille) [60]. The NME is evaluated using nuclear models, although with considerable uncertainty ($\S$2.5). In other words, the value of the effective neutrino Majorana mass, $m_{\beta\beta}$, can be inferred from a non-zero $0\nu\beta\beta$-rate measurement, even though with some nuclear physics uncertainties. Conversely, if a given experiment does not observe the $0\nu\beta\beta$ process, the result can be interpreted in terms of an upper bound on $m_{\beta\beta}$.

If light Majorana neutrino exchange is the dominant mechanism for $0\nu\beta\beta$ decay, it is clear from Eq. (2.7) that the decay is then directly connected to neutrino oscillations phenomenology, and that it also provides direct information about the absolute neutrino mass scale, as cosmology and $\beta$-decay experiments do ($\S$1.1). The relationship between $m_{\beta\beta}$ and the actual neutrino masses $m_i$ is affected by the uncertainties in the measured oscillation parameters, the
unknown neutrino mass ordering (normal or inverted) and the unknown phases in the neutrino mixing matrix (both Dirac and Majorana). For example, the relationship between $m_{\beta\beta}$ and the lightest neutrino mass, $m_{\text{light}}$, is shown in Figure 2.5. The width of the two bands is due to the unknown CP violation phases and the uncertainties in the measured oscillation parameters ($3\sigma$ ranges quoted in Table 1.1). Figure 2.5 also shows the upper bound on $m_{\text{light}}$ from cosmology discussed in §1.1 ($m_{\text{light}} < 0.07$ eV), and an upper bound on $m_{\beta\beta}$ from $0\nu\beta\beta$-decay searches ($m_{\beta\beta} < 0.2$ eV) [62,63], which we will review in §3.3. As can be seen from the figure, the bound from $0\nu\beta\beta$-decay data on the absolute mass scale is almost as stringent as that from cosmological observations.

In Figures 1.2 and 2.5, we have shown only upper bounds on various neutrino mass combinations coming from current data. The detection of positive results for absolute neutrino mass scale observables would open up the possibility to further explore neutrino properties and lepton number violating processes. We give three examples in the following. First, the successful determination of both $m_{\beta}$ in Eq. (1.5) and $m_{\beta\beta}$ in Eq. (2.7) via $\beta$ and $0\nu\beta\beta$-decay experiments, respectively, can in principle be used to determine or constrain the phases $\alpha_i$ [64]. Second, measurements of $m_{\beta}$ or $\Sigma m_\nu$ in Eq. (1.3) may yield a constraint on $m_{\text{light}}$ that is inconsistent with a $m_{\beta\beta}$ upper limit. In this case, the non-observation of $0\nu\beta\beta$ decay could imply that neutrinos are Dirac particles. Third, measurements of $m_{\beta}$ or $\Sigma m_\nu$ may yield a constraint on $m_{\text{light}}$ that is inconsistent with a measured non-zero $m_{\beta\beta}$. This scenario would demonstrate that additional lepton number violating physics, other than light Majorana neutrino exchange, is at play in the $0\nu\beta\beta$ process. We briefly describe some of these possible $0\nu\beta\beta$-decay alternative mechanisms in the following section.

### 2.4 Alternative mechanisms

A number of alternative $0\nu\beta\beta$ mechanisms have been proposed. The realization of $0\nu\beta\beta$ decay can differ from the standard mechanism in one or several aspects:

- The Lorentz structure of the currents: Positive-chirality currents mediated by a $W_R$ boson can arise, for example, in left-right symmetric

---

[For a comprehensive discussion, the reader can refer, for instance, to the review paper by W. Rodejohann 58.]}
Figure 2.6. Examples of non-standard mechanism for 0νββ-decay: (a) heavy neutrino exchange with positive chirality currents \[65\]; (b) neutralino exchange in R-parity violating supersymmetry \[66\]; and (c) Majoron emission \[67\].

- The mass scale of the exchanged virtual particles: One example would be the presence of heavy neutrinos in the propagator of Fig. \[2.4\] in addition to the three light active neutrinos we are familiar with. Another example would be the exchange of heavy supersymmetric particles, as in Figure \[2.6\](b).

- The number of particles in the final state: A popular example involves decay modes where additional Majorons, that is, very light or massless particles which can couple to neutrinos, are produced in association with the two electrons, as in Figure \[2.6\](c).

In non-standard 0νββ mechanisms, the scale of the lepton number violating physics is often larger than the momentum transfer, in which case one speaks of short-range processes. This is in contrast to the standard 0νββ mechanism of light Majorana neutrino exchange, where the neutrino is very light compared to the energy scale, resulting in a long-range process. Non-standard, long-range 0νββ processes are, nevertheless, also possible.

In general, several contributions to the total 0νββ-decay amplitude can add coherently, enhancing the observed signal and possibly exhibiting additional effects due to their interference. In certain cases, neutrinoless double beta decay observables alone may be able to identify the dominant mechanism responsible...
for the decay. For instance, if Majorons are also emitted in association with the two electrons, energy conservation alone requires the energy spectrum to be a continuous distribution with $Q_{\beta\beta}$ as endpoint. This spectrum is potentially distinguishable from the $2\nu\beta\beta$ one (see Fig. 2.2), provided that the Majoron-neutrino coupling constant is large enough. Likewise, if positive chirality current contributions dominate the $0\nu\beta\beta$-decay rate, electrons will be emitted predominantly as positive helicity states. As a consequence, both the energy and angular correlations of the two emitted electrons will be different from the ones of the standard $0\nu\beta\beta$ mechanism. A detector capable of reconstructing individual electron tracks may therefore be able to distinguish this type of non-standard $0\nu\beta\beta$ mechanism from light Majorana neutrino exchange.

### 2.5 Nuclear matrix elements

All nuclear structure effects of neutrinoless double beta decay are included in the nuclear matrix element (NME). Its knowledge is essential in order to relate a possible half-life measurement to the neutrino masses (§ 2.3), and to compare the sensitivity and results of experiments using different $\beta\beta$ isotopes. NMEs cannot be separately measured and must be evaluated theoretically. Unfortunately, due to the many-body nature of the nuclear problem, only approximate estimates can be obtained at present time. A variety of techniques — differing in their choice of the nuclear valence space, interaction Hamiltonian and in the way the corresponding equations of motion are solved — are used for this; namely: the *interacting shell model* (ISM) \[68\][69], the *quasiparticle random-phase approximation* (QRPA) \[70][71], the *interacting boson model* (IBM-2) \[72] and the *energy density functional method* (EDF) \[73][74]\[sup 2\].

Figure 2.7 summarizes the results of the most recent calculations. The reliability of the calculations has greatly improved in the last few years, and although the results from the different techniques are not yet completely convergent, differing by up to a factor of 2, they seem to be at least fairly insensitive to the broad range of approximations made \[75\]. Therefore, if $0\nu\beta\beta$ decay were observed in one nucleus, one would be able to predict its lifetime in a different candidate nuclei with some confidence, increasing the chances that a reliable and confirmed result is obtained. In any case, it is clear that further

---

2 It is beyond the scope of this work to provide a detailed discussion of these methods. The interested reader may consult, for instance, the recent review paper by P. Vogel \[75\] and the specialized references given above.
Figure 2.7. Nuclear matrix elements (NMEs) for $0\nu\beta\beta$ decay to the ground state as calculated in four different frameworks: *interacting shell model* (ISM) [69], *quasiparticle random-phase approximation* by the Tübingen (QRPA Tü) [76] and the Jyväskylä (QRPA Jy) [77] groups, *interacting boson model* (IBM-2) [72] and *energy density functional* method (EDF) [74]. The IBM-2 calculation uses the Jastrow Miller-Spencer short-range correlations; ISM, EDF and QRPA Jy use UCOM; and QRPA Tü uses the Argonne potential. IBM-2 and QRPA Tü NMEs are evaluated with $g_A = 1.269$, and the rest are evaluated using $g_A = 1.25$.

Progress in the calculation of NMEs is needed to reduce the overall theoretical uncertainty. In particular, the origin of the differences between the models should be better understood. Work in that direction has already started [78].

The dependence of the NME on the effective axial-vector coupling constant, $g_A$, introduces another significant uncertainty in the calculated rates. Barea et al. [72] and Ejiri [79] have fitted the known half-lives for $2\nu\beta\beta$ decay and find effective values of $g_A$ of about 0.8 for ISM calculations and 0.6 for IBM-2, with a mild dependence on the mass number of the isotope. In contrast, the calculated phase-space factors for neutrinoless decay are generally presented with the free-nucleon value $g_A \simeq 1.27$ [31], measured in the weak interactions and decays of nucleons. The difference between this value and 0.6 corresponds
to a factor of 20 in decay rate. The extent of the renormalization of $g_A$ in $0\nu\beta\beta$ decay remains a topic of discussion among nuclear theorists.
The search for neutrinoless double beta decay

3.1 Introduction

The discovery of neutrinoless double beta decay would represent a major breakthrough in particle physics. A single unequivocal observation of the decay would prove the Majorana nature of neutrinos and the violation of total lepton number. Unfortunately, this is by no means an easy task. The design of a detector capable of identifying efficiently and unambiguously such a rare signal poses a considerable experimental problem. To begin with, one needs a large mass of the scarce $\beta\beta$ isotopes to probe in a reasonable time the extremely long lifetimes predicted for the process. For instance, for a Majorana neutrino mass of 50 meV, we can estimate using Eq. (2.8) that half-lives in the range of $10^{26}$ to $10^{27}$ years must be explored, i.e. 17 orders of magnitude longer than the age of the universe!

A better sense of what such long half-lives mean can be grasped with a simple calculation. Consider the radioactive decay law in the approximation $T_{1/2} \gg t$, where $t$ is the observation time. In that case, the expected number of $0\nu\beta\beta$ decays is given by

$$N = \log 2 \frac{\epsilon \cdot M \cdot N_A}{W} \frac{t}{T_{1/2}^{0\nu}},$$

where $M$ is the mass of the $\beta\beta$-emitting isotope, $W$ is its molar mass, $N_A$ is the Avogadro constant and $\epsilon$ is the detection efficiency. It follows from the above equation that in order to observe one decay per year, assuming perfect detection efficiency and no disturbing background, and for a Majorana neutrino mass of 50 meV, macroscopic masses of $\beta\beta$ isotope of the order of 100 kg are needed.

The situation becomes even more desperate when considering real experimental conditions. The detectors used in double beta decay experiments are
designed, in general, to measure the energy of the radiation emitted by a $\beta\beta$ source. In a neutrinoless double beta decay the sum of the kinetic energies of the two released electrons is always equal to the $Q$ value of the process (§2.1). However, due to the finite energy resolution of any detector, $0\nu\beta\beta$ events spread over an energy range centred around $Q_{\beta\beta}$, typically following a Gaussian distribution. Other processes occurring in the detector can fall within that energy window, becoming a background and compromising drastically the sensitivity of the experiment.

The background processes that can mimic a $0\nu\beta\beta$-decay signal in a detector are abundant. To begin with, the experiments have to deal with the intrinsic background from the standard two-neutrino double beta decay, which can only be distinguished from the signal by measuring the energy of the emitted electrons, since the neutrinos go undetected. Good energy resolution is, therefore, essential to prevent the $2\nu\beta\beta$ spectrum tail from spreading over the $0\nu\beta\beta$ peak. However, this energy signature is not enough per se, since the continuous energy spectrum arising from natural radioactivity can easily overwhelm the signal peak. For this reason, additional experimental signatures and careful selection of radiopure materials are crucial.

Several other factors, like the detection efficiency or the scalability to large masses, must be taken into account as well during the design of a double beta decay experiment. The simultaneous optimization of all these parameters is often conflicting, if not impossible, and hence many different experimental techniques have been proposed. In order to compare them, a figure of merit, the experimental sensitivity to $m_{\beta\beta}$, is normally used. It will be derived in the next section, followed by a discussion on its main parameters.

### 3.2 Sensitivity of a double beta decay experiment

#### 3.2.1 Frequentist confidence intervals and upper limits

All double beta decay experiments have to deal with non-negligible backgrounds and an only partially efficient event selection. It is instructive for our purposes, however, to imagine an ideal background-free experiment. The expected number of $0\nu\beta\beta$ events that such an experiment should observe after accumulating an exposure $Mt$ is given by Eq. (3.1). However, due to the stochastic nature of radioactivity, the experiment may actually yield a different
result, with a probability described by a *Poisson distribution*:

\[
\text{Po}(n; \mu) = \frac{\mu^n}{n!} e^{-\mu},
\]

where \( \mu = N \) is the *expected value* and \( n \) is the observed number of events. For example, for an expected value \( \mu = 5 \), if the experiment were repeated a large number of times, one would observe 5 events only in 17.5% of the cases. In fact, the same percentage would yield 4 events, and in a small but non-null number of tries, 0.7%, one would observe 0 events. What can be said, therefore, about the parameter \( \mu \) from any given measurement?

A standard way to report this type of results, proposed by Jerzy Neyman in 1937 [80], is to give a *confidence interval*, calculated from the observations, within which there is a given probability that the parameter \( \mu \) falls:

\[
P(\mu_1(n) \leq \mu \leq \mu_2(n)) = \gamma,
\]

where \( \gamma \), usually expressed as a percentage, is the *coverage probability* or *confidence level* (CL) of the interval. The relationship between the coverage probability and the endpoints of the interval, \( \mu_1 \) and \( \mu_2 \), known as the *confidence limits*, can be better understood by splitting Eq. (3.3) into two separate expressions:

\[
P(\mu_1(n) \geq \mu) = \alpha,
\]

\[
P(\mu_2(n) \leq \mu) = \beta,
\]

with \( \alpha + \beta = 1 - \gamma \). These equations do not determine the confidence interval uniquely, and one must choose additional criteria. For instance, a common choice is to require the tail probabilities \( \alpha \) and \( \beta \) be equal,

\[
\alpha = \beta = (1 - \gamma)/2,
\]

resulting in the so-called *central* confidence intervals.

Equations (3.4) and (3.5) can only be approximately satisfied in the case of a Poisson distribution, due to its discreteness. In this case, the convention is to build a *conservative* confidence interval — that is, one whose actual coverage probability is greater than the nominal — using the following inequalities:

\[
\alpha \gtrsim \sum_{n'=n}^{\infty} \text{Po}(n'; \mu_1) = 1 - \sum_{n'=0}^{n-1} \text{Po}(n', \mu_1),
\]

\[
\beta \gtrsim \sum_{n'=0}^{n} \text{Po}(n'; \mu_2).
\]
Figure 3.1. Central confidence limits $\mu_{1,2}$ computed for an observed number of events $n = 4$ and $\alpha = \beta = 0.05$. Two Poisson distributions are shown: $\text{Po}(n'; \mu_1)$, blue circles, and $\text{Po}(n'; \mu_2)$, red squares.

Figure 3.1 illustrates the above definition. The confidence limits are computed for $n = 4$ and $\alpha = \beta = 0.05$. The blue circles represent $\text{Po}(n'; \mu_1 = 1.37)$, whereas the red squares represent $\text{Po}(n'; \mu_2 = 9.15)$. Adding the circles corresponding to $n' < 4$ yields a probability of 0.95 (the integrated probability to find $n' \geq n$ is 0.05). Adding the squares corresponding to $n' > 4$ yields a probability of 0.95 (the integrated probability to find $n' \leq n$ is 0.05). We would state then that the parameter $\mu$ is in the interval $[1.37, 9.15]$ at 90% CL.

What should an experiment report if no evidence of the decay is found, i.e. $n = 0$? Clearly, a lower confidence limit cannot be found, but an upper limit can still be determined by setting $n = 0$ in Eq. (3.8) and taking logarithms of both sides:

$$
\mu_{\text{up}} \equiv \mu_2 = -\log(\beta).
$$

Therefore, an ideal experiment that observes no events after a given exposure, rather than saying that nothing was found, would report an upper limit (at a given confidence level) on the expected number of events, $N$. For example, setting $\beta = 0.1$ in Eq. (3.9), we obtain $\mu_{\text{up}} = 2.3$ at 90% CL.

Consider now the more realistic case of an experiment with background.
Define $b$ as the expected value of the background, and assume (unrealistically) that it is known with no uncertainty. The relevant pdf will then be

$$\text{Po}(n; \mu + b) = \frac{(\mu + b)^n}{n!} e^{-(\mu+b)} ,$$

where $\mu$ is the unknown mean signal expectation and $b$ is the known mean background expectation. The Poisson variable $n$ is such that $n = n_s + n_b$, where the signal and background Poisson variables $n_s$ and $n_b$ have mean expectation values $\mu$ and $b$, respectively. A priori, it appears as if we could treat the problem in exactly the same way as for the case without background, and construct a confidence interval using the following expressions:

$$\alpha \gtrsim \sum_{n'=n}^{\infty} \frac{(\mu_1 + b)^{n'}}{n'!} e^{-(\mu_1+b)} = 1 - \sum_{n'=0}^{n-1} \frac{(\mu_1 + b)^{n'}}{n'!} e^{-(\mu_1+b)} ,$$

$$\beta \gtrsim \sum_{n'=0}^{n} \frac{(\mu_2 + b)^{n'}}{n'!} e^{-(\mu_2+b)} .$$

Let us work out an explicit example. Suppose that we run an experiment in which the predicted background is $b = 3$, and we observe 3 events. Solving numerically the above equations for $\alpha = \beta = 0.05$ and $n = 3$, we obtain an upper limit $\mu_2 = 3.30$ and a lower limit $\mu_1 = -2.18$. Part of this confidence interval is meaningless, since negative values of $\mu$ are unphysical. Therefore, we may decide to quote instead a strict upper limit, solving Eq. (3.12) for $\beta = 0.1$; in this case, we obtain $\mu \leq 3.68$ at 90% CL. However, this flip-flopping between two-sided intervals and upper limits depending on the measured data results in improper coverage and, for that reason, should be avoided.

Consider now the cases in which $n = 2, 1$ and 0 for our experiment with expected background $b = 3$. The respective upper limits at 90% CL are $\mu_{up} = 2.32$, 0.89 and $-0.70$. Obviously, as the number of observed events becomes smaller, the Poisson distribution has to shift left to guarantee that the cumulative percentage to the left of $n$ gets to 10%, resulting in negative, non-physical upper limits. In short, the classical Neyman construction of confidence limits fails to provide meaningful results when the observed number of events is small compared to the expected number of background events.

The two problems described above can be circumvented using an alternative procedure for constructing confidence intervals. Neyman’s definition leaves freedom as to which particular values of the parameter $\mu$ are included in the intervals as long as proper coverage is obtained. Therefore, one can
consider a test of the hypothesis that the parameter’s true value is \( \mu \), excluding all values of \( \mu \) where the hypothesis would be rejected in a test of size \( \alpha \) or less. The remaining values constitute the confidence interval at confidence level \( 1 - \alpha \). The famous Feldman-Cousins construction follows this approach, using a test statistic based on the likelihood ratio

\[
\lambda \equiv \frac{P(n; \mu)}{P(n; \hat{\mu})},
\]

where \( \hat{\mu} \) is the value of the parameter which, out of all allowed values, maximizes \( P(n; \mu) \). This procedure, which has become the de facto standard frequentist approach to compute confidence intervals, solves the two problems we have described: it provides meaningful intervals even when the observable falls close to physical boundaries, and it offers a natural transition between upper confidence limits for null results and two-sided confidence intervals for non-null results. Figure illustrates the differences between the classical construction and Feldman-Cousins for the case of an unknown Poisson signal mean \( \mu \) in the presence of Poisson background with known mean \( b = 3 \).

### 3.2.2 Sensitivity to \( m_{\beta\beta} \)

The sensitivity of an experiment searching for new phenomena is a measure of the result that would be obtained in the absence of a true signal. More precisely, it is defined as the average confidence limit one would get from a large ensemble of experiments with the same expected background and no signal. Accordingly, the sensitivity of a double beta decay experiment to \( m_{\beta\beta} \) can be expressed combining Equations (2.8) and (3.1) as

\[
S(m_{\beta\beta}) = A \sqrt{\frac{N}{\varepsilon M t}},
\]

where

\[
A \equiv \left( \frac{W}{N_A \log 2} \frac{m_e^2}{G^{0\nu} |M^{0\nu}|^2} \right)^{1/2}
\]

is a constant that depends only on the considered \( \beta\beta \) isotope and \( \bar{N} \) is the average upper limit on the expected number of events in the absence of signal.

For our ideal 0\( \nu \beta\beta \)-decay experiment, the expected background is exactly \( b = 0 \). Therefore, under the no-signal hypothesis, the observed number of
3.2. SENSITIVITY OF A DOUBLE BETA DECAY EXPERIMENT

Figure 3.2. Confidence belts (90% CL) for an unknown Poisson signal mean $\mu$ in the presence of Poisson background with known mean $b = 3$ using three different procedures: classical central confidence intervals (dashed, blue lines), classical upper limits (solid, red line) and Feldman-Cousins intervals (shaded region). Notice that the classical construction results in an empty set for $n = 0$.

events would always be equal to zero, with no fluctuations. The average upper limit, $\bar{N}$, is in this case simply given by Equation (3.9), and consequently the sensitivity is only a function of $\left(Mt\right)^{-1/2}$.

For an experiment with expected background $b$, the average upper limit is given by

$$\bar{N}(b) = \sum_{n=0}^{\infty} \text{Po}(n; b) \ U(n; b),$$

where $U(n; b)$ is a function that returns the Feldman-Cousins upper limit for a given observation $n$ and a known background level $b$. Figure 3.3 shows the value of $\bar{N}$ as a function of $b$ for 4 different confidence levels. In the case of large background, these curves are approximately given by

$$\bar{N}(b) \approx k \sqrt{b}.$$  

Substituting Eq. (3.17) into Eq. (3.14), we obtain

$$S(m_{\beta\beta}) = \mathcal{A}' \left( \frac{b^{1/2}}{\varepsilon M t} \right)^{1/2},$$
with
\[ A' \equiv A \sqrt{k}. \] (3.19)

Usually, the background is approximately proportional to the exposure, \( Mt \), and to the width of the energy window \( \Delta E \) defined by the resolution of the detector:
\[ b = c \cdot M \cdot t \cdot \Delta E, \] (3.20)

with the background rate \( c \) typically expressed in counts/(keV \cdot kg \cdot year). Equation (3.18) then becomes
\[ S(m_{\beta\beta}) = A' \sqrt{1/\epsilon} \left( \frac{c \Delta E}{M t} \right)^{1/4}. \] (3.21)

In short, the background limits dramatically the sensitivity of a double beta decay experiment, improving only as \((Mt)^{-1/4}\) instead of the \((Mt)^{-1/2}\) expected in the background-free case.

### 3.3 History and status of double beta decay searches

For almost half a century the only evidence of the existence of double beta decay came from geochemical methods that measure the concentration of the stable daughter isotopes produced over geologic times (\(\sim 10^9\) years). An excess of the daughter isotope over its natural concentration is interpreted as evidence for double beta decay, either \(2\nu\beta\beta\) or \(0\nu\beta\beta\), since the method cannot distinguish between them.

The first direct measurement of the two-neutrino mode, in \(^{82}\text{Se}\), did not happen until 1987. It was done with a fairly large (\(\sim 1\) m\(^3\)) time projection chamber. The source, 14 g of selenium enriched to 97% in \(^{82}\text{Se}\), was deposited on a thin Mylar foil forming the central electrode of the chamber. The trajectories of the electrons emitted from this source foil were recorded by the TPC and analyzed to infer their energy and kinematic characteristics. Since this initial detection, the two-neutrino mode has been directly observed by several experiments in 9 isotopes (see Table 3.1).

---

1Traditionally, the exposure and background rate have been expressed in units of detector mass. However, given the disparity of detector masses used by the new generation of \(0\nu\beta\beta\)-decay experiments (for example, KamLAND-Zen deploys 13 tonnes of \(^{136}\text{Xe}\)-loaded liquid scintillator, whereas GERDA will use about 40 kg of enriched germanium diodes in its seconds phase), in this work, in order to facilitate the comparison between experiments, we will express the relevant quantities in units of \(\beta\beta\) isotope mass unless stated otherwise.
Figure 3.3. Average upper limit on the expected number of events observed by a large ensemble of experiments with the expected background and no true signal as a function of the expected background for the case of a measurement of a Poisson variable.

<table>
<thead>
<tr>
<th>Isotope</th>
<th>$T_{1/2}^{2\nu} (10^{19} \text{ yr})$</th>
<th>Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{48}\text{Ca}$</td>
<td>$4.4^{+0.6}_{-0.5}$</td>
<td>Irvine TPC [83], TGV [84], NEMO-3 [85]</td>
</tr>
<tr>
<td>$^{76}\text{Ge}$</td>
<td>$160^{+13}_{-10}$</td>
<td>ITEP/YePI [86], IGEX [87], H-M [88], GERDA [89]</td>
</tr>
<tr>
<td>$^{82}\text{Se}$</td>
<td>$9.2 \pm 0.7$</td>
<td>Irvine TPC [90], NEMO-2 [91], NEMO-3 [92]</td>
</tr>
<tr>
<td>$^{96}\text{Zr}$</td>
<td>$2.3 \pm 0.2$</td>
<td>NEMO-2 [93], NEMO-3 [94]</td>
</tr>
<tr>
<td>$^{100}\text{Mo}$</td>
<td>$0.71 \pm 0.04$</td>
<td>Irvine TPC [95], NEMO-2 [96], NEMO-3 [92]</td>
</tr>
<tr>
<td>$^{116}\text{Cd}$</td>
<td>$2.85 \pm 0.15$</td>
<td>ELEGANT [97], Solotvina [98], NEMO-3 [85]</td>
</tr>
<tr>
<td>$^{130}\text{Te}$</td>
<td>$69 \pm 13$</td>
<td>CUORICINO [99], NEMO-3 [100]</td>
</tr>
<tr>
<td>$^{136}\text{Xe}$</td>
<td>$220 \pm 6$</td>
<td>EXO-200 [101], KamLAND-Zen [102]</td>
</tr>
<tr>
<td>$^{150}\text{Nd}$</td>
<td>$0.82 \pm 0.09$</td>
<td>Irvine TPC [95], ITEP TPC [103], NEMO-3 [104]</td>
</tr>
</tbody>
</table>
During the 1990s, $0\nu\beta\beta$-decay searches were dominated by germanium calorimeters. In particular, for many years the best limit to the half-life of the process was the one set by the Heidelberg-Moscow (HM) experiment, $T_{1/2}^{0\nu}(^{76}\text{Ge}) > 1.9 \times 10^{25}$ yr (90% CL) \[105\]. The detector was composed of five high-purity germanium diodes enriched to 86% in $^{76}\text{Ge}$. The experiment ran in the Laboratori Nazionali del Gran Sasso (LNGS), Italy, from 1990 to 2003, totaling an exposure of 71.7 kg · year. The background rate reached by the experiment in the energy region around $Q_{\beta\beta}$ was, in units of $\beta\beta$ emitter mass, 0.22 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. Pulse shape discrimination was used in a subset of the data (35.5 kg yr) to separate single-site events (like $0\nu\beta\beta$ decays) from multi-site events (like $\gamma$ interactions), reducing the background rate in about a factor of 3 to 0.07 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$.

A subset of the Heidelberg-Moscow Collaboration re-analyzed the experiment’s data and claimed to find evidence of $0\nu\beta\beta$ decay \[106\], sparking an intense debate in the community \[107\]. The latest publication by this group reported a 6σ evidence for $0\nu\beta\beta$ decay and a half-life measurement of $T_{1/2}^{0\nu} = (2.23^{+0.44}_{-0.31}) \times 10^{25}$ yr \[108\], corresponding to a neutrino Majorana mass of about 300 meV.

The International Germanium Experiment (IGEX) \[109\] also searched for $0\nu\beta\beta$ decay using enriched germanium detectors. It ran in the Homestake Mine (USA), the old Laboratorio Subterráneo de Canfranc (Spain) and the Baksan Neutrino Observatory (Russia) from 1991 to 2000, accumulating a total exposure of 8.87 kg year. Using pulse shape discrimination, the experiment reached a background rate of 0.12 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$ \[110\]. Its final sensitivity was slightly lower than that of Heidelberg-Moscow, and thus insufficient to disprove the claim of a positive signal.

The Cuoricino experiment \[111\], an array of 62 TeO$_2$ bolometric crystals, ran for five years at LNGS searching for $0\nu\beta\beta$ decay in $^{130}\text{Te}$. The average FWHM resolution for these crystals was (6.3 ± 2.5) keV at 2615 keV. Their physical efficiency — mostly due to the geometrical effect of particles escaping the detector — was estimated to be 0.874 ± 0.011. The average background rate for the Cuoricino crystals computed in a 60-keV wide region centred around $Q_{\beta\beta}$ was 0.58 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$ \[112\]. Its final sensitivity was comparable to that of the HM experiment, but it was not able disprove the claim due to the uncertainties in the nuclear matrix elements.

The NEMO-3 experiment \[113\] searched for the double beta decay of seven isotopes, running from 2003 to 2010 at the Modane Underground Laboratory.
3.4. INGREDIENTS FOR THE ULTIMATE DOUBLE BETA DECAY EXPERIMENTS

(LSM), France. The detector consisted of 20 sectors arranged in a cylindrical geometry containing thin (40–60 mg/cm²) source foils of ββ emitters suspended between two concentric cylindrical tracking volumes. The largest ββ source was 100Mo, with a mass of 6.914 kg. The tracking detector was surrounded by a calorimeter made of large blocks of plastic scintillator coupled to low-radioactivity photomultiplier tubes. A solenoid generated a magnetic field that allowed the measurement of the charge sign of tracked particles. By reconstructing the full topology of the events, NEMO-3 achieved the lowest background rate of the previous generation of 0νββ-decay experiments, 1.3 × 10⁻³ cts keV⁻¹ kg⁻¹ yr⁻¹ [114]. The FWHM resolution of the detector at the Q value of 100Mo (3034 keV) was 11.5%, and the 0νββ detection efficiency was estimated to be 4.7% [114]. After accumulating an exposure of 34.7 kg yr, no evidence for the 0νββ decay of 100Mo was found, setting the following half-life limit: $T_{1/2}^{0\nu}(100\text{Mo}) > 1.1 \times 10^{24}$ (90% CL) [114].

The conclusion of Cuoricino and NEMO-3 marked in some way the transition toward a new generation of experiments characterized by bigger detectors (tens to hundreds of kilograms of ββ isotope) designed and constructed by international collaborations. The first experiments of this generation to produce physics results, EXO-200 [115] and KamLAND-Zen [102], are both searching for 0νββ decay in 136Xe. The combination of the 0νββ half-life limits reported by the two experiments is 3.4 × 10²⁵ years (90% CL), corresponding to $m_{\beta\beta} < 120–250$ meV. This result refutes the detection claim in 76Ge at > 97.5% CL [102]. The GERDA experiment, a new-generation germanium experiment, has also published its first results. No evidence of a 0νββ peak has been found in the data, setting the following half-life limit: $T_{1/2}^{0\nu}(76\text{Ge}) > 2.1 \times 10^{25}$ yr (90% CL) [116]. The 76Ge claim is hence strongly disfavoured. These three experiments, together with the other main projects from the present generation, are described in more detail in §3.5.

Table 3.2 summarizes the best present limits to the half-life of 0νββ decay achieved in 9 different isotopes.

3.4 Ingredients for the ultimate double beta decay experiments

The first objective of the new generation of double beta decay experiments, to confirm or refute experimentally the claim of a 0νββ-decay signal in 76Ge, has already been tackled by EXO-200, KamLAND-Zen and GERDA (§3.3), and in the next 2 to 3 years, more data from other experiments (§3.5) will
Table 3.2. Best present limits (at 90% CL) to the half-life of neutrinoless double beta decay.

<table>
<thead>
<tr>
<th>Isotope</th>
<th>$T_{1/2}^{0\nu}$ (10$^{25}$ yr)</th>
<th>Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{48}$Ca</td>
<td>$&gt; 0.0058$</td>
<td>ELEGANT $^{[117]}$</td>
</tr>
<tr>
<td>$^{76}$Ge</td>
<td>$&gt; 3.0$</td>
<td>Heidelberg-Moscow $^{[105]}$, GERDA $^{[116]}$</td>
</tr>
<tr>
<td>$^{82}$Se</td>
<td>$&gt; 0.036$</td>
<td>NEMO-3 $^{[118]}$</td>
</tr>
<tr>
<td>$^{96}$Zr</td>
<td>$&gt; 0.00092$</td>
<td>NEMO-3 $^{[94]}$</td>
</tr>
<tr>
<td>$^{100}$Mo</td>
<td>$&gt; 0.11$</td>
<td>NEMO-3 $^{[114]}$</td>
</tr>
<tr>
<td>$^{116}$Cd</td>
<td>$&gt; 0.017$</td>
<td>Solotvina $^{[98]}$</td>
</tr>
<tr>
<td>$^{130}$Te</td>
<td>$&gt; 0.28$</td>
<td>Cuoricino $^{[111]}$</td>
</tr>
<tr>
<td>$^{136}$Xe</td>
<td>$&gt; 3.4$</td>
<td>EXO-200 $^{[62][115]}$, KamLAND-Zen $^{[102]}$</td>
</tr>
<tr>
<td>$^{150}$Nd</td>
<td>$&gt; 0.0018$</td>
<td>NEMO-3 $^{[104]}$</td>
</tr>
</tbody>
</table>

be available as well. The ultimate goal of all these projects is the exploration of the inverted hierarchy of neutrino masses, a very ambitious objective that will require exposures close to $10^4$ kg yr and background levels of the order of $10^{-3}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$ or better. Given the scale, cost and risk that characterize these experiments, it seems prudent to build as a first step an $\mathcal{O}(100)$-kg detector that demonstrates the performance of the experimental technique. Nevertheless, scaling up these detector to large $\beta\beta$ masses will not be straightforward in most cases.

We saw in §3.2.2 that the sensitivity of any $0\nu\beta\beta$-decay experiment depends, ultimately, on a few parameters; namely: energy resolution, background rate, detection efficiency, exposure and $\beta\beta$ isotope. In this section, we discuss the requirements that the exploration of the inverted-hierarchy region of neutrino masses imposes on them. Some of the parameters (the energy resolution, for instance) depend only on the experimental technique and cannot be improved at will. Others are also determined by factors unrelated to the detection technique (e.g. the background rate may depend on the availability of radiopure materials or the depth of the underground laboratory), leaving more room for improvement.

### 3.4.1 Choice of the isotope

Thirty-five naturally-occurring isotopes are $\beta\beta$ emitters. Which ones are the most adequate for neutrinoless double beta decay searches? Let us start with
considerations of the most favourable phase-space factors and nuclear matrix
elements. To a first approximation, the phase-space factor $G^{0\nu}$ varies as $Q_{\beta\beta}^5 \ [60]$. Isotopes with large $Q$ values are therefore strongly favoured. For this
reason, only the eleven with $Q_{\beta\beta} > 2$ MeV (see Table 3.3) have usually been
considered for $0\nu\beta\beta$-decay searches. As for the NMEs, all calculation methods
agree that the value of $|M^{0\nu}|$ does not change abruptly from one candidate
isotope to another one ($\S\ 2.5$). The isotopic constant $A$ defined in Eq. (3.14)
shows variations of about a factor of 2 (see Figure 3.4). In other words, within
a factor of 2, the decay rate per unit mass does not depend on the $\beta\beta$ isotope.

Another advantage in choosing a $\beta\beta$ isotope with a high $Q$ value comes
in the form of background control. As we will discuss later ($\S\ 3.4.3$), back-
grounds from natural radioactivity populate the energy region below 3 MeV.
The possibility to use an isotope with $Q_{\beta\beta}$ above these energies is therefore
desirable.

Experimental techniques with modest energy resolution (say, more than
3–4% FWHM at the $Q$ value) are also interested in choosing an isotope with
a relatively slow two-neutrino decay mode. The significance of the $2\nu\beta\beta$
spectrum as a background depends on its spectral form and rate. It is irrelevant
if the energy resolution of the experiment is effectively perfect, but as the energy
resolution degrades, it can become a serious background. This is illustrated
in Figure 3.3 where the $m_{\beta\beta}$ sensitivity (90% CL) as a function of the energy
resolution is shown for experiments using $^{136}$Xe and $^{82}$Se. The experiments
are assumed to have perfect detection efficiency and be affected only by $2\nu\beta\beta$
backgrounds. In these idealized conditions, it is clear that $^{136}$Xe is preferable
to $^{82}$Se for low-resolution experiments thanks to its much longer $2\nu\beta\beta$-decay half-life (see Table 3.1).

The general conclusion is that there is no magic candidate, no specially
favoured or disfavoured isotope among those with a $Q$ value larger than 2 MeV.
The choice of the $\beta\beta$ isotope is, therefore, ultimately driven by its procurement
cost ($\S\ 3.4.5$) and by the experimental technique.

3.4.2 Energy resolution

High energy resolution is a necessary condition (but not sufficient) for an
ultimate $0\nu\beta\beta$-decay experiment: it is the only protection against the intrin-
sic $2\nu\beta\beta$ background (see Fig. 3.5), and improves the signal-to-noise ratio in
the region of interest around $Q_{\beta\beta}$. Figure 3.6 illustrates the latter point: the
energy region of interest around the $Q$ value is represented for three Monte
Figure 3.4. Isotopic constant $A$ — see Eq. (3.14) — of the eleven $\beta\beta$ isotopes with a $Q$ value larger than 2 MeV. Variations of about a factor of 2 can be found for a given set of NMEs.

Figure 3.5. Dependence on the energy resolution of the sensitivity to $m_{\beta\beta}$ (at 90% CL) of experiments with 500 kg · year of exposure and affected only by $2\nu\beta\beta$-decay backgrounds.
Table 3.3. Relevant properties of the eleven $\beta\beta$-emitter isotopes with a $Q$ value greater than 2 MeV: atomic weight, $W$; natural isotopic abundance, IA; $Q$ value of the $\beta\beta$ process, $Q_{\beta\beta}$; phase-space factor, $G^{0\nu}$; and nuclear matrix element (NME) in different nuclear models, $M^{(0\nu)}$. The $Q$ values have been measured with very good precision in Penning-trap spectrometers [119–125]. The phase-space factors, $G^{0\nu}$, are those recently calculated by Kotila and Iachello [60]. The quoted NMEs are the latest results from the models and groups cited in §2.5 (see also Fig. 2.7).

<table>
<thead>
<tr>
<th>Isotope</th>
<th>$W$ (g mol(^{-1}))</th>
<th>IA (%)</th>
<th>$Q_{\beta\beta}$ (keV)</th>
<th>$G^{0\nu}$ (10(^{-15}) yr(^{-1}))</th>
<th>ISM</th>
<th>QRPA Tü</th>
<th>QRPA Jy</th>
<th>IBM-2</th>
<th>EDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{48})Ca</td>
<td>47.95</td>
<td>0.19</td>
<td>4263</td>
<td>24.81</td>
<td>0.85</td>
<td>0.541</td>
<td>1.98</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td>(^{76})Ge</td>
<td>75.92</td>
<td>7.73</td>
<td>2039</td>
<td>2.36</td>
<td>2.81</td>
<td>5.157</td>
<td>5.18</td>
<td>5.42</td>
<td>5.55</td>
</tr>
<tr>
<td>(^{82})Se</td>
<td>81.92</td>
<td>8.73</td>
<td>2998</td>
<td>10.16</td>
<td>2.64</td>
<td>4.642</td>
<td>4.20</td>
<td>4.37</td>
<td>4.67</td>
</tr>
<tr>
<td>(^{96})Zr</td>
<td>95.91</td>
<td>2.80</td>
<td>3346</td>
<td>20.58</td>
<td>2.717</td>
<td>3.12</td>
<td>2.53</td>
<td>6.50</td>
<td></td>
</tr>
<tr>
<td>(^{100})Mo</td>
<td>99.91</td>
<td>9.82</td>
<td>3034</td>
<td>15.92</td>
<td>5.402</td>
<td>3.93</td>
<td>3.73</td>
<td>6.59</td>
<td></td>
</tr>
<tr>
<td>(^{110})Pd</td>
<td>109.91</td>
<td>11.72</td>
<td>2018</td>
<td>4.82</td>
<td>5.762</td>
<td>5.63</td>
<td>3.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(^{116})Cd</td>
<td>115.90</td>
<td>7.49</td>
<td>2814</td>
<td>16.70</td>
<td>4.040</td>
<td>3.93</td>
<td>2.78</td>
<td>5.35</td>
<td></td>
</tr>
<tr>
<td>(^{124})Sn</td>
<td>123.91</td>
<td>5.79</td>
<td>2287</td>
<td>9.04</td>
<td>2.62</td>
<td>2.558</td>
<td>4.57</td>
<td>3.50</td>
<td>5.79</td>
</tr>
<tr>
<td>(^{130})Te</td>
<td>129.91</td>
<td>34.08</td>
<td>2528</td>
<td>14.22</td>
<td>2.65</td>
<td>3.888</td>
<td>4.76</td>
<td>4.03</td>
<td>6.41</td>
</tr>
<tr>
<td>(^{136})Xe</td>
<td>135.91</td>
<td>8.86</td>
<td>2458</td>
<td>14.58</td>
<td>2.19</td>
<td>2.177</td>
<td>3.16</td>
<td>3.33</td>
<td>4.77</td>
</tr>
<tr>
<td>(^{150})Nd</td>
<td>149.92</td>
<td>5.64</td>
<td>3371</td>
<td>63.03</td>
<td>2.32</td>
<td>2.19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Carlo experiments with the same signal and background, but different energy resolution. The signal is distributed normally around $Q_{\beta\beta}$, whereas the background is assumed flat in the window. The signal strength (50 counts) and the background rate (1 count/keV) correspond to typical values for a tonne-scale experiment and a Majorana neutrino mass of the order of 100 meV. A clear peak rising above the background is visible in the case of the experiment with better energy resolution (1% FWHM). However, the peak is hardly discernible for the experiment with a 3% FWHM resolution, and it has disappeared completely for the worst-resolution case (10% FWHM). In conclusion, the better the energy resolution, the greater the discovery potential of an experiment given certain conditions of signal and background. Therefore, the experimental techniques with modest or poor energy resolution have to compensate this deficiency reaching lower background rates and higher exposures.

3.4.3 Low background

We have seen already (§3.2.2) that the presence of background in the region of interest around $Q_{\beta\beta}$ changes the regime of the sensitivity to $m_{\beta\beta}$ of an experiment from an inverse square-root dependence on the exposure to an inverse fourth root. For this reason, the main developmental challenges for any double beta decay experiment are all concerned with the suppression of backgrounds.

Figure 3.7 shows the sensitivity (at 90% CL) of a $^{136}$Xe-based\textsuperscript{2} experiment with four different assumptions for the background rate within the energy region of interest: $10^{-1}$, $10^{-2}$ and $10^{-3}$ counts kg$^{-1}$ yr$^{-1}$, and zero background. EXO-200 and KamLAND-Zen have achieved a background rate of approximately 0.19 and 0.04 cts kg$^{-1}$ yr$^{-1}$, respectively (see the corresponding descriptions in §3.5). Both experiments have accumulated about 100 kg yr of exposure, and thus, according to the figure, their sensitivities are slightly above 100 meV. In order to probe Majorana neutrino masses down to 20 meV, the tonne-scale versions of these experiment must improve the quoted background rates in more than one order of magnitude, reaching values close to $10^{-3}$ counts kg$^{-1}$ yr$^{-1}$. The exploration of the normal-hierarchy region of neutrino masses ($m_{\beta\beta} < 20$ meV) will only be possible with gigantic, background-free experiments.

\textsuperscript{2}The conclusions reached here would be the same regardless of the $\beta\beta$ isotope used for the discussion.
Figure 3.6. Signal and background (red and grey stacked histograms, respectively) in the region of interest around $Q_{\beta\beta}$ for three Monte Carlo experiments with the same signal strength (50 counts) and background rate (1 count/keV), but different energy resolution (top: 1% FWHM; centre: 3.5% FWHM; bottom: 10% FWHM). The signal is distributed normally around $Q_{\beta\beta}$, while the background is assumed flat.
Figure 3.7. Sensitivity to $m_{\beta\beta}$ (at 90% CL) of a $^{136}$Xe-based experiment with perfect signal detection efficiency and four different assumptions for the background rate within the energy region of interest around $Q_{\beta\beta}$: 0.1 cts kg$^{-1}$ yr$^{-1}$ (green, dotted line); 0.01 cts kg$^{-1}$ yr$^{-1}$ (blue, dash-dot line); 0.001 cts kg$^{-1}$ yr$^{-1}$ (red, dashed line) and background-free (grey, solid line). The IBM-2 nuclear matrix element has been used to convert the half-life limits to $m_{\beta\beta}$. The grey band represents the inverted-hierarchy region of neutrino masses.

The natural radioactivity of detector components is usually the main background in $0\nu\beta\beta$-decay experiments. Even though the half-lives of the natural decay chains are comparable to the age of the universe, they are very short compared to the half-life sensitivity of the new-generation experiments. Consequently, even traces of these nuclides can become a significant background. Particularly pernicious are $^{208}$Tl and $^{214}$Bi — decay products of the thorium and uranium series, respectively (see Figures 3.8 and 3.9) — due to the high energy of the particles emitted in their decays. These isotopes are present at some level in all materials. Therefore, careful selection of radiopure materials is mandatory for all $0\nu\beta\beta$-decay experiments. New-generation detectors are being fabricated from components with activities as low as a few microbecquerels per kilogram or less.

Radon, another intermediate decay product of the uranium and thorium
3.4. INGREDIENTS FOR THE ULTIMATE DOUBLE BETA DECAY EXPERIMENTS

Figure 3.8. Thorium decay series, from $^{232}$Th to $^{208}$Pb.

NOTES:
The symbols $\alpha$ and $\beta$ indicate alpha and beta decay, and the times shown are half-lives.
An asterisk indicates that the isotope is also a significant gamma emitter.
Figure 3.9. Uranium decay series, from $^{238}$U to $^{206}$Pb.
series, is also a concern for most experiments. It is one of the densest substances that remains a gas under normal conditions, and it is also the only gas in the atmosphere that only has radioactive isotopes. While the average rate of production of $^{220}\text{Rn}$ (from the thorium decay series) is about the same as $^{222}\text{Rn}$, the longer half-life of the latter (3.8 days versus 55 seconds) makes it much more abundant. Being a noble gas, radon is chemically not very reactive and can diffuse easily through many materials, infiltrating into the active region of the detectors. Radon progenies, also radioactive, tend to be charged and adhere to surfaces or dust particles. The impact of radon can be mitigated by flushing the detector surroundings with pure nitrogen or by installing radon traps in the laboratory air circulation systems.

In addition to the backgrounds coming from radioactive impurities in detector components, there are external backgrounds originating outside the detector. These can be suppressed by placing the detector underground and by enclosing it in a shielding system. Very efficient shielding and additional detection signatures such as track reconstruction can compensate the benefits of a very deep location. Several underground facilities are currently available to host physics experiments around the world. Figure 3.10 arranges them according to their depth and corresponding cosmic ray muon flux. In addition to depth, the other important factor characterizing the underground sites is the size of the excavated halls, since future experimental proposals at the tonne scale will need large underground volumes.

At the depths of underground laboratories, muons and neutrinos are the only surviving radiation from the atmosphere and outer space. Future very massive detectors will have to deal with the irreducible external background due to elastic electron scattering of solar neutrinos. Muon interactions can produce high-energy secondaries such as neutrons or electromagnetic showers. Charged backgrounds can be easily eliminated using a veto system. Neutrons, on the other hand, are a more serious problem. They can have sizable penetrating power, impinging on the detector materials and activating them, ultimately resulting in radioactive nuclides. Detectors can be shielded against neutrons with layers of hydrogenous material. Cosmogenic activation is, of course, more severe on surface. Therefore, for experiments using materials that can get activated (like germanium or copper), underground fabrication and storage of the detector components may be essential.

Natural radioactivity in the rock of the underground caverns results in a gamma-ray flux that can interact in the detector producing background. Dense, radiopure materials such as lead or copper are used as shielding to
Figure 3.10. Total muon flux measured at various underground sites currently available to host physics experiments \cite{126, 127}. An empirical parameterization \cite{127} is shown as a dashed line. Facilities shown in red, blue and green are located, respectively, in North America, Europe and Asia. The full names and host countries of the facilities shown in the figure, from top to bottom, are the following: Waste Isolation Pilot Plant (WIPP), USA; Laboratorio Subterráneo de Canfranc (LSC), Spain; Soudan Underground Laboratory (SUL), USA; Kamioka Observatory (Kamioka), Japan; Boulby Palmer Laboratory (Boulby), United Kingdom; Laboratorio Nazionale del Gran Sasso (LNGS), Italy; Laboratoire Souterrain de Modane (LSM), France; Sanford Underground Science and Engineering Laboratory (SUSEL), USA; Baksan Neutrino Observatory (BNO), Russia; and Sudbury Neutrino Observatory Laboratory (SNOLAB), Canada.
attenuate this background. Water, being inexpensive and easy to purify, is also a good alternative for shielding against $\gamma$ rays.

Besides the passive background reduction techniques mentioned above, most experiments use now active methods for the discrimination of signal and background: reconstruction of the event topology, pulse-shape discrimination, combination of detection signatures, etc. A unique possibility offered by xenon-based experiments is that all backgrounds except the two-neutrino decay mode could be effectively removed by identification of the daughter barium ion using atomic laser resonant spectroscopy [130].

### 3.4.4 Detection efficiency

Neutrinoless double beta decay is extremely rare, if existent at all. A high detection efficiency is, therefore, an important requirement for a $\beta\beta$ experiment, as clearly stated by Eq. (3.21). To obtain the same increase in $m_{\beta\beta}$ sensitivity attained by doubling the efficiency, the mass would have to be increased by a factor of 4, assuming the same background. In general, the simpler the detection scheme, the higher the detection efficiency. For instance, pure calorimetric approaches such as germanium diodes or bolometers have detection efficiencies in excess of 80%. This is to be contrasted with experiments performing, for example, particle tracking, which will typically result in significant efficiency loss. Homogeneous detectors, where the source material is the detection medium, provide in principle higher efficiency than the separate-source approach for a number of reasons, including geometric acceptance or absorption in the $\beta\beta$ source. That being said, some homogeneous detectors may use part of the mass close to the detector boundaries for self-shielding against external backgrounds, paying it with efficiency loss.

### 3.4.5 Exposure

Thousands of kilograms of $\beta\beta$ source will be needed to explore the extremely long $0
\nu\beta\beta$ half-lives corresponding to the inverted hierarchy of neutrino masses. Most collaborations searching for $0\nu\beta\beta$ decay are advertising already future tonne-scale versions of their experiments. However, not all the technologies are equally suitable for that purpose. The scalability of each experimental technique will be, therefore, one of the key points — together with the detector performance at the 100-kg stage — for the evaluation of these proposals.
Large-scale production of the $\beta\beta$ isotopes will represent a technical and logistic challenge, as well as a significant fraction of the total cost of the detectors. Some of the most popular $\beta\beta$ isotopes are, in fact, quite rare in Earth. For example, the annual world production of germanium or tellurium is a few hundred tonnes, and much less, a few tens of tons, in the case of xenon. Moreover, with the exception of $^{130}\text{Te}$, which represents approximately one third of all tellurium, the isotopic abundance of the $\beta\beta$-decaying isotopes is around or below 10% (see Table 3.3), requiring isotopic enrichment in order to obtain large, concentrated masses. In fact, this has been so far the driving cost in the procurement of the isotopes. The most cost-effective enrichment technology is centrifugal separation, but it is only possible for elements with a stable gas compound. Affordable enrichment of large quantities of those species with no gas compound, such as $^{48}\text{Ca}$ or $^{150}\text{Nd}$, is not possible at present. Centrifugation of xenon, being a noble gas, is, of course, simpler (and hence cheaper) than that of metalloids such as germanium. Therefore, from this point of view, $^{136}\text{Xe}$ would be a particularly favourable isotope to use for a tonne-scale experiment.

### 3.5 Current generation of experiments

Three experiments of the current generation — EXO-200, KamLAND-Zen and GERDA — have been operating for a few years already, and at least five other — CUORE, MAJORANA, NEXT-100, SNO+ and SuperNEMO — plan to start taking data soon. All these experiments are described below, with the exception of NEXT, which is the focus of this work and will be discussed in depth in the following chapters. The descriptions here do not pretend to be exhaustive, but focused on the main features of each technique.

#### 3.5.1 CUORE

The *Cryogenic Underground Observatory for Rare Events* (CUORE) will search for the $0\nu\beta\beta$ decay of $^{130}\text{Te}$ using TeO$_2$ crystal bolometers. When these crystals are cooled to 10 mK, their heat capacity becomes so small that the energy deposited by interacting particles is measurable as a rise in temperature. The crystals, therefore, function as highly sensitive calorimeters. This technique was used for the first time in $0\nu\beta\beta$-decay searches by the MiBDDB and Cuoricino $^{130}\text{Te}$ experiments.
Figure 3.11. Cutaway view of the CUORE bolometers inside the cryostat, consisting of six nested copper vessels at 300 K (outer vacuum chamber), 40 K, 4 K (inner vacuum chamber), 0.6 K (still), 0.05 K (heat exchanger), and 0.01 K (mixing chamber). Several layers of radiopure lead shield the bolometers from external radiation. Reproduced from Artusa et al. (2014) [134].

CUORE, currently under construction at the Laboratori Nazionali del Gran Sasso, will consist of 988 bolometers arranged in 19 vertical towers held by a copper frame. The basic detector element is a $5 \times 5 \times 5$ cm$^3$ TeO$_2$ crystal of 750 g instrumented with a temperature sensor and a resistive heater. The total mass of the bolometers will be 741 kg, of which 206 kg are $^{130}$Te. The bolometer towers will be housed in a cryostat composed of six nested copper vessels (see Figure 3.11). Two cold lead shields will shield the bolometers from radiation originating in the cryostat: a layer of ancient Roman lead 6 cm thick located between the two middle copper vessels will shield the detectors from radioactivity in the outer vessels, and a disc 31 cm thick made of modern and Roman lead and located below the mixing chamber plate will shield the detectors from radioactivity in the overhead apparatus. The cryostat will be surrounded by a 73-tonnes octagonal external shield designed to screen
the detector from environmental $\gamma$ rays and neutrons. The shield has three layers: an outermost layer 20 cm thick consisting of a floor and sidewalls of polyethylene to thermalize and absorb neutrons; a side layer 2 cm thick of boric-acid powder to absorb neutrons; and an innermost layer of lead bricks of at least 25 cm thickness to absorb $\gamma$ rays.

CUORE aims at improving the sensitivity of Cuoricino ($§$ 3.3), its predecessor, by more than a factor of 30 by operating a larger, cleaner, better-shielded detector with enhanced energy resolution. The expected energy resolution (FWHM) of the CUORE crystals is 5 keV at the $Q$ value of $^{130}$Te (2528 keV) [134]. This resolution has already been achieved in tests performed during the R&D phase. In Cuoricino, the average background rate in the region of interest was 0.58 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. Three main contributions were identified [112]: (30$\pm$10)\% of the measured background in the region of interest was due to multi-Compton events due to the 2615-keV gamma ray from the thorium decay chain from the contamination of the cryostat shields; (10$\pm$5)\% was due to surface contamination of the crystals (primarily degraded alphas from the natural decay chains); and (50$\pm$20)\% is ascribed to similar surface contamination of inert materials surrounding the crystals, most likely copper. On the basis of this result, the R&D for CUORE has pursued two major complementary lines: the reduction of surface contamination and the selection of extremely radiopure construction materials. The goal is achieving a background rate in the region of interest of 0.03–0.04 cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$ [112].

CUORE-0, a single tower of CUORE containing 52 TeO$_2$ crystals, is in operation inside the Cuoricino cryostat since March 2013. It will directly test the level of backgrounds of the CUORE experiment and improve the Cuoricino sensitivity to the $^{130}$Te $0$\nu$\beta\beta$-decay half-life. CUORE is now in an advanced state of construction. The Collaboration plans to complete the integration and commissioning of the detector at the end of 2014, and commence data taking in the first half of 2015 [134].

3.5.2 EXO

The Enriched Xenon Observatory (EXO) is an experimental program searching for neutrinoless double beta decay using $^{136}$Xe. The first phase of the experiment, EXO-200 [135], consists in a 200-kg liquid xenon (LXe) time projection chamber that has been taking physics data at the Waste Isolation Pilot Plant (WIPP), in New Mexico, USA, since early May 2011. The results produced so far by the experiment include the first observation of the $2\nu\beta\beta$ decay of $^{136}$Xe [101].
3.5. CURRENT GENERATION OF EXPERIMENTS

and some of the most stringent limits so far on the effective Majorana neutrino mass \[ m_{\text{eff}} \] \cite{62, 115}. Building on the success of EXO-200, the EXO Collaboration has started the R&D work for a future multi-tonne LXe experiment called nEXO.

The EXO-200 detector is a cylindrical TPC, about 40 cm in diameter and 44 cm in length, with two drift regions separated in the centre by a transparent cathode. The TPC measures the 3D coordinates and energy of ionization deposits in the LXe by simultaneously collecting the scintillation light and the charge. Charge deposits spatially separated by about 1 cm or more are individually observed with a position accuracy of a few millimetres. A pair of crossed wire planes collects the ionization charge and measures its amplitude and transverse coordinates, and arrays of avalanche photodiodes (APDs) located behind the wire planes measure the scintillation light. The sides of the chamber are covered with teflon sheets that act as VUV reflectors improving the light collection. The xenon, enriched to 80.6\% in \(^{136}\)Xe, is held inside a thin copper vessel immersed in a cryofluid that also shields the detector from external radioactive backgrounds. The cryofluid is maintained at \(~167\) K inside
a vacuum-insulated copper cryostat. Further shielding is provided by at least 25 cm of lead in all directions. The entire assembly is housed in a clean-room located underground at WIPP. Four of the six sides of the clean-room are instrumented with plastic scintillator panels recording the passage of cosmic ray muons. Figure 3.12 shows the overall detector and shielding arrangement. The fiducial volume selected in the data analysis contains a $^{136}$Xe mass of 76.5 kg. Signal detection efficiency is estimated to be 84.6%. Thanks to the simultaneous measurement of the ionization and scintillation signals (see

Figure 3.13. Energy spectra recorded by the EXO-200 detector. Main panels show single-site (a) and multi-site (b) events (black points) with a zoom-in (inset) around $Q_{\beta\beta}$. Lower panels in a and b show residuals between data and best fit normalized to the Poisson error, ignoring empty bins. The vertical red lines in the SS spectra indicate the $\pm 2\sigma$ energy region of interest. Several background model components are indicated in the main panel of b to show their relative contributions to the spectra. Figure reproduced from Albert et al. (2014) [62].
3.5. CURRENT GENERATION OF EXPERIMENTS

Figure 4.6). EXO-200 reaches an energy resolution of 3.6% FWHM at the Q value of $^{136}$Xe. The data of the last published analysis [62], corresponding to an exposure of 100 kg yr (736 mol yr), is shown in Figure 3.13. The estimate of the background in a $\pm 2\sigma$ window around $Q_{\beta\beta}$ is $31.1 \pm 1.8$ (stat) $\pm 3.3$ (syst) counts, or $(2.1 \pm 0.3) \times 10^{-3}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. The dominant backgrounds arise from the thorium series (16.0 counts), the uranium series (8.1 counts) and $^{137}$Xe (7.0 counts). This amount of $^{137}$Xe is consistent with estimates from studies of the activation of $^{136}$Xe in muon-veto-tagged data. The EXO Collaboration reports a 90% CL lower limit on the half-life of $^{136}$Xe of $1.1 \times 10^{25}$ yr. This corresponds to an upper limit on the Majorana neutrino mass of 190–450 meV.

The EXO Collaboration is planning a new, next-generation detector, nEXO, that would use 5 tonnes of enriched xenon [136]. Many of the detector concepts and implementation in EXO-200 are, in principle, scalable to a larger mass of xenon, and the self-shielding improves with larger mass. The detector would be placed in a large water shield instead of the lead shield used for EXO-200 and a deeper site would be chosen to reduce the cosmogenic neutron backgrounds.

3.5.3 GERDA and Majorana

The GERmanium Detector Array (GERDA) experiment, located in Hall A of the Laboratori Nazionali del Gran Sasso (LNGS), is searching for the $0\nu\beta\beta$ decay of $^{76}$Ge using bare high-purity germanium (HPGe) diodes immersed in a cryogenic bath of liquid argon (LAr) [137]. The HPGe detectors are arranged in strings and mounted in special low-mass holders made of ultra-pure copper and PTFE. The strings are suspended inside a vacuum-insulated stainless steel cryostat of 4.2 m diameter and 8.9 m height filled with LAr. A copper lining 6 cm thick covers the inner cylindrical shell of the cryostat. The cryostat is placed in a 590-m$^3$ water tank instrumented with PMTs which serves as a Cherenkov muon veto as well as a gamma and neutron shield. A drawing of the entire system is shown in Figure 3.14.

The GERDA experiment was planned in two physics stages. Phase I, which started in November 2011 and ended in March 2013, has used refurbished semi-coaxial HPGe detectors from the Heidelberg-Moscow and IGEX experiments (§ 3.3), which are isotopically enriched to 86% in $^{76}$Ge, plus a non-enriched detector from the GENIUS-TF project [139], totalling a mass of 17.67 kg. In addition to these detectors, 5 broad-energy germanium (BEGe) diodes foreseen for the second phase of the experiment were deployed in July 2012. The exposure-averaged energy resolution (FWHM) of the detectors at the Q value
CHAPTER 3. THE SEARCH FOR NEUTRINOLESS DOUBLE BETA DECAY

Figure 3.14. Artist’s view (HPGe strings not to scale) of the GERDA detector at LNGS. Reproduced from Macolino (2013) [38].

of $^{76}$Ge is $(4.8 \pm 0.2)$ keV for the semi-coaxial detectors and $(3.2 \pm 0.2)$ keV for the BEGe detectors.

The GERDA Collaboration has published a measurement of the $2\nu\beta\beta$ half-life of $^{76}$Ge, $T^{2\nu}_{1/2} = 1.84^{+0.09}_{-0.08} \text{ (stat)} \times 10^{21} \text{ years}$ [89], and a limit to the $0\nu\beta\beta$ half-life, $T^{0\nu}_{1/2}(^{76}\text{Ge}) > 2.1 \times 10^{25} \text{ yr (90\% CL)}$ [116]. This analysis was done with 17.9 kg yr of exposure. The achieved background rate was $(11 \pm 2) \times 10^{-3} \text{ cts keV}^{-1} \text{ kg}^{-1} \text{ yr}^{-1}$ (see Figure 3.15), and the $0\nu\beta\beta$ signal acceptance was 86\%. The GERDA result is consistent with the limits by Heidelberg-Moscow and IGEX (§3.3). The combination of the results of the three experiments yields a limit of $3.0 \times 10^{25}$ years (90\% CL) [116]. Therefore, the long-standing claim for a $0\nu\beta\beta$ signal in $^{76}$Ge is strongly disfavoured.

In Phase II, besides the increase of the active mass by about 20 kg (30 BEGe detectors), the main goal is to further reduce the background by one order of magnitude thanks to several improvements in the detector setup (instrumentation of the LAr bath, materials of higher radiopurity in the vicinity of the detectors, etc.) [140].

For the very long term, it is foreseen a third phase of the experiment with about 1 tonne of $^{76}$Ge together with further reduction of background. Such an effort would be feasible only in a word-wide collaboration with the
Figure 3.15. The combined energy spectrum registered by the enriched HPGe detectors of GERDA before pulse-shape discrimination (open histogram) and after (filled histogram). In the upper panel, the spectrum zoomed to $Q_{\beta\beta}$ is superimposed with the expectations based on the $^{76}$Ge claim and with the limit derived from these data. Reproduced from Agostini et al. (2013) [116].

The Majorana project [141], which is following a more classic approach than GERDA in the design of a germanium-based experiment. The Majorana Collaboration is building a modular setup composed of two cryostats built from ultra-pure electroformed copper, with each cryostat capable of housing over 20 kg of HPGe detectors. The baseline plan calls for 30 kg of the detectors to be built from Ge material enriched to 86% in isotope 76 and 10 kg fabricated from natural Ge. Starting from the innermost cavity, the cryostats will be surrounded by an inner layer of electroformed copper, an outer layer of oxygen-free copper, high-purity lead, an active muon veto, polyethylene, and borated polyethylene. The cryostats, copper, and lead shielding will all be enclosed in a radon exclusion box. The entire experiment will be located in a clean room at the Sanford Underground Research Facility (SURF) in South Dakota, USA. The goal is to demonstrate a background rate of 3 counts per tonne and per year in the 4-keV wide region of interest. The detector should be in operation in 2015.
Figure 3.16. Schematic drawing of the KamLAND-Zen detector. Reproduced from Gando et al. (2012) [144].

3.5.4 KamLAND-Zen

The KamLAND-Zen experiment is searching for the $0\nu\beta\beta$ decay of $^{136}\text{Xe}$ using enriched xenon dissolved in liquid scintillator, a technique first proposed by R. Raghavan in 1994 [142]. The experiment reuses the neutrino KamLAND detector [143], located at the Kamioka Observatory, Japan. The KamLAND-Zen detector, shown in Figure 3.16, is composed of two concentric transparent balloons. The inner one, 3.08 m in diameter and fabricated from 25 $\mu$m thick nylon film, contains 13 tonnes of Xe-loaded liquid scintillator. The outer balloon, 13 m in diameter, contains 1 kilotonne of pure liquid scintillator, and serves as an active shield for external gamma background as well as a detector for internal radiation from the inner balloon. Buffer oil between the outer balloon and an 18 m diameter spherical stainless-steel containment tank shields the detector from external radiation. Scintillation light is recorded by 1325 17-in and 554 20-in photomultiplier tubes mounted on the stainless-steel tank, providing 34% solid-angle coverage. The containment tank is surrounded by a 3.2-kt water-Cherenkov outer detector. The Xe-loaded scintillator consists of 82% decane and 18% pseudocumene by volume, 2.7 g/litre of the fluor PPO, and 2–3% by weight of enriched xenon gas [144], corresponding to approximately 350 kg of $^{136}\text{Xe}$. 
3.5. CURRENT GENERATION OF EXPERIMENTS

KamLAND-Zen, which has been collecting physics data since late 2011, has published a measurement of the half-life of the $2\nu\beta\beta$ decay of $^{136}\text{Xe}$, $2.38 \pm 0.02 \text{ (stat)} \pm 0.14 \text{ (syst)} \times 10^{21}$ years \cite{144}, and a limit to the half-life of the $0\nu\beta\beta$ decay, $2.6 \times 10^{25}$ years (90% CL) \cite{63,102}. The energy resolution of the detector is 9.9% FWHM at the $Q$ value of $^{136}\text{Xe}$. The achieved background rate in the region of interest is approximately $1.4 \times 10^{-4}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, thanks to a tight selection cut in the fiducial volume and the identification of $^{214}\text{Bi}$ events via Bi-Po tagging \cite{63}.

3.5.5 SNO+

SNO+, the follow-up of the Sudbury Neutrino Observatory (SNO) \cite{145}, is a multipurpose liquid scintillator experiment housed in SNOLAB (Ontario, Canada). The detector, shown in Figure 3.17, reuses many of the components of its predecessor, replacing the heavy water by 780 tonnes of liquid scintillator in order to obtain a lower energy threshold. The detector consists of a 12 m diameter acrylic vessel surrounded by about 9500 8-in photomultiplier tubes that provide a 54% effective photocathode coverage. The acrylic vessel is immersed in a bath of ultra pure water that fills the remaining extent of the underground cavern, attenuating the background from external media such as the PMTs and surrounding rock. The density of the liquid scintillator (0.86 g/cm$^3$) being lower than that of the surrounding water leads to a large buoyant force on the acrylic vessel. To keep it in place, a hold-down rope net has been installed over the detector and anchored to the cavity floor.

The physics program of SNO+ includes the search for neutrinoless double beta decay in $^{130}\text{Te}$, which will be loaded into the liquid scintillator in the form of (non-enriched) telluric acid. A loading of 0.3%, equivalent to 780 kg of $^{130}\text{Te}$, is planned for the first phase of the experiment, which will start towards the end of 2015 or beginning of 2016.

The energy resolution of the SNO+ detector is expected to be 10.5% FWHM at the $Q$ value of $^{130}\text{Te}$ \cite{146}. Consequently, the $2\nu\beta\beta$ spectrum will be an important source of background. The expected levels of uranium and thorium in the liquid scintillator can also result in substantial activity near the $0\nu\beta\beta$ endpoint, mostly from the decays of $^{214}\text{Bi}$ and $^{212}\text{Bi}$. Nevertheless, these can be, in principle, actively suppressed via Bi-Po $\alpha$ tagging \cite{146}. External backgrounds (not originating in the liquid scintillator) can be suppressed with a tight fiducial volume selection, which will cut, however, about 70–80% of the signal.
3.5.6 SuperNEMO

The new instalment of the NEMO detector series (§ 3.3), SuperNEMO, will consist of up to 20 modules 6.2 m long, 4.1 m high and 2.1 m wide, each one containing a thin source foil of 5–7 kg of $^{82}$Se ($^{150}$Nd and $^{48}$Ca are also contemplated, in case their isotopic enrichment becomes feasible) [147]. A drawing of a SuperNEMO module can be seen in Figure 3.18. Each module will consist of a central source foil, 3 m long and 4.5 m high, with a surface density of about 40 mg/cm$^2$, placed in between two tracking chambers — drift cells operating in Geiger mode — surrounded by calorimeter walls made of plastic scintillator blocks coupled to PMTs.

The expected sensitivity of SuperNEMO to the 0νββ half-life of $^{82}$Se is at the level of $1 \times 10^{26}$ years for an exposure of 500 kg \cdot yr [148]. Achieving this implies several significant improvements over the performance of the NEMO-3 detector. The energy resolution of SuperNEMO is expected to be 4% FWHM at 3 MeV, almost a factor of 2 better than in NEMO-3 [148]. The projected...
background rate in the energy region around $Q_{\beta\beta}$ is $5 \times 10^{-5}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, about 25 times better than in NEMO-3, requiring $^{214}$Bi and $^{208}$Tl impurities in the source foils to be reduced to less than 2 $\mu$Bq/kg and 10 $\mu$Bq/kg, respectively \cite{148}. A dedicated setup, the BiPo detector \cite{149}, installed in the Laboratorio Subterráneo de Canfranc, will measure the radiopurity of the SuperNEMO source foils to make sure that the mentioned levels are achieved.

The first SuperNEMO module, called the Demonstrator, will house a slightly denser source of approximately 7 kg of $^{82}$Se. The construction of the Demonstrator has already started, and the beginning of its operations is projected for 2015. It will be installed in the former location of the NEMO-3 apparatus. Future modules will be located in the recently approved extension of the Modane laboratory.

**Figure 3.18.** Exploded view of a SuperNEMO module. The source foil (A), with approximately 5 kg of $^{82}$Se, is sandwiched between two tracking chambers (B) and two calorimeter walls (C).
High-pressure xenon gas for double beta decay searches

4.1 Introduction

In the previous chapter we concluded that the difficult challenge of exploring the neutrino-mass region corresponding to the inverse hierarchy requires experiments that can be extrapolated to large masses of $\beta\beta$ emitter, in the range of a few tonnes, while keeping the background rate as low as possible, typically well below 1 count keV$^{-1}$ tonne$^{-1}$ year$^{-1}$. Most of the existing techniques may not be suitable for this purpose, either because of the difficulty of reaching large detector masses, or due to their intrinsic poor energy resolution or limited background-rejection capabilities. In contrast, a xenon gas experiment, we will argue in this chapter, offers a priori the potential to build a near-ideal experiment.

Two naturally-occurring isotopes of xenon, $^{134}$Xe and $^{136}$Xe, can undergo $\beta\beta$ decay. The latter, having a higher Q value (2458 keV [119][150]), is preferred for $0\nu\beta\beta$-decay searches. It constitutes only 8.86% of natural xenon, but the enrichment process is relatively simple and cheap compared to that of other $\beta\beta$ isotopes. The two-neutrino decay mode of $^{136}$Xe is slow, $2.2 \times 10^{21}$ years [101][102], and hence the experimental requirement for energy resolution is less severe than for other $\beta\beta$ sources. Furthermore, xenon is a suitable detection medium with strong scintillation and ionization primary signals. In its gaseous phase, xenon can provide very good energy resolution, better, in principle, than 0.5% at the Q value of $^{136}$Xe [151].

The use of xenon gas detectors for $0\nu\beta\beta$-decay searches is actually an old idea. The Milano experiment [152], running at LNGS in the late 1980s, made use of a multi-wire proportional chamber filled with xenon gas (enriched to 64% in $^{136}$Xe) at a pressure of 9.5 bars. The detector, with an active volume of 79 litres, contained about 4.4 kg of xenon (32.3 moles of $^{136}$Xe). The detection efficiency, computed with Monte Carlo simulations, was $\sim 35\%$, and the measured energy
CHAPTER 4. HIGH-PRESSURE XENON GAS FOR DOUBLE BETA DECAY SEARCHES

Figure 4.1. Reconstructed particle tracks in the Gotthard time projection chamber [153]. Left: A typical $\beta\beta$-decay candidate: a continuous track with *blobs* at both ends. The $xz$ and $yz$ projections, as well as the extracted $xy$ projection in the lower frame, are represented; the time evolution of the anode signal is displayed beside them. Scales are in cm. Right: $\beta$-$\alpha$ coincidence: a single-electron track is followed by an alpha emission (with a characteristic drop-like track and fast anode pulse) 28 $\mu$s later, at the same $x$ and $y$ position. The $\beta$ particle is emitted in the direction of increasing $z$.

resolution was 4.5% FWHM at 2.5 MeV. The background rate in the $0\nu\beta\beta$ energy window was $\sim$11 counts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, 4 orders of magnitude worse than current detectors. After accumulating almost 2 kg·yr of exposure, no evidence of any of the $\beta\beta$ decay modes was found in the data. Instead, the following limits to the decay half-lives were reported: $T^{2\nu}_{1/2} > 6.0 \times 10^{19}$ years and $T^{0\nu}_{1/2} > 2.0 \times 10^{22}$ years, both at 90% CL.

The other pioneer experiment was the Gotthard TPC [153,154], built by a Caltech-PSI-Neuchâtel collaboration and operated at the St. Gotthard road tunnel (Switzerland) in the 1990s. The key idea of the experiment was the use of the tracking capabilities of the TPC to identify the detected events as signal or background according to their energy deposition pattern ($dE/dx$), which is unique to a given particle type. For example, Figure 4.1 (left) shows a typical $\beta\beta$-candidate event as registered in the Gotthard TPC: a continuous track with extra depositions (*blobs*) at both ends. Single-track events were selected and scanned visually to classify them as signal or background according to their energy-deposition pattern. The rejection efficiency of single electrons was estimated to be above 98%.
The detector was filled at a pressure of 5 atm with a 96:4 mixture of xenon gas (enriched to 62.5% in $^{136}$Xe) and methane, which acted as a quencher increasing the drift velocity and reducing the diffusion of the ionization electrons. The active volume of the detector, of about 180 litres, contained 3.3 kg (24.2 moles) of $^{136}$Xe. The TPC readout system was a classic set of wire grids located behind the anode. The detector had no start-of-event trigger, and hence the absolute longitudinal position of the tracks could not be reconstructed. The achieved energy resolution, 6.5% FWHM at 2500 keV, was rather modest for xenon. According to simulations, the probability for a $\beta\beta$ event at 2.5 MeV to be completely contained in the active volume of the TPC was 30%. The measured background rate in the region of interest was about 0.01 counts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, slightly higher than that registered by the Heidelberg-Moscow experiment (§3.3). The dominant background was concluded to be Compton-scattered electrons from natural radioactivity misidentified as two-electrons events. The following limits to the $2\nu\beta\beta$ and $0\nu\beta\beta$-decay half-lives were reported: $T_{1/2}^{2\nu} > 3.6 \times 10^{20}$ years and $T_{1/2}^{0\nu} > 4.4 \times 10^{23}$ years at 90% CL.

The Gotthard TPC demonstrated the effectiveness of tracking in xenon to discriminate signal from background. However, the sensitivity of the experiment was limited by its poor energy resolution. In this chapter we will present the most important physical and operational parameters affecting the energy resolution in gaseous xenon, with some comparisons to the liquid phase. Using arguments based on these experimental factors, we will present a detector design chosen to optimize the energy resolution while keeping the powerful tracking signature.

### 4.2 Gaseous xenon as detection medium

#### 4.2.1 Primary signals in xenon: scintillation and ionization

All radiation detectors are based on the same fundamental principle: the transfer of part or all of the radiation energy to the detector mass, where it is converted into some other form more adequate for human perception or electronic processing. In the case of a gaseous detector, charged particles transfer their energy to the atoms of a gas through two processes: excitation and ionization. In the former, the energy absorbed by the atom causes an electron to raise to a higher energy level, whereas in the latter, the transferred energy is high enough to remove the electron from the atom, resulting in the formation of an electron-ion pair. Recombination of these ionization pairs
CHAPTER 4. HIGH-PRESSURE XENON GAS FOR DOUBLE BETA DECAY SEARCHES

Figure 4.2. Main processes responsible for the ionization and scintillation primary signals in xenon (the symbol X indicates ionizing radiation). Reproduced from Álvarez et al. [155].

and atomic de-excitation processes both lead to the emission of scintillation photons with characteristic energies. All these processes are summarized in Figure 4.2 for the case of xenon gas.

The transfer of the radiation energy into measureable ionization and scintillation can be expressed mathematically in the form of an energy-balance equation, as originally proposed by Robert L. Platzman [156]:

$$ E_0 = N_i \langle E_i \rangle + N_{ex} \langle E_{ex} \rangle + N_i \langle E \rangle. \quad (4.1) $$

In this equation, $E_0$ is the energy absorbed by the gas, $N_i$ is the number of ionization pairs produced with an average energy expenditure $\langle E_i \rangle$, $N_{ex}$ is the number of atoms excited at an average energy expenditure $\langle E_{ex} \rangle$, and $\langle E \rangle$ is the average kinetic energy of sub-excitation electrons (i.e. electrons with energy lower than the excitation potential), which ultimately is released as heat.

Both the electron-ion pairs and the scintillation photons can be used to acquire information on the identity, energy and kinematics of the ionizing radiation. Consequently, the suitability of a gas as a detection medium depends, among other things, on the amplitude and variance of those signals. The
ionization and scintillation amplitudes are usually expressed in terms of the average energies required to produce, respectively, an electron-ion pair ($W_i$) and a scintillation photon ($W_{sc}$). From Equation (4.1), $W_i$ can be written as

$$W_i \equiv \frac{E_0}{N_i} = \langle E_i \rangle + \frac{N_{ex}}{N_i} \langle E_{ex} \rangle + \langle \mathcal{E} \rangle.$$

This so-called ionization $W$ value depends weakly on the type and energy of the radiation, but is sensitive to the density of the media, as shown in Figure 4.3. The ratio of $W_i$ to the ionization potential of the atom, $I$, is between 1.7 and 1.8 for noble gases (at low pressure), and slightly smaller, around 1.65, for their liquid phases. All of the quantities on the right-hand side of Eq. (4.2) can be evaluated in comparison with $I$ using information that is independent of the absolute measurement of $W_i$. First, the average energy lost in the ionization
Table 4.1. Detection properties of noble gases and their liquid phases: ionization potential, $I$; average required to produce an electron-ion pair, $W_i$; average energy required to produce a scintillation photon, $W_{sc}$; and Fano factor, $F$. Values taken from Aprile et al. [158].

<table>
<thead>
<tr>
<th>Element</th>
<th>$I$ (eV)</th>
<th>$W_i$ (eV)</th>
<th>$W_{sc}$ (eV)</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>24.587</td>
<td>41.3</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Neon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>21.565</td>
<td>29.2</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Argon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>15.75</td>
<td>26.4</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>13.4</td>
<td>23.6</td>
<td>19.5</td>
<td>0.107</td>
</tr>
<tr>
<td>Krypton</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>14.00</td>
<td>24.2</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>11.55</td>
<td>18.4</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Xenon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>12.13</td>
<td>22.1</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>11.67</td>
<td>15.6</td>
<td>13.8</td>
<td>0.041</td>
</tr>
</tbody>
</table>

The detection process, $\langle E_i \rangle$, is somewhat larger than $I$ because of the energy wasted in the production of excited ions and multiply-charged ions. Second, the average energy spent in an excitation process, $\langle E_{ex} \rangle$, is not much smaller than $I$ since all atomic levels in noble gases lie fairly close to the ionization limit. The ratio $N_{ex}/N_i$, however, is about 0.4–0.6 in the gaseous phase and 0.1–0.2 in the liquid phase, making the second term on the right-hand side of Eq. (4.2) comparatively small. Finally, the energy transferred to sub-excitation electrons, $\langle \mathcal{E} \rangle$, is about 30% of the ionization potential. These estimations agree well with the experimental measurements of $W_i$, shown in Table 4.1. Xenon has the smallest $W$ value, hence the largest ionization yield, of all noble gases.

Analogously to Eq. (4.2), we can define $W_{sc}$ as

$$ W_{sc} \equiv E_0 / N_{sc} $$

where $N_{sc}$ is the total number of scintillation photons produced:

$$ N_{sc} = r N_i + N_{ex} = N_i \left( r + N_{ex}/N_i \right). $$
Here, \( r \) is the recombination fraction of electron-ion pairs. Substituting Equations (4.2) and (4.4) into Eq. (4.3), we obtain

\[
W_{sc} = \frac{E_0}{N_i (r + N_{ex}/N_i)} = \frac{W_i}{r + N_{ex}/N_i}.
\]  

(4.5)

The scintillation \( W \) value depends strongly on the linear energy transfer (LET) of the ionizing radiation, that is, the amount of energy transferred on average to the medium per unit length for a given particle type. A higher LET translates into a higher density of ionization pairs, and this, in general, into a higher probability of recombination \([158]\). The presence of an external electric field in the detector also affects the recombination process \([155, 158]\). Very intense fields prevent recombination \((r \approx 0)\), whereas in the absence of a field, all freed electrons and ions eventually recombine \((r \approx 1)\). These two situations correspond, respectively, to the minimum and maximum scintillation yields, that is, the maximum and minimum possible values of \( W_{sc} \).

Several measurements of \( W_{sc} \) for xenon gas are available in the literature. The value \((76 \pm 12) \text{ eV}\) was reported by Parsons and collaborators \([159]\) for a 90:10 xenon-helium mixture at 15 bar and for a drift field of order 0.1 kV cm\(^{-1}\) bar\(^{-1}\), with similar results for pure xenon. Do Carmo et al. \([160]\) measured \( W_{sc} = (111 \pm 16) \text{ eV} \) at 1 bar pressure and 0.35 kV cm\(^{-1}\) bar\(^{-1}\) drift field, whereas Fernandes et al. \([161]\) obtained the value \((72 \pm 6) \text{ eV}\) for gas pressures in the 1–3 bar range and for drift fields in the 0.15–0.6 kV cm\(^{-1}\) bar\(^{-1}\) range, with no significant variations with pressure and field. Finally, Resnati and collaborators \([161]\) measured \( W_{sc} = (26 \pm 7) \text{ eV} \) at 40 bar pressure and no electric field. The difference between these results is presently not fully understood, as it can be attributed only partially to the different gas density and drift field conditions. Furthermore, the theoretical estimations of the ratio \( N_{ex}/N_i \) appear to be incompatible with the experimental measurements of \( W_{sc} \). Clearly, more experimental data are needed in order to reach a complete understanding of the scintillation yield of gaseous xenon.

The scintillation spectra of noble gases extend from the infrared to the far ultraviolet. They are the result of a complex system of discrete atomic lines, bands and continua originating from many excited states and from various collision and transfer processes \([158, 163]\). At atmospheric pressure and above, the vacuum-ultraviolet (VUV) emission dominates the spectra (see Figure 4.4). Noble gases are relatively fast scintillators, with a slow decay component of tens of nanoseconds and a fast component of a few nanoseconds \([158]\).
CHAPTER 4. HIGH-PRESSURE XENON GAS FOR DOUBLE BETA DECAY SEARCHES

Figure 4.4. Scintillation spectra of noble gases and their condensed phases in the vacuum ultraviolet region. \[158\] [162].

4.2.2 Detection of the ionization and scintillation signals

The VUV scintillation photons emitted by noble gases are difficult to detect because they are strongly absorbed by most materials. Therefore, the VUV light is usually shifted to the visible or near-visible band using photo-fluorescent coatings deposited onto the surfaces exposed to the gas volume. At these wavelengths, the efficiency of photosensors — devices that transform an optical signal into an electric current through the photoelectric effect — is optimal, and highly transparent or reflective materials are available. Popular wavelength shifting (WLS) substances used in xenon detectors are p-terphenyl (TPH) and tetraphenyl butadiene (TPB) \[158\].

The detection of the ionization signal generally involves the transport of the ionization pairs through the gas under the influence of an external electric field. In this process, known as drift, the electrons and ions liberated by ionizing radiation are accelerated along the electric-field lines in opposite directions towards, respectively, the anode and cathode. The acceleration is intermittently interrupted by collisions with the gas atoms, limiting the maximum average velocity attainable by the drifting charges. For sufficiently low fields, the drift velocity of electrons, \( v_d \), is proportional to the field strength. At high fields, the drift velocity saturates, becoming independent of the field. The drift velocity
of ions is several orders of magnitude smaller than that of the electrons.

During the drift, due to the collisions with the gas atoms, the ionization charges deviate from the trajectories defined by the field lines spreading gaussianly in both the longitudinal and transverse directions. This diffusion limits the intrinsic position resolution of gaseous detectors. The magnitude of the spread is proportional to the drift time, \(t_d\):

\[
\sigma_L = \sqrt{D_L t_d}, \quad \sigma_T = \sqrt{D_T t_d},
\]

where \(D_L\) and \(D_T\) are, respectively, the longitudinal and transverse diffusion coefficients of the gas.

Electron attachment to electronegative impurities dissolved in the gas may lead to a significant decrease of the ionization signal during drift. This effect can be described, in general, by an exponential distribution:

\[
N(t_d) = N(0) \exp\left(-{t_d}/\tau\right),
\]

where \(N\) is the number of drifting electrons, which is a function of the drift time, \(t_d\), and \(\tau\) is the electron lifetime in the gas, inversely proportional to the concentration of impurities. Sufficiently long electron lifetimes can be achieved by circulating the gas continuously through appropriate filters.

In most applications, the total ionization charge collected is too low to be distinguished by the electronics or leads to poor signal-to-noise ratio. In these cases, the primary-electron signal can be amplified using electric fields of higher intensity than those typically applied for the drift. If the amplification field is such that the electrons gain energy above the excitation threshold of the gas but below the ionization threshold, they will excite gas atoms along their drift path. These will decay later, emitting light known as electroluminescence (EL) or secondary scintillation. In this way, each primary ionization electron produces a measurable, proportional optical signal (see Figure 4.5). If the amplification field is intense enough for the electrons to gain energy above the ionization threshold, they will produce new electron-ion pairs while drifting. The secondary ionization electrons can also gain energy from the field and ionize further atoms. Eventually, a charge avalanche is generated, resulting in an electron yield orders of magnitude higher than the number of primary ionization electrons. This charge signal is then large enough to be picked up by the electronics.

Detector designs exploiting electroluminescence in noble gases were developed rapidly in the last third of the 20th century, especially in Coimbra.
CHAPTER 4. HIGH-PRESSURE XENON GAS FOR DOUBLE BETA DECAY SEARCHES

Figure 4.5. Reduced electroluminescence yield \((Y/p)\) of gaseous xenon as a function of the reduced electric field, \(E/p\). The results of Garfield/Magboltz microscopic simulations \cite{164} are compared to the experimental measurements by Monteiro et al. \cite{165}. A linear fit to the simulation data points is also shown. The EL yield is linearly proportional to the reduced electric field above a threshold of 0.83 kV cm\(^{-1}\) bar\(^{-1}\) and up to approximately 5.5 kV cm\(^{-1}\) bar\(^{-1}\), where secondary-ionization effects become visible. Figure redrawn from Oliveira et al. \cite{164}.

Intrinsic energy resolution in xenon

The ionization signal is a good measure of the energy deposited in the medium, given that \(W_i\) is almost independent of the type and kinetic energy of the (Portugal) \cite{166,167} and at CERN \cite{168}. High-resolution spectrometry of low-energy X-rays was the main motivation. These detectors have had traditionally very little application in particle physics, with the notable exception of two-phase liquid noble gas detectors for direct WIMP searches \cite{158}. In contrast, avalanche-based detectors have been ubiquitous in high-energy physics experiments since the introduction of the multi-wire proportional chamber (MWPC) in 1968 \cite{169}. Nowadays, micro-pattern devices such as GEMs \cite{170} or Micromegas \cite{171} have largely replaced wire chambers thanks to their superior performance.

4.2.3 Intrinsic energy resolution in xenon

The ionization signal is a good measure of the energy deposited in the medium,
ionizing radiation. Nevertheless, stochastic fluctuations in the number of electron-ion pairs produced limit the energy resolution that can be achieved in a gaseous detector. If each ionization could be considered independent of the others, the fluctuations would then be described by a Poisson distribution with variance \( \sigma_i^2 = N_i \), where \( N_i \) is the mean number of electron-ion pairs produced for a given deposited energy. However, Ugo Fano demonstrated \[72\] that the processes leading to the creation of ionization pairs are not independent, and that, in general, the associated fluctuations can be described by the variance

\[
\sigma_i^2 = F N_i,
\]

(4.8)

where the number \( F \), known as Fano factor, depends on the detection medium. The ultimate energy resolution (FWHM) achievable with an ionization detector — often called intrinsic resolution or Fano limit — is, therefore, given by

\[
\delta E/E = 2.35 \sqrt{F N_i W_i / E} = 2.35 \sqrt{F W_i / E}
\]

(4.9)

where \( E = N_i W_i \) is the total energy deposited by radiation.

The Fano factor of noble gases is relatively well understood. For xenon, a number of experimental measurements report values between 0.13 and 0.17 \[73\]–\[75\], in good agreement with Monte Carlo calculations \[76\]. Conversely, for the liquid phase, the best energy-resolution measurements \[77\]–\[78\] imply a Fano factor \( F > 20 \), far worse than Poissonian and in complete contradiction with the theoretical calculations \[79\], which predict an intrinsic resolution for liquid xenon at least 3.5 times better than that of the gaseous phase. Although the reasons for this discrepancy are not entirely clear \[158\]–\[160\], a plausible explanation is that at higher densities the fluctuations associated to the recombination become more important than the fluctuations in the generation of the ionization pairs \[151\]–\[158\]. The energy in recombination is returned largely as scintillation light, but some unknown fraction of energy is converted to heat through excimer-excimer collisions. With the imposition of an electric field, which makes the recombination fraction somewhat smaller, and the simultaneous measurement of the scintillation and ionization signals, which are strongly anti-correlated (see Figure 4.6), the energy measurement can be improved \[177\], but large fluctuations persist.

Bolotnikov and Ramsey reported \[157\] the dependence with the density of the energy resolution obtained in a xenon gas detector by measuring the ionization signal generated by 662-keV gamma rays from \(^{137}\)Cs (see Figure 4.7). For
the density range between 0.12 and 0.6 g/cm³, they measured an approximately-constant resolution, $\delta E/E = 0.6\%$ FWHM at 662 keV, a value close to the Fano limit. For xenon densities above 0.55 g/cm³, however, the energy resolution quickly deteriorated, approaching the typical values measured for liquid xenon. It has been suggested [151,158] that this behaviour may be due to the appearance, as density increases, of globs of liquid xenon coexisting with the gas. Below the apparent density threshold at 0.55 g/cm³, the liquid phase fraction would be insufficient to contribute measurably to the processes introducing anomalous fluctuations. With increasing density, the liquid phase fraction would grow, and so would the impact of anomalous fluctuations related to recombination, reaching a maximum at the density of the liquid phase.

### 4.2.4 Impact of amplification fluctuations on energy resolution

An ideal gaseous detector would provide an exact measurement of the number of ionization electrons ($N_i$) generated by interacting radiation, returning a
Figure 4.7. Density dependence of the energy resolution (% FWHM) measured in xenon using the ionization signal generated by 662-keV gamma rays from $^{137}$Cs. As the photo-conversion peak was asymmetric, the resolution was extracted from its upper half only, and the electronic noise was subtracted in quadrature. Reproduced from Bolotnikov and Ramsey (1997) [157].

A perfectly proportional electric signal $S = N_i s_i$, where $s_i$ is the signal generated by one ionization electron. In practice, however, all detectors are affected by losses, noise and fluctuations that may worsen the intrinsic energy resolution of the gas (§ 4.2.3). It seems reasonable to assume that any possible fluctuations associated to the signal $s_i$ are uncorrelated with those intrinsic to the ionization process, described by Eq. (4.8). In that case, the overall variance of the detection process can be written as

$$
\sigma_E^2 = \sigma_i^2 + \sigma_G^2 = F N_i + G N_i = (F + G) N_i,
$$

(4.10)

where $G$ is defined as the variance of signal $s_i$. Therefore, the energy resolution (FWHM) of a gaseous detector, including the effect of fluctuations in the detection, is given by

$$
\delta E/E = 2.35 \sqrt{(F + G)} \frac{W_i}{E}.
$$

(4.11)

Although the factors $G$ and $F$ enter the above equation in a symmetric way, they are fundamentally different: $G$ reflects the impact of statistical fluctuations
in the detection process for a single electron, while $F$ reflects a constraint on fluctuations in energy partitioning for a fixed total energy [151].

Let us examine the value of $G$ for the two amplification schemes of the ionization signal discussed in §4.2.2, namely charge avalanche and electroluminescence. In both cases, the detection process consists of a number of steps in which an input signal $s_k$ is transformed into an output signal $s_{k+1}$. These signals are characterized by a probability distribution with mean $\mu(s_k)$ and variance $\sigma(s_k)$. It can be demonstrated [181] that the overall energy resolution of the detection process is given by the quadratic sum of the resolution of each step, $\sigma/\mu$, evaluated at the mean value of their input signal:

$$\delta E/E = 2.35 \left[ \sum_k \left( \frac{\sigma_k(\mu_{k-1})}{\mu_k(\mu_{k-1})^2} \right)^2 \right]^{1/2}. \quad (4.12)$$

The generation of the ionization pairs, common to both amplification schemes, is a zero-order step with no input signal and, recalling the discussion in §4.2.3 an output signal with mean value $\mu_0 = N_i$ and variance $\sigma_0^2 = F N_i$.

In a detector using charge-avalanche amplification, each ionization electron is, on average, multiplied by a factor $A$. This electron multiplication is subject to important fluctuations (besides those of the ionization process) associated to the spatial development of the charge avalanche, since the number of ionization pairs generated depends strongly on the length of the drift paths followed by the electrons across the amplification region of the detector [182]. For this reason, detector geometries with shorter average amplification drift paths (and higher potential gradients), such as those found in micro-pattern gaseous detectors, present lower avalanche fluctuations [183,184]. In the absence of other effects that may add to the avalanche fluctuations (such as attachment or avalanches initiated by scintillation photons) and assuming that each primary ionization electron develops an independent avalanche, the output signal of the amplification process has, then, mean value $\mu = N_i A$ and variance $\sigma^2 = N_i \sigma_A^2$, where $\sigma_A^2$ is the variance of the multiplication factor $A$. Therefore, according to Eq. (4.12), the energy resolution of a gaseous detector with charge-avalanche amplification is given by

$$\delta E/E = 2.35 \left( \frac{F N_i}{N_i^2} + \frac{N_i \sigma_A^2}{N_i^2 A^2} \right)^{1/2} = 2.35 \left( \frac{F}{N_i} + \frac{1}{N_i A^2} \sigma_A^2 \right)^{1/2}. \quad (4.13)$$

Comparing Eq. (4.13) with Eq. (4.11), we find that

$$G = \left( \frac{\sigma_A}{A} \right)^2. \quad (4.14)$$
4.2. GASEOUS XENON AS DETECTION MEDIUM

That is, for an avalanche-based detector, the factor $G$ that describes the fluctuations in the amplification of the ionization signal is equal to the single-electron multiplication resolution. Under common operating conditions, $G$ is larger than the Fano factor of gaseous xenon, with typical values between 0.2 and 0.9 [183, 185].

Consider now the case of a detector using electroluminescence for the amplification of the ionization signal. Each electron drifting through the amplification region generates an average of $Y$ photons (see Figure 4.5). The fluctuations in this optical multiplication factor can be described by a Fano-like distribution with variance $\sigma_Y^2 = J_{CP} Y$, where $J_{CP}$ is the so-called Conde-Policarp factor. Therefore, the first step in the amplification — the production of the electroluminescence light — has a mean output signal $\mu_{EL} = N_i Y$ and variance $\sigma_{EL}^2 = N_i \sigma_Y^2 = N_i J_{CP} Y$.

Due to practical limitations (absorption losses in the detector materials, finite photosensor coverage, etc.), only a small fraction of the emitted EL photons is typically detected. This step in the detection process can be modelled by a binomial distribution with mean value $\mu_{pe} = \eta N_i Y$ and variance $\sigma_{pe}^2 = \eta N_i Y (1 - \eta)$, where $\eta$ is defined as the probability of an EL photon to be detected by the photosensors. This probability is usually very small ($\lesssim 1\%$), and includes both the light collection efficiency of the system (i.e. the probability that an EL photon propagates from the emission point to the surface of a photosensor) and the detection efficiency of the photosensors themselves.

A photon can only be detected by a photosensor if it liberates an electron from its active surface via the photoelectric effect. When this occurs, a charge avalanche develops inside the photosensor, multiplying the initial charge by several orders of magnitude. This process, therefore, is analogous to that described above for the avalanche-based detector, and can be characterized by the charge resolution of the photosensor, $\sigma_q/q$.

Adding all of the above contributions together, the resolution of an EL-based detector is given by

$$\delta E/E = 2.35 \left( \frac{F}{N_i} + \frac{J_{CP}}{N_i Y} + \frac{1 - \eta}{\eta N_i Y} + \frac{(\sigma_q/q)^2}{\eta N_i Y} \right)^{1/2}. \quad (4.15)$$

The factor $G$ can thus be written in this case as

$$G = \frac{J_{CP} - 1}{Y} + \frac{(\sigma_q/q)^2 + 1}{\eta Y}. \quad (4.16)$$
The condition that we seek for near-intrinsic energy resolution in gaseous xenon is $G < F = 0.15$, which can be translated into a condition for the number of EL photons that must be detected per ionization electron, $\eta Y$. The Conde-Policarpo factor, $J_{CP}$, is much smaller than 1 over a wide range of operating conditions (see Figure 4.8). Therefore, the first term on the right-hand side of Eq. (4.16) can be made much smaller than the second one. For most modern PMTs, $(\sigma_q/q)^2$ is around 0.5; sometimes much less. Thus, in order to obtain $G < 0.15$, the product $\eta Y$ must be higher than 10. In other words, the system has to detect at least 10 photons per primary ionization electron.

### 4.3 The SOFT detector concept

A xenon gas detector optimized for neutrinoless double beta decay searches can provide both very good energy resolution (§§4.2.3 and 4.2.4) and, as proven by the Gotthard experiment (§4.1), a powerful discriminator between signal and
4.3. THE SOFT DETECTOR CONCEPT

**Figure 4.9.** Working principle of a TPC: the ionization electrons produced along the path of a charged particle (represented by the red trace) are drifted under the influence of an external electric field towards an amplification and readout plane that registers the signal amplitudes and transverse positions. The longitudinal coordinate is obtained from the arrival time of the signals.

background based on their distinct energy deposition pattern (see Figure 4.1). Gotthard made use of a *time projection chamber* (TPC), a gaseous detector invented by David Nygren in the 1970s [186][187], for the reconstruction of the ionization tracks left in the xenon by charged particles. Today, the TPC is present in a large number of nuclear and particle physics experiments [188], including some of the leading rare-event searches [135][189][190]. Its working principle is illustrated in Figure 4.9. A uniform electric field is established along the symmetry axis of a cylindrical chamber filled with a gaseous (or liquid) detection medium. The ionization electrons produced along the path of a charged particle traversing the cylinder are drifted under the influence of the field towards an amplification and readout plane with 2D spatial segmentation. This plane registers the amplitudes and transverse positions of the ionization signals. The longitudinal coordinate is obtained from the arrival time of the signals. The TPC, therefore, provides a full three-dimensional measurement of the trajectory and energy deposition pattern (dE/dx) of a charged particle.
Most TPCs built to date, including the Gotthard TPC, have used charge avalanche amplification schemes (typically, MWPCs) for their readout planes \[^{[188]}\]. However, from the discussion in §4.2.4 it appears that signal amplification by electroluminescence is the appropriate option to realize optimal energy resolution in gaseous xenon. The traditional implementation of an EL TPC includes an additional electric field region after the drift volume with intensity between the excitation and ionization thresholds of the gas. The electroluminescence generated in that gap by the ionization electrons can then be detected by a plane of photosensors located nearby. The transverse coordinates of the signals can be reconstructed from the detected light pattern, as in Anger cameras \[^{[191]}\]. Such a detector concept, called sometimes scintillating drift chamber (SDC) \[^{[168]}\], has become popular in recent years due to its successful use in direct searches for dark matter \[^{[189]}[190]\].

The potential of the SDC for optimal energy resolution in xenon gas was demonstrated — although at lower energies than those of interest for $0
\nu
\beta
\beta$-decay searches — by Bolozdynya et al. \[^{[192]}\] using a relatively large system (a drift volume 4 cm long and 30 cm diameter) filled with xenon gas at 9 bar pressure and instrumented with an array of 19 PMTs. They achieved an energy resolution of 2.5% FWHM at 122 keV, which extrapolates to 0.56% at the Q value of $^{136}$Xe assuming an $E^{-1/2}$ dependence. In addition, good enough tracking for our application should be possible with an EL TPC. Convoluted topologies may not be reconstructed well, but the essential track features for background rejection (blob versus no blob, for instance) can be retained using methods similar to those applied in medical imaging. Notice that a commitment to detect EL light with high efficiency also ensures the detection of the primary scintillation light from xenon. This signal can be used as a start-of-event trigger to reconstruct the absolute position of tracks along the longitudinal coordinate. The Gotthard TPC lacked this feature, and thus could not reject background events emanating from the cathode of the TPC, such as the charged progeny of radon.

The SDC, therefore, seems to be the optimal detector concept for a $0
\nu
\beta
\beta$-decay experiment based on gaseous xenon, offering the two features that we sought: high-resolution calorimetry and track reconstruction for background rejection. Nevertheless, the instrumental requirements of these two functions are somewhat conflicting. The energy measurement demands an efficient and precise light collection system with single-photon sensitivity. Photomultiplier tubes seem best suited for the task: they have low noise and high gain, and can cover large areas at a reasonable cost. Unfortunately, most PMTs contain
excessive levels of radioactive contaminants to be used in low-background detectors, and even the most radiopure models can easily become the leading background source of the experiment. Therefore, our design should contain the minimum number of PMTs necessary to obtain a robust measurement of the optical signals. This enters in conflict with the requirements for tracking: recording the light pattern that defines a track demands a dense array of small photosensors — silicon photomultipliers (SiPMs), for instance — placed close to the EL gap. This dilemma can be overcome if the two measurements, calorimetry and tracking, occur through separated, specialized readouts. Figure 4.10 shows two possible implementations of such a detector concept. In the symmetric scheme (the top panel in the figure), a central cathode divides the active volume in two halves. Both ends of the chamber are equipped with an EL amplification region and a tracking plane consisting of small, low-gain photosensors. The sides of the chamber are instrumented with photomultipliers for a precise energy measurement. The asymmetric design (bottom panel in Fig. 4.10) has a single drift volume. Electroluminescence generated at the anode is recorded in the tracking plane right behind it and in the energy plane behind the transparent cathode, at the other end of the chamber. The internal walls of the active volume can be covered with a reflector to improve the light collection efficiency of the detector. In both cases, the sensors used for calorimetry would detect as well the primary scintillation, required in order to establish the start-of-event time.
Figure 4.10. The SOFT (Separated, Optimized Functions TPC) detector concept, with two possible configurations: symmetric (top panel) and asymmetric (bottom). The detection process is the same in both cases: the interaction of ionizing radiation (represented by the grey trace) is immediately followed by the emission of scintillation light (S1). Its detection in the calorimetry photosensors constitutes the start-of-event. The ionization electrons left behind by the ionizing particle drift under the influence of an external electric field towards the TPC anode, entering another region of the detector with an electric field of higher strength. There, electroluminescence light (S2) of intensity proportional to the number of ionization electrons is emitted isotropically. Therefore, both scintillation and ionization produce an optical signal. The S2 light generated at the anode is recorded in the photosensor plane right behind it and used for tracking; it is also detected by the calorimetry photosensors and used for a precise energy measurement.
5.1 Introduction

The NEXT experiment will search for the neutrinoless double beta decay of $^{136}$Xe using a pressurized xenon gas time projection chamber that implements the (asymmetric) SOFT detector concept described in the previous chapter. The detector will be installed at the Laboratorio Subterráneo de Canfranc, located in an old railway tunnel under the Spanish Pyrenees [193]. The experiment is a collaboration involving about 60 physicists and engineers from 13 different research institutes and universities in Spain, Portugal, USA, Colombia and Russia.

The first phase of the project (2009–2014) has been largely devoted to R&D with two medium-size prototypes, NEXT-DEMO and NEXT-DBDM, built with the double aim of demonstrating the detector concept and gaining technical expertise to facilitate the design, construction and operation of a larger system. Both prototypes and their main results are described in §5.2.

The goal of the second and current phase of the project is the construction, commissioning and operation of the NEXT-100 detector, which is expected to start taking data in 2017. The design of NEXT-100, described in some detail in §5.3, follows closely that of the DEMO and DBDM prototypes, incorporating most of the technical solutions found during the R&D phase. The detector is a cylindrical TPC with an active volume of approximately 1.15 m$^3$. Two arrays of photosensors are located at opposite ends of the chamber: the energy plane, composed of 60 low-background photomultiplier tubes sealed into individual pressure-resistant copper enclosures, is positioned behind the cathode, whereas the tracking plane, consisting of 1-mm$^2$ silicon photomultipliers (SiPMs) spaced at a pitch of 1 cm, is located close behind the EL amplification region. The active volume is shielded from external radiation by at least 12 cm of copper in all directions. This inner shield, the electric-field cage and both
sensor planes are housed in a stainless steel pressure vessel designed to withstand up to 20 bar. The vessel sits on top of an anti-seismic pedestal and inside a lead castle shield. This design will be validated with a 1:2 scale prototype called NEW that the Collaboration will operate at LSC during 2015–2016.

Xenon-136 is the cheapest $\beta\beta$ isotope \[^{[133]}\], and hence the most obvious candidate for a future tonne-scale $0\nu\beta\beta$-decay experiment\[^{[1]}\]. In this respect, NEXT-100, in addition to contributing to the current exploration of the degenerate region of neutrino masses, will also serve as a springboard for a possible tonne-scale xenon gas detector.

### 5.2 Results from the R&D phase

#### 5.2.1 NEXT-DEMO

The NEXT-DEMO detector was designed as a proof-of-concept prototype and testbed of the NEXT technology. It is an asymmetric EL TPC with dimensions approximately 4.5 times smaller than those of NEXT-100. It is instrumented with an energy plane composed of 19 1-inch PMTs and a tracking plane with 256 1-mm$^2$ SiPMs. The TPC and photosensor arrays are contained within a stainless steel vessel 60 cm long and 30 cm diameter that can withstand pressure up to 15 bar. The detector is neither radiopure nor shielded against natural radioactivity. It is installed in a clean-room at the Instituto de Física Corpuscular (IFIC), in Valencia, Spain. In this section, we will briefly describe NEXT-DEMO and its main results. Further details can be found in several dedicated publications \[^{[155][194][197]}\].

A cutaway drawing of NEXT-DEMO is shown in Figure 5.1. The two electric-field regions of the detector, the 300 mm long drift volume and the 5 mm long EL gap, are delimited by three metallic rings (cathode, gate and anode), 229 mm in diameter, with stainless steel wire grids stretched over them. The cathode grid is made of parallel wires spaced 0.5 cm, whereas gate and anode are made with wire mesh of, respectively, 76% and 88% open area. The anode ring is set at ground potential, while the gate is at a negative voltage such that an electric field above the excitation threshold of xenon (0.86 kV cm$^{-1}$ bar$^{-1}$) is created in the EL gap. A moderate, uniform electric field (0.5 kV cm$^{-1}$, typically) is established in the drift region by supplying a large negative voltage to the cathode, then grading it using a series of metallic rings (also 229 mm

\[^{1}\text{In fact, there is already about a tonne of enriched xenon in the world, considering the amounts owned by the EXO, KamLAND-Zen and NEXT collaborations.}\]
5.2. RESULTS FROM THE R&D PHASE

Figure 5.1. Cross-section drawing of the NEXT-DEMO prototype with major parts labelled.

diameter) that enclose the volume and are connected electrically via 0.5-GΩ resistors. The high voltage is supplied to cathode and gate through custom-made feed-throughs tested to high vacuum and 100 kV without leaks or sparking. Inside the electric-field rings, six reflecting panels made of polytetrafluoroethylene (PTFE) are mounted forming a tube of hexagonal cross section with an apothem length of 8 cm. The panels were vacuum-evaporated with tetraphenyl butadiene (TPB), which shifts the VUV light emitted by xenon to blue, so as to improve the light collection efficiency of the detector.

The tracking plane of NEXT-DEMO sits 5–10 mm away from the anode. It consists of 256 Hamamatsu S10362-11-050P SiPMs (1 mm² active area) regularly positioned at a pitch of 1 cm and distributed between four boards, each with 64 sensors arranged as a 8 × 8 matrix. Since the SiPMs are not directly sensitive at the emission wavelength of xenon, the boards were coated to a TPB thickness of 0.1 mg cm⁻² prior to their introduction in the detector. An automated active voltage control developed by the Collaboration stabilizes the gain of the SiPMs — which is very sensitive to temperature variations — at an approximate value of 7.5 × 10⁵.

The energy plane is equipped with 19 Hamamatsu R7378A photomultiplier tubes. These are 1-inch, pressure-resistant (up to 20 bar) PMTs with acceptable quantum efficiency (~15%) in the VUV region and higher efficiency at the TPB emission wavelengths (~25%). The resulting photocathode coverage
of the energy plane is 39%. The PMTs are inserted into a PTFE holder following a hexagonal pattern. A stainless steel wire grid is screwed on top of the holder and set to +500 V. This protects the PMTs from the high-voltage set in the cathode, and ensures that the electric field is below the excitation threshold in the 10 cm of buffer between the energy plane and the cathode. The gain of the PMTs is adjusted to about $5 \times 10^6$ to place the mean amplitude of a single photoelectron pulse well above the electronic noise.

The pressure vessel of NEXT-DEMO is a stainless steel (grade 304L) cylindrical shell 3 mm thick, 30 cm diameter and 60 cm long, welded to Conflat flanges on both ends. The two end-caps are stainless steel plates 3 cm thick, with standard Conflat knife-edge flanges. Flat copper gaskets are used for sealing. The vessel was certified to 15 bar operational pressure. On top of the vessel and along the vertical plane there are three Conflat ports used for high-voltage feeding and connection to a mass spectrometer. On the opposite side, at the bottom, another port connects the pressure vessel to the vacuum pumping system. A guillotine valve closes this connection when the vessel is under pressure. The end-caps include the connections to the gas system, and the power and signal feed-throughs for the photosensors. The side of the vessel includes additional ports used as radioactive-source windows. Six bars manufactured from PEEK (a low outgassing plastic) hold the electric-field cage and the energy plane together. The whole structure is attached to one of the vessel end-caps, with the tracking plane attached to the other.

Natural xenon circulates in a closed loop through the vessel and a system of purifying filters that remove the electronegative impurities from the gas. The recirculation loop is powered by an oil-less, single-diaphragm compressor that recirculates the full volume of NEXT-DEMO in about 5 minutes. The standard procedure during normal operation of the detector starts with the evacuation of the vessel to vacuum levels around $10^{-5}$ mbar, followed by an argon purge. A second vacuum step exhausts the argon from the system. The detector is then filled with xenon gas to pressures up to 15 bar. The xenon can be cryogenically reclaimed to a stainless-steel bottle connected to the gas system.

Signal processing, digitisation and read-out are performed using an electronics chain based on the Scalable Readout System developed by the RD51 Collaboration [199]. After amplification and high-frequency filtering, all PMT signals are sampled at 40 MHz using 12-bit digitizer cards, whereas SiPM signals are integrated in intervals of 1 µs. The digitizers are read out by FPGA-based DAQ modules that buffer, format and send event fragments to DAQ
5.2. RESULTS FROM THE R&D PHASE

PCs where the data are stored.

The response of NEXT-DEMO was studied using a 1-μCi $^{22}$Na calibration source placed at a lateral port. Sodium-22 is a long-lived $\beta^+$ radioactive isotope, and the annihilation of the positron emitted in its decay — which rarely leaves the source — results in two back-to-back 511-keV gammas. The coincident detection of the forward gamma in the TPC and of the backward gamma in an external NaI scintillator coupled to a photomultiplier was used to trigger the detector readout. This arrangement optimized the acquisition of useful calibration data.

In the data analysis, the signals from the 19 PMTs are summed after removing their pedestals, and peaks corresponding to PMT pulses are located and classified as a candidate S1 or S2 pulse based on their integral, width and location (see Figure 5.2). S1 pulses are typically shorter than 3 μs (this width mostly due to the shaping of the front-end electronics) and have an average charge of at least 0.5 pe/PMT, whereas S2 peaks are longer than 3 μs and their integrated charge is at least 10 pe/PMT. Once a single S1 pulse and one or more S2 pulses are identified, the drift time of the ionization produced in the event can be determined as the difference in time between the leading edges of the pulses. While the $z$ position of the events in the chamber is unambiguously defined by the drift time and the drift velocity of the electrons in the gas, a number of possible methods can be used to reconstruct the $xy$ position of a charge deposit. For the data presented here, the barycentre of the tracking-plane’s signals was employed.

The raw-energy spectrum (total charge of each S2 peak detected in the selected sample of events) recorded in DEMO is shown in Figure 5.3. Even though the main features of the spectrum (the K-shell xenon X-ray peak at low energies and the photoelectric peak at high energy) are clearly discernible, these data are affected by several spatially dependent factors that contribute to degrade the intrinsic energy resolution of gaseous xenon. Among these are: tracks leaving the active volume; charge losses along the drift due to electron attachment; charge losses at the edge of the fiducial region due to inhomogeneities in the electric field; light absorption in the meshes or other detector surfaces with limited reflectivity; and imperfect rotational symmetry of the light tube.

In order to discard events that are not fully contained in the active volume

---

Notice that more than one S2 peak are possible in a single event if the ionization track had multiple components that arrived at the amplification region at distinct times.
Figure 5.2. Top: Typical waveform of a 511-keV gamma-ray event in the NEXT-DEMO detector obtained by summing the signals of the energy plane (and dividing by the number of PMTs, 19). Centre and bottom: Close-ups of the S1 and S2 pulses in the waveform.
5.2. RESULTS FROM THE R&D PHASE

Figure 5.3. Raw-energy spectrum recorded in NEXT-DEMO using a $^{22}$Na gamma source, with its main features identified.

Figure 5.4. Energy dependence with drift time due to electron attachment in the gas.

and those whose transverse position is expected to be reconstructed with greater error, a fiducial volume is defined selecting only those event with reconstructed position $|x|, |y| < 60$ mm. Data-MC comparison shows that within this region the transverse position is reconstructed accurately. Additionally, to avoid events very close to anode and cathode, which could lose charge, an additional requirement that events be reconstructed with $20 < z < 290$ mm is implemented.

The energy dependence on $z$ due to charge attachment in the gas (§ 4.2.2)
Figure 5.5. NEXT-DEMO $xy$-response map built using the point-like energy depositions of the $K_\alpha$ X-rays emitted by xenon. The corrections are normalised to the value of an arbitrary bin. The differences in $xy$ response, of up to a factor of $\sim 20\%$, are attributed to inhomogeneities in the wavelength shifter coating and in the reflectivity.

is expected to be uncorrelated with other factors, and thus can be treated independently. The effect is well described by an exponential decay. In the data runs presented here, the decay constant extracted (see Figure 5.4) is $(-3.5 \pm 0.4) \times 10^{-5} \, \mu s^{-1}$, equivalent to a mean electron lifetime of about 28 ms, far larger than the maximum drift time of the TPC.

The K-shell X-rays emitted by xenon are a useful tool for the calibration of the $xy$ response of the detector due to their point-like nature (electrons produced at these energies travel less than 1 mm) and abundance in the data. In the range of sensitivity of NEXT-DEMO, the most important lines are the $K_\alpha$ and $K_\beta$ emissions with 29.7 keV and 33.8 keV respectively. The response of the detector to the $K_\alpha$ X-rays was studied sorting them in $xy$ bins of $1 \, \text{cm}^2$ after pre-selection in the raw-energy spectrum. The signal distribution for each PMT in each bin was fitted to a Gaussian, and its mean and variance
5.2. RESULTS FROM THE R&D PHASE

Figure 5.6. Energy spectrum after all spatial corrections. The energy resolution for the $K_{\alpha}$ peak at 30 keV is $(5.691 \pm 0.003)\%$ FWHM, and $(1.62 \pm 0.01)\%$ FWHM for the $^{22}$Na photopeak at 511 keV.

extracted. The following quantity was then calculated for each $xy$ bin:

$$Q_{xy} = \sum_{i} \frac{\bar{q}_{xy,i}}{\sigma_{xy,i}^2},$$  \hspace{1cm} (5.1)

where $\bar{q}_{xy,i}$ is the mean charge recorded by PMT $i$ and $\sigma_{xy,i}^2$ is its variance. The resulting map of the quantity $Q_{xy}$ was then normalized to the value of an arbitrary bin, and interpolated using an algorithm based on the Delaunay triangulation in order to obtain a finer binning. The result is shown in Figure 5.5. The differences in response, of up to a factor of $\sim 20\%$, are attributed to inhomogeneities in the wavelength shifter coating and in the reflectivity.

Using this map, the following estimator of the energy of a track was defined:

$$E = \frac{E_0}{f(x, y)} \sum_{i} \frac{q_i}{\sigma_i^2(x, y)},$$  \hspace{1cm} (5.2)

where $q_i$ is the charge recorded by PMT $i$, $\sigma_i$ is the variance of the PMT response, $f(x, y)$ is the geometrical correction factor for the reconstructed $xy$ position, and the quantity $E_0$ is an overall conversion factor from photoelectrons to energy.
Applying all the corrections described above, the resultant spectrum is that shown in Figure 5.6, with energy resolution for the Kα peak of (5.691 ± 0.003)% FWHM. The 22Na photopeak has a resolution of (1.62±0.01)% FWHM. Extrapolating these two values to the Q value of 136Xe assuming an $E^{-1/2}$ dependence results in a predicted energy resolution of 0.63% FWHM and 0.74% FWHM, respectively. The discrepancy can be attributed to the gradual breakdown of the point-like model as extended tracks become more probable. Future work will involve the generalisation of these corrections to higher energies by slicing long tracks into shorter segments. This, however, will require a more sophisticated xy reconstruction than the barycentre.

Using NEXT-DEMO data, a first approximation to the event topology reconstruction was also performed by subdividing the registered signals in 4-µs time slices (z dimension) and reconstructing a single xy point per slice using the barycentre. The width of the slices was chosen so that the charge collected in a such a window would be sufficient to achieve a reliable xy reconstruction. An energy measurement was then associated to each slice so that the event $dE/dz$ could be studied. The energy and position information were used to calculate a cubic spline between the individual points in order to obtain a finer description of the path. An example of this procedure is shown in Figure 5.7.

### 5.2.2 NEXT-DBDM

The main objective of the NEXT-DBDM prototype, built and operated at the Lawrence Berkeley National Laboratory (USA), was the demonstration of near-intrinsic energy resolution in high-pressure xenon gas. A thorough description of the system and the first set of results can be found in a dedicated paper [200]. We offer here a brief summary.

A cutaway drawing of the DBDM prototype is shown in Figure 5.8. The TPC consists of a hexagonal field cage separated into a drift region of length 8 cm and an amplification region of length 5 mm by grids of wire mesh (88% transparent) stretched tightly across stainless-steel frames. The active region is enclosed by teflon panels bare on the side facing the active volume and with copper strips parallel to the mesh planes on the other side. Adjacent strips are connected through 100-MΩ resistors to grade the potential and produce a uniform electric field. The panels are supported by thin plastic frames. An

---

3 *Double Beta and Dark Matter.*
Figure 5.7. Example of $^{137}$Cs track reconstruction (top panel: $yz$ projection; bottom panel: 3D). The signal of the tracking plane is split into slices of 4 mm width in the longitudinal coordinate. One $xy$ point is calculated for each slice using the barycentre of the SiPM signals. The energy measured in the energy plane (after corrections) is then associated to the point. A cubic spline is used to interconnect the different points.
array of 19 1-inch diameter Hamamatsu 7378A [198] photomultiplier tubes arranged in a hexagonal pattern is located at the end opposite the amplification region (13.5 cm away) and observes the active region.

A stainless steel cylindrical pressure vessel (20 cm diameter, 33.5 cm length) houses the TPC and the PMT array. One end of the vessel is closed in an ellipsoidal shape and the other sealed via a Conflat flange to a stainless steel lid to which the internal components are attached. The lid is connected by a long tube to a stainless steel octagon with 8 Conflat ports, several of which are occupied by multi-pin feedthroughs through which PMT high voltages are input to the interior of the detector and through which the PMT signals are output. An opening of diameter 1.7 cm extends through the center of the octagon and down the tube to a 2 mm calibration-source window to the interior of the pressure vessel. High voltages for the wire meshes are passed into the pressure vessel through the lid via commercial feedthroughs rated to 20 kV at 17 bar and connected via teflon-coated wire to the mesh frames. The pressure vessel is connected to a gas system allowing for full system pump-down to pressures of the order of $5 \times 10^{-5}$ Torr, reclamation and reintroduction of the xenon gas in the system, and constant recirculation of the xenon gas during operation through a heated zirconium-based getter to remove electronegative impurities (O$_2$, H$_2$O, N$_2$, et cetera).

The energy resolution achievable with the TPC was studied using gamma
rays of 662 keV from a $^{137}$Cs radioactive source. A central fiducial cut was made according to the average position of the event determined using the distribution of EL light registered on the PMT plane. In addition, the events were corrected for electron attachment by multiplying the collected signal by a longitudinal-dependent exponential factor corresponding to an electron lifetime of $\tau = 8.3$ ms. Energy resolutions of 1% FWHM for 662-keV gamma rays 5% FWHM for 30-keV fluorescence xenon X-rays were obtained at 10 and 15 atm. These results extrapolate to a 0.5% FWHM resolution at 2459 keV, the energy of the hypothetical neutrinoless double beta decay peak of $^{136}$Xe.

Figure 5.9 summarizes our measurements and understanding of the reached resolution. The lower diagonal line represents the Poisson statistical limit in resolution ($2.35 \text{ FWHM}/\sqrt{N}$) for the measurement of a certain number of detected photons, while the upper diagonal line includes, additionally, the degradation due to PMT response (mostly from afterpulsing). The circle data points show the resolutions obtained for dedicated LED runs with varying light intensities per LED pulse. The points follow the expected resolution over the two-decades range studied. The two horizontal lines represent the nominal intrinsic resolution of xenon gas, Eq. (4.9), for 30 and 662 keV, respectively. Finally, the two curved lines are the expected overall energy resolutions with contributions from the intrinsic limit and the measurement of photons. Our 662-keV data (squares) and xenon X-ray data (triangles) taken with various EL gains follow the expected functional form of the resolution, but are 20–30% larger, possibly because of $xy$-response non-uniformity within the fiducial volume.

### 5.3 The NEXT-100 detector

The NEXT experiment was formally proposed to the Laboratorio Subterráneo de Canfranc (LSC) in 2009 in a Letter of Intent [201] that outlined the physics case and basic design concepts of an EL TPC for $0\nu\beta\beta$-decay searches with a source mass of the order of 100 kg of $^{136}$Xe. The detector design was narrowed down in a Conceptual Design Report [202], published in 2011, and fixed a year later in a Technical Design Report (TDR) [203]. The rationale behind most of the choices made during the design process was that of minimizing the R&D required to produce a competitive detector. Consequently, the Collaboration favoured the configuration that had been tested with the DEMO and DBDM prototypes.
Figure 5.9. Energy resolution (FWHM) in the NEXT-DBDM prototype. Data points show the measured resolutions as a function of the number of photons detected for 662-keV gamma rays from $^{137}$Cs (squares), 30-keV xenon X-rays (triangles) and LED light pulses (circles). The expected resolution — including the intrinsic Fano factor, the statistical fluctuations in the number of detected photons and the PMT charge resolution — is shown for X-rays (dash-dot-dot line) and for 662-keV gammas (dash-dot-dot-dot line). The measured resolutions are 20–30% larger than expected, possibly due to non-uniformities in the $xy$ response of the detector.

Figure 5.10 shows a longitudinal section of NEXT-100, with its main components labelled. These subsystems and others not shown in the figure (gas system, electronics and shielding) are described in the following.

5.3.1 Electric-field cage

The main body of the field cage is an open-ended high-density polyethylene (HDPE) cylinder of 148 cm length, 107.5 cm inner diameter and 2.5 cm wall thickness that provides structural stiffness and electric insulation. A series of copper rings for electric field shaping are fixed to the inner surface of the cylinder. The rings are covered by PTFE tiles coated with TPB to shift the xenon VUV light to blue and improve the light collection efficiency. One of the ends of the HDPE cylinder is closed by a fused-silica window 1 cm thick. The inner surface of the window functions as the TPC anode thanks to a transparent, conductive, wavelength-shifting coating of indium tin oxide.
5.3. THE NEXT-100 DETECTOR

Figure 5.10. Cross-section view of the NEXT-100 detector inside its lead castle shield. A stainless-steel pressure vessel (A) houses the electric-field cage (D) and the two sensor planes (energy plane, B, and tracking plane, E) located at opposite ends of the chamber. The active volume is shielded from external radiation by at least 12 cm of copper in all directions.

(ITO) and TPB. The two other electrodes of the TPC, EL gate and cathode, are positioned 0.5 cm and 106.5 cm away from the anode, respectively. They are built with highly-transparent stainless steel wire mesh stretched over circular frames. The electrodes will be set at voltages such that a moderate electric field of 0.3–0.5 kV cm\(^{-1}\) is established in the drift region between cathode and gate, and another field of higher intensity, 2–3 kV cm\(^{-1}\) bar\(^{-1}\), is created in the EL gap, between gate and anode, for the amplification of the ionization signal. The high voltage is supplied to the electrodes via custom-made feed-throughs similar to those built for NEXT-DEMO.

5.3.2 Energy plane

The energy plane of NEXT-100 will be composed of 60 Hamamatsu R1410-10 photomultiplier tubes located behind the cathode of the TPC and covering
approximately 30% of its area. This coverage is a compromise between the need to collect as much light as possible for a robust measurement of the energy and $t_0$, and the need to minimize the number of sensors to reduce cost, complexity and radioactivity. Simulations show that this coverage will allow optimal detection of events with energies well below 30 keV in the full chamber range.

The R11410-10 is a 3-in PMT specially developed for low-background operation [204]. It is equipped with a synthetic silica window and a photocathode made of low temperature bialkali with quantum efficiency above 30% for the emission wavelengths of xenon and TPB [204]. Pressure-resistance tests run by the manufacturer showed that the R11410-10 cannot withstand pressures above 6 atmospheres [204]. Therefore, in NEXT-100 they will be sealed into individual pressure-resistant, vacuum-tight copper enclosures coupled to sapphire windows. The window, of 5 mm thickness, is silver brazed to the front-end of the enclosure. The back side is closed with a 12 cm thick copper cap sealed with an O-ring. The PMT is optically coupled to the window using an optical gel with a refractive index intermediate between those of fused silica and sapphire. A spring on the backside pushes the PMT against the enclosure window. The external face of the enclosure windows is coated with TPB. The copper enclosures are all mounted to a common carrier plate that attaches to an internal flange of the pressure vessel head. The enclosures are all connected via vacuum-tight tubing conduits to a central manifold, and maintained at vacuum well below the Paschen minimum, avoiding sparks and glow discharge across PMT pins. The PMT cables route through the conduits and the central manifold to a feedthrough in the pressure vessel nozzle.

5.3.3 Tracking plane

The tracking function in NEXT-100 will be provided by a matrix of silicon photomultipliers (SiPMs) at a pitch of 1 cm located behind the fused-silica window that closes the EL gap. The SiPM chosen is the Hamamatsu model S10362-11-050P. This device has an active area of 1 mm$^2$, 400 sensitive cells of 50 $\mu$m size, and high photon detection efficiency in the blue region (about 50% at 440 nm). SiPMs are very cost-effective and their radioactivity can be very low, given their composition (mostly silicon) and very small mass. The SiPMs will be mounted on 8 × 8 flexible circuit boards made of Kapton and copper. The boards have long tails that carry the signals through zigzagging slits — so as to avoid a straight path for external gammas — made in the copper plates
that shield the active volume (see Figure 5.11). The tails are connected to flat shielded cables that extract the signals from the vessel via large custom-made feed-throughs. Like in NEXT-DEMO, all the front-end electronics will be outside the chamber. In total, the NEXT-100 tracking plane will be composed of 7168 SiPMs distributed between 112 boards.

5.3.4 Pressure vessel and inner copper shield

The pressure vessel of NEXT-100, shown in Figure 5.12, consists of a cylindrical central section of 160 cm length, 136 cm inner diameter and 1 cm wall thickness, and two identical torispherical heads of 35 cm height, 136 cm inner diameter and 1 cm wall thickness. It has been fabricated with stainless steel Type 316Ti due to its low levels of natural radioactive contaminants. Designed almost entirely by the Collaboration following the ASME Pressure Vessel Code, the vessel has been built by a specialized company based in Madrid.

The $\gamma$ radioactivity from the vessel will be attenuated by three orders of magnitude by an inner copper shield of 12 cm thickness.

5.3.5 Electronics

The NEXT-100 data acquisition system (DAQ) follows a modular architecture known as the Scalable Readout System (SRS), developed by CERN and the
Figure 5.12. The NEXT-100 pressure vessel during the final stages of fabrication.
NEXT Collaboration under the auspices of the RD51 Collaboration [199]. At the top of the hierarchy, a PC farm running the DAQ software, DATE, receives event data from the DAQ modules via Gigabit Ethernet links. The DATE PCs (Local Data Concentrators, LDCs) assemble incoming fragments into sub-events, which are sent to one or more additional PCs (Global Data Concentrators, GDC). The GDCs build complete events and store them to disk for offline analysis. The DAQ modules used are Front-End Concentrator (FEC) cards, which serve as the generic interface between the DAQ system and application-specific front-end modules. The FEC module can interface different kinds of front-end electronics by using the appropriate plug-in card. Three different FEC plug-in cards are used in NEXT-100 (energy plane readout digitization, trigger generation, tracking plane readout digitization).

5.3.6 Gas system

The gas system of NEXT-100 must be capable of carrying out its functions (evacuation of the detector to vacuum levels around $10^{-5}$ Torr, pressurization and depressurization with xenon and argon, and re-circulation of the gas through purification filters to remove electronegative impurities) with negligible losses of the very expensive enriched xenon and without damage to the detector.

The most vulnerable component of the gas system is the re-circulation compressor, which must have sufficient redundancy to minimize the probability of failure and leakage. The compressor chosen by the Collaboration is made with metal-to-metal seals on all the wetted surfaces. The gas is moved through the system by a triple stainless steel diaphragm. Between each of the diaphragms there is a sniffer port to monitor for gas leakages.

In addition, an automatic recovery system is needed to evacuate the chamber in case of an emergency condition, such as an over-pressure that could potentially cause an explosion or an under-pressure indicating a leak in the system. An expansion tank of 35 m$^3$ will be placed inside the laboratory to quickly reduce the gas pressure in the system. For controlled evacuation, the gas will be cryo-pumped into an external vessel cooled with liquid nitrogen.

5.3.7 External shield and supporting structures

NEXT-100 will be shielded from the external flux of high-energy gamma rays by a lead castle of 20 cm wall thickness. It is made of staggered lead bricks held by a stainless steel frame. The lead castle is made of two halves mounted
on a system of wheels that move on rails with the help of an electric engine. The movable castle has an open and a closed position. The former is used for the installation and service of the pressure vessel; the latter position is used in normal operation. A lock system fixes the castle to the floor in any of the two configurations to avoid accidental displacements.

Due to the mild seismic activity of the part of the Pyrenees where the LSC is located, a comprehensive seismic study has been conducted as part of the project risk analysis. As a result, an anti-seismic structure that will hold both pressure vessel and shielding has been designed. This structure will be anchored directly to the ground and independent of the working platform to allow seismic displacements in the event of an earthquake.

An elevated working platform has been built prior to the installation of NEXT-100. It is designed to withstand a uniform load of 1500 kg/m² and a concentrated load of 200 kg/m². It is anchored to the hall ground and walls. The platform floor tiles are made of galvanized steel and have standard dimension to minimize cost.
5.3.8 NEW

The NEXT-White\(^4\) detector, currently under fabrication, is a 1:2 scaled-down version of NEXT-100, implying about 10 kg of xenon mass at 15 bar and readout planes of 20% the size of the final ones. The design of NEW follows closely that of NEXT-100, implementing the same technical solutions and using the same materials and photosensors. In this way, NEW serves as a demonstrator of the NEXT technology, and its operation will provide valuable information to prevent possible future problems with NEXT-100. In addition to these technological goals, NEW has several physics objectives:

- Refinement of detector calibration techniques for NEXT-100.

- The DEMO and DBDM prototypes are too small to contain tracks of the energies of interest for 0νββ searches in \(^{136}\)Xe. Therefore, NEW data will be extremely useful for the optimisation of our reconstruction and pattern-recognition algorithms at those energies. In particular, NEW will allow the accurate characterization of the 2-electron tracks.

- The background rate predicted for the NEXT-100 (§7.4) is estimated using detector simulations and the data from material screening. NEW deploys the same materials chosen for NEXT-100, thus allowing a stringent test of the background model. Understanding the effects of radon is of uppermost importance.

- Depending on the actual background level and the length of the data-taking period with enriched xenon, it may be possible to measure the half-life of the two-neutrino decay mode of \(^{136}\)Xe.

NEW, shown in Figure 5.14, is a xenon gas time projection chamber with separate detection planes for calorimetry and tracking. The active volume of the TPC is a cylinder 55 cm long and 48 cm in diameter. Two metallic wire meshes mounted in circular frames and placed as basis of the cylinder act as the TPC electrodes. An electric drift field of about 500 V cm\(^{-1}\) is created between them by supplying a large negative voltage to the cathode, then degrading it using a series of metallic rings spaced 12 mm and connected via resistors. These rings are mounted on the inner surface of a polyethylene tube that provides electric isolation to the system. A second electric field of higher intensity is

\(^4\)Named after our late collaborator Prof. James T. White.
established in a 5-mm region between the anode and a quartz plate (coated with ITO) set to ground potential. This is the so-called electroluminescence gap where the ionization signal is amplified.

The energy plane, placed behind the cathode, is equipped with 12 Hamamatsu R11410-10 photomultiplier tubes. The resulting photocathode coverage of the energy plane is about 35%. Since the PMTs cannot withstand high pressure, the volume containing them is separated from the rest of the detector and maintained at a light vacuum. The PMTs are coupled to the field cage via sapphire windows coated with TPB (to shift the VUV light emitted by xenon to blue). A 7 cm buffer region between the PMTs and the cathode protects these from the high voltage and ensures that the electric field in that region is below the EL threshold. The tracking plane consists of 1408 P SiPMs distributed between 22 boards, each with 8×8 sensors spaced 1 cm. The plane is positioned behind the 5-mm thick quartz plate that closes the EL gap. The quartz plate is also coated with TPB.

Both sensor planes and the TPC itself are contained within a cylindrical stainless-steel (316Ti alloy) designed to withstand more than 20 bars. The sensor planes are mounted in 12-cm thick copper plates, whereas the TPC is
surrounded by copper bars 6 cm thick that attenuate the external radiation entering the detector. The vessel is mounted on a seismic pedestal and surrounded by a lead castle (20 cm thick) that shields the detector against the high-energy gamma flux from the rocks of the laboratory cavity.
NEXUS, the end-to-end
detector simulation of NEXT

6.1 Introduction

Detector simulation is an essential tool for the NEXT experiment. It has guided the detector design process providing estimations of the performance of proposed configurations as well as the acceptable levels of radioactive impurities in all detector materials and components. Besides, the development of event reconstruction and selection algorithms — required to complete our understanding of the experiment’s performance and sensitivity — relies on the use of realistic simulation datasets. Moreover, since the data produced by this type of detectors are often complex, simulations have been helpful for the extraction of physics measurements during the data analysis.

In this chapter we describe NEXUS\textsuperscript{1}, the detector simulation program of the NEXT experiment, developed by the Collaboration. It is based on the Geant4 toolkit \cite{Agostinelli2003}, arguably the standard simulation code in particle physics, and is integrated into the general software framework of the NEXT experiment. NEXUS covers the entire simulation process, from the generation and transport of particles interacting in the detector to the production of the associated primary signals (ionization electrons and scintillation photons) and their collection, amplification and conversion into electric signals in the photosensors. The output of the full simulation chain is written in a format identical to the output of the NEXT data acquisition system (DAQ). In this way, both the simulated and real data can be run through the same reconstruction algorithms.

The chapter is organized as follows. Section \ref{sec:nexus-structure} describes the basic structure of NEXUS and the simulation process. Sections \ref{sec:nexus-geometry} and \ref{sec:nexus-primary} describe, respectively, the geometries and primary generators implemented in the simulation. The physics models used in NEXUS, including those written by us

\footnote{An acronym for NEXr Utility for Simulation.}
for the simulation of the ionization charge drift and EL amplification, are described in §6.5. Section 6.6 briefly describes the digitization of the simulated signals. Finally, in §6.7 we showcase several data-simulation comparison results.

6.2 Overview of Geant4 and NEXUS

Geant4 is a software toolkit for the simulation of the passage of particles through matter using Monte Carlo methods. It is written in C++ and designed with an emphasis on modularity and expandability. All aspects of the simulation process are considered in the toolkit: description and navigation of detector geometries, models for the interaction of particles with matter, particle tracking, detector response, visualization, user interface, etcetera.

Geant4 divides a detector simulation into several basic components that the user must define; namely:

- **Detector geometry**: the physical layout of the simulated system, including a description (microscopic and macroscopic) of the materials used in its construction.

- **Physics list**: a collection of physics processes and their associated particles to be considered in the simulation.

- **Event generation**: the initial conditions of each event to be simulated; that is, number and type of particles in the event, their position within the detector geometry and their initial three-momentum.

- **User actions**: operations that retrieve data of interest to the user about the trajectory and interactions of particles as they propagate through the detector. This information is available at several processing stages during the simulation.

NEXUS is a Geant-based computer program that comprises a collection of such components. The user selects at runtime a component of each of the above categories via a macro file. These components may instantiate others, and have their own macro commands that let the user further refine the simulation configuration. The output information from the simulation is saved in the form of high-level data objects — such as particles, tracks or detector hits — using ROOT [206]. Such a design, illustrated in Figure 6.1, allows the simulation
6.2. OVERVIEW OF GEANT4 AND NEXUS

![Diagram](image)

**Figure 6.1.** Basic structure of the NEXUS detector simulation program. The user selects via a macro file the components — geometry, physics list, primary generator and user actions — that wants to load at runtime. These components may have their own macro commands that let the user further refine the configuration of the simulation. The output information is saved to disk in the form of high-level data objects such as particles, tracks or detector hits.

... of many different systems to be performed with the same program, using exactly the same physics models, geometries and services (job configuration, persistency, et cetera), reducing as a result the coding and debugging efforts, and facilitating the cross-comparison and validation of results.

### 6.2.1 Event data flow

The basic simulation unit in Geant4 is the *event*, consisting of a set of *primary particles* that interact with the detector materials producing an associated *detector response*. The data flow during the simulation of an event in NEXUS is outlined in Figure 6.2. A *generator* produces *primary particles* with a given...
momentum and position within the detector. The events can be filtered at
generation time so that only those with a certain property (e.g. energy above a
certain threshold) are kept. During this step, the run number for the simulated
dataset and event numbers for each event are established. Event numbers are
consecutive in a single run, but events may be omitted because of filtering. Run
numbers for simulated datasets derive from the job options used to generate
the sample and mimic real run numbers used during data taking. A record of
the particles produced by the generator is stored in the simulation output file.

All primary particles are propagated through the detector geometry by
Geant4 down to zero range, and in their interactions with the detector materi-
als further particles may be generated. The trajectories and associated energy
deposits of charged particles in the active volume of the detector are recorded
as track hits and written to the output file. The number of photons converted
in a given photosensor is stored as a zero-suppressed time histogram called
photosensor hit. Its sampling time (bin width) can be configured by the user.
In general, the electroluminescence photons are not tracked individually, but
its detection probability is looked up in a table indexed by its xy production
location in the amplification plane. This table is produced in an independent
Monte-Carlo run in which a large number of photons is generated per point
on a grid dividing the EL plane, and the detection probability for each pho-
tosensor recorded based on the number of photons collected out of those generated. The genealogy of each simulated particle and the history of their interactions are also stored in the output file, excepting the optical photons and ionization electrons. A post-event filtering step decides whether to store an event depending on the results of the simulation. A record of the number of filtered events is stored in the output file together with other basic statistics of the run.

The digitization constructs *digits*, i.e. data structures equivalent to those generated by the NEXT DAQ, using the photosensor hits. Detector noise is added at this stage. The Monte Carlo (MC) *truth* information — that is, primary particles and tracks — is kept in the reconstruction jobs so that it can be used during the analysis of simulated data to quantify the success of the reconstruction algorithms.

### 6.2.2 Fast simulation

The full simulation and reconstruction of events in the energy region of interest for $0\nu\beta\beta$-decay searches is a very CPU-time-consuming task. This makes many physics studies unviable without the use of considerable computing resources, limiting the utility of the simulation program. As a solution, we developed a *fast-simulation* mode in which the Monte Carlo truth information is smeared by the detector spatial and energy resolutions to provide *pseudo-reconstructed* events as similar as possible to those obtained with the full software chain.

In this simplified mode, neither the primary scintillation nor the ionization charge collection and amplification are simulated. In the pseudo-reconstruction, the active volume of the detector is divided into a regular 3D grid of small cuboids known as *voxels*. The recorded energy deposits left by charged particles in the active volume (the MC-truth tracks) are then accumulated in the voxels that contain their positions. The dimensions of the voxels, which are configurable, are typically proportional to the spatial resolution of the detector. The total energy deposited in the active volume is smeared according to a user-defined probability density function. The energy accumulated in each voxel is multiplied by the ratio of the smeared energy to the truth one.
Figure 6.3. The NEXT-DEMO detector as modelled in NEXUS. The rings that form the electric-field cage have been made transparent to allow for a better visualization of the other elements.

6.3 Detector geometries and materials

A detector geometry in Geant4 is made of a number of volumes, each one defined by its shape, dimensions, material and position inside the simulated world. The physics processes simulated by Geant4 in a given volume depend on the material used for its construction. A material is defined by its chemical composition, density, physical state, temperature and pressure. Some specific physics processes may require additional information. NEXUS complements the internal database of materials of Geant4 with several definitions relevant for NEXT. Especially important is the compilation of properties of optical media (refractive index, reflectivity, absorption, etc.), essential for a precise simulation of the signals registered by the photosensors. Likewise, the detection properties of xenon (drift velocity, diffusion, etc.) were extracted from the literature or from the Magboltz program [207] (see Figure 6.4), and tabulated for their use in the simulation.

The geometries included in NEXUS range from the the most basic elements
used in NEXT, such as the silicon photomultipliers, to full detectors like NEXT-DEMO or NEXT-100 (see Figure 6.3). All geometry classes derive from a base class, thus allowing for the reuse of existing geometries in the implementation of others. For example, a photomultiplier tube already coded can be used in the geometry of several detectors. Each geometry includes code for the generation of 3D points within its volumes. These points, which are used by the event generators, may correspond to fixed positions relevant for a simulation (for instance, the position of a calibration source) or be chosen at random inside a certain volume. The geometries may also have configuration parameters that can be set by the user at runtime via a macro file.

Two geometry representations of the NEW and NEXT-100 detectors are available. One includes the external shielding and all detector components that are relevant for background simulations. Since an approximate geometry is used for the shape of some complex components, a check is made to verify that the mass of materials is accurately reproduced. Although these distinctions could lead to minor differences in gamma attenuation between data and simulation, the effect is expected to be small. The second geometry implements only the optically active region of the detector (drift volume, detector planes and field cage). Less computationally demanding than the full model, this is useful for light collection studies. Both produce, nevertheless, exactly the same optical response.

### 6.4 Event generators

NEXUS offers a number of choices for event generation. The single-particle generator can produce particles — selected by their Geant4 name or PDG identifier — with a configurable position and momentum. This generator is especially useful for testing and debugging purposes. A variation of the single-particle generator is employed for the production of radioisotopes in the simulation of calibration sources or the natural radioactivity of detector components. Geant4 uses data retrieved from the ENSDF [208] to simulate the decay of radioactive nuclei by α, β± and γ emission and by electron capture. The thorium and uranium series can be simulated down to the stable isotopes that end the decay chains. If the user is only interested in certain isotopes of the series, the simulation can be interrupted once the decay chain reaches a specified point, saving computation time on simulated events that would ultimately be discarded in the analysis.
The DECAY0 Monte Carlo event generator [209], a separate Fortran program written by V. I. Tretyak and collaborators, is used for the production of $\beta\beta$-decay events. DECAY0 can simulate 14 different $\beta\beta$ decay modes — including $2\nu\beta\beta$ decay and $0\nu\beta\beta$ decay mediated by the exchange of light Majorana neutrinos — for 21 candidate isotopes. The program outputs text files with the initial time, energy and angular distribution of the particles emitted in the decays. NEXUS includes an interface to read those files and pass the information to Geant4.

6.5 Physics

6.5.1 Physics lists

By default, NEXUS loads the low-energy physics list provided by Geant4 that makes use of the Penelope models [210] and the Lawrence Livermore National Laboratory data libraries [211] for an accurate description of the electromagnetic interactions of charged particles and $\gamma$-rays with matter down to 250 eV. These physics models also incorporate the description of the atomic de-excitations (via fluorescence and Auger effect) that sometimes follow an ionization, Compton interaction or photoelectric emission. The validity of the Geant4 physics models has been systematically assessed by the Geant4 Collaboration [212] and by other groups [213] at the few-percent level.

To avoid infrared divergences and for optimization reasons, some of the physics processes require a production threshold below which the energy loss is assumed continuous and no secondary particle is generated. Geant4 defines such a threshold as a range cut-off, which is internally converted to an energy for individual materials. Each application must find the threshold that results in the right compromise between accuracy and computing time. In NEXUS, the range is set to 750 $\mu$m for $\gamma$ particles, electrons and positrons. For gaseous xenon at 15 bar, that range corresponds to an energy threshold of 2 keV for gammas, and 30 keV for electrons and positrons.

In addition to the low-energy electromagnetic physics processes, NEXUS employs the Geant4 optical physics list, which contains models describing the production and propagation of optical photons. Geant4 does not include code for simulating the drift and amplification of ionization charges. Consequently, we have developed our own models, which are described below.
6.5.2 Simulation of the primary signals

NEXUS utilizes the default Geant4 physics model for the simulation of the primary scintillation. This model requires several sources of input data for each scintillating material considered in a simulation: the average number of photons emitted by the scintillator per unit of deposited energy, that is, \(1/W_{sc}\); the time constants of the fast and slow components of the scintillation; the ratio of photons produced with the fast scintillation time constant to those produced with the slow one; and a probability density function describing the scintillation emission spectrum. The scintillation process is simulated at the end of every tracking step in which a charged particle has deposited energy in a scintillating material. The number of photons to be generated is sampled from a Poisson distribution with mean equal to the ratio of the deposited energy to \(W_{sc}\). The photons are generated evenly along the track segment and are emitted uniformly into \(4\pi\) with a random linear polarization. Their energy is sampled from the input spectrum.

As mentioned before, the scintillation properties of gaseous xenon were compiled from the literature and stored in NEXUS in a format adequate for Geant4. The emission spectrum and time constants were taken from Koehler et al. (1974) [214], where measurements for several gas pressures are presented. For the scintillation light yield, we have used our own measurement in NEXT-DEMO: \(W_{sc} = (60 \pm 12)\) eV [197]. No scintillation quenching is considered for heavy particles, in agreement with our results for alpha particles [155].

The simulation of the collection and amplification of the ionization signal starts with the generation of the ionization electrons, which are defined in NEXUS as a particle different from the standard electrons, so that they can have separate physics processes assigned. The generation process is very similar to that described above for the primary scintillation: at the end of every tracking step of a charged particle in the active volume of the detector, we generate a number of ionization electrons sampled from a Gaussian distribution with mean \(E_{dep}/W_i\) and variance \(F_{GXE} \cdot N_i\), where \(E_{dep}\) is the energy deposited in the xenon by the particle during the step, \(F_{GXE}\) is the Fano factor of gaseous xenon, and \(W_i\) is the average energy required to produce an ionization pair. The ionization electrons are given initial positions generated randomly along the step. The maximum step length is limited by default in the active volume to 1 mm in order to closely follow the trajectory of the particles.

Once the ionization electrons have been generated, they are transported one by one towards the TPC anode with the assistance of a drift-field map. For
any given initial position in the active volume, the map returns the corresponding final position according to the drift-field lines and the total drift length. Using tabulated values of the drift velocity in xenon versus the electric field strength (see the left panel of Figure 6.4), the drift length is converted into time units. Electron attachment (§ 4.2.2) is simulated by sampling a random number from an exponential distribution with a decay constant equal to a user-defined electron lifetime — see Eq. (4.7) —, and comparing it to the total drift time of the electron. If the random number is shorter than the drift time, the electron is considered lost. To simulate the effect of diffusion (§ 4.2.2), the final position of the electron is smeared according to a 3D Gaussian with longitudinal and transverse standard deviations given by Eq. (4.6). The diffusion coefficients of xenon gas in terms of the drift field strength (see Fig. 6.4, right panel) are also available in NEXUS in tabulated form.

The drift-field maps have another property attached to them: the electroluminescence yield per unit of drift length, $Y/d$. In the case of xenon, the EL yield is linearly proportional to the reduced electric field, $E/p$, above a threshold of about 0.83 kV cm$^{-1}$ bar$^{-1}$ and up to 6 kV cm$^{-1}$ bar$^{-1}$ [165]:

$$Y/d = 140 \left( E/p - 116 \right) p,$$

(6.1)
where \( d \) is the drift length in cm, \( E \) is the electric field strength in kV/cm, and \( p \) is the gas pressure in bar. If the drift field in a certain detector region is above the emission threshold, the electroluminescence is simulated along the drift path of an ionization electron by generating a number of photons sampled from a Gaussian distribution with mean \( Y \) and variance \( J_{CP} \cdot Y \), where \( J_{CP} \) is the Conde-Policarp factor (see Fig. 4.8). The photons are positioned uniformly along the drift path and are given a random, isotropic direction of emission. We assume as well that the emission spectra of electroluminescence and scintillation are the same.

### 6.5.3 Light transport and collection

Optical photons in Geant4 can undergo four different types of interactions: elastic (Rayleigh) scattering, bulk absorption, wavelength shifting, and boundary interactions (reflection, refraction and absorption). The accuracy in the simulation of all these processes depends largely on the correct characterization of the optical properties (refractive index, absorption length, reflectivity, etc.) of the detector materials, which must be provided to Geant4.

The optical properties of gaseous xenon are reasonably well known. The refractive index of xenon as a function of its density and the photon energy was taken from Baldini et al. (2006) [215] expressed in the form of the Lorentz-Lorenz equation:

\[
\frac{n^2 - 1}{n^2 + 2} = -A(h\omega) \frac{\rho}{M},
\]

where \( n \) is the refractive index, \( \rho/M \) is the molar density of the gas and \( A \) is the so-called *first virial coefficient*, which is a function of the photon energy. From a fit to the available experimental data, the following expression for \( A \) was obtained:

\[
A(E) = \sum_{i=1}^{3} \frac{P_i}{E_i^2 - E_{\text{th}}^2},
\]

where \( E_i \) are the energies of the absorption lines (\( E_1 = 8.4 \) eV, \( E_2 = 8.81 \) eV and \( E_3 = 13.2 \) eV) and \( P_i \) are the parameters resulting from the fit (\( P_1 = 71.23 \) eV² cm³ mol⁻¹, \( P_2 = 77.75 \) eV² cm³ mol⁻¹ and \( P_3 = 1384.89 \) eV² cm³ mol⁻¹). Figure 6.5 shows the refractive index of xenon at 15 bar as a function of the photon energy.

Light absorption by gaseous xenon over the wavelength range of interest for NEXT (160–500 nm) can be assumed negligible for all practical purposes [158].
Figure 6.5. Refractive index of xenon gas at 15 bar in terms of the photon energy, with the value at the energy (7.08 eV) of the peak of the xenon scintillation spectrum indicated.

Bulk absorption in the gas is usually due to the presence of impurities, in particular water and oxygen. However, these can be reduced to insignificant levels (< 1 ppb) circulating continuously the gas through heated getters.

The transparency and refractive index of the glasses commonly used in optical equipment are either made available by the manufacturers or easily found in the literature. A correct simulation of the characteristics of these materials is especially important in the case of the NEXT-100 detector. Before reaching the photocathode of the photomultiplier tubes, the photons emitted from the xenon will have to cross three different volumes: the sapphire enclosures, an optical gel and the fused-silica windows of the PMTs (§5.3.2). Similarly, the SiPMs are separated from the active volume of the detector by a fused-silica plate (§5.3.3). Fresnel reflections at the interfaces between the different media can have a sizeable impact in the light collection efficiency of the sensor planes. The refractive index of a glass is usually provided as a Sellmeier equation:

\[ n^2(\lambda) = 1 + \sum_{i=1}^{3} \frac{B_i \lambda^2}{\lambda^2 - C_i}, \]  

(6.4)
Table 6.1. Sellmeier coefficients for several glasses defined in NEXUS [216].

<table>
<thead>
<tr>
<th>Material</th>
<th>$B_1$</th>
<th>$B_2$</th>
<th>$B_3$</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(μm²)</td>
<td>(μm²)</td>
<td>(μm²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sapphire</td>
<td>1.43135</td>
<td>0.65055</td>
<td>5.3414</td>
<td>0.00528</td>
<td>0.01424</td>
<td>325.01783</td>
</tr>
<tr>
<td>Fused silica</td>
<td>0.69617</td>
<td>0.40794</td>
<td>0.89748</td>
<td>0.00468</td>
<td>0.01351</td>
<td>97.934</td>
</tr>
<tr>
<td>Borosilicate</td>
<td>1.08079</td>
<td>0.23179</td>
<td>1.01047</td>
<td>0.00004</td>
<td>0.0004</td>
<td>103.56065</td>
</tr>
</tbody>
</table>

where $n$ is the refractive index, $\lambda$ is the wavelength, and $B_i$ and $C_i$ are empirically-determined coefficients different for each material. The Sellmeier coefficients of several glasses defined in NEXUS are shown in Table 6.1.

The transparency of glasses is typically expressed by the manufacturers as the transmittance fraction $T$ (i.e. number of photons passing through a certain thickness of material divided by the number of incident photons) as a function of the wavelength of light. On the other hand, Geant4 simulates bulk absorption of light as an exponential decay:

$$T \equiv \frac{I}{I_0} = e^{-x/\ell},$$

where $I_0$ is the intensity of the incident radiation, $I$ is the intensity of the radiation coming out of a layer of material of thickness $x$, and $\ell$ is the absorption length, which is the parameter that characterizes the material in Geant4. Therefore, given the transmittance of a material, we can easily obtain the absorption length by taking logarithms of both sides of Eq. (6.5):

$$\ell(\lambda) = -\frac{x}{\log T(\lambda)},$$

where $\lambda$ is the wavelength of light.

The wire meshes used as TPC electrodes are simulated as thin layers of a dielectric material with a refractive index matching that of xenon — therefore, no Fresnel reflection or refraction occurs at the interface between them — but with an absorption length such that the transmittance at normal incidence corresponds to the open area of the mesh. Bulk absorption mimics the obscuration of light with angle of incidence to good approximation for all but the largest angles with respect to the normal, as shown in Figure 6.6.

Polytetrafluoroethylene (PTFE) is extensively used as a reflector in noble gas detectors in order to improve their light collection efficiency. However, its
properties had not been measured in detail in the VUV region of the spectrum until very recently. Silva et al. measured the reflectivity of several samples of PTFE manufactured by different processes, obtaining results that range from about 47% to 66% for VUV light, and more than 95% for visible light. In addition, they observed that, although predominantly a Lambertian diffuser, PTFE has in the VUV a significant specular component of reflection.

The largest uncertainty in the optical model of NEXUS is the behaviour of tetraphenyl butadiene (TPB), the wavelength shifter (WLS) used in NEXT. TPB is used very often in noble gas detectors, and consequently its properties have been studied by many groups. These data, however, are neither easy to interpret in the context of the Geant4 WLS model — which was developed for wavelength shifting fibres and not for uniform coatings — nor perfectly consistent with each other. Therefore, we adapted the Geant4 model to our needs, adjusting its parameters using the work by Gehman et al. (2011) and our own measurements in the NEXT-DEMO prototype. In our model, a WLS material is characterized by its photon absorption and emission spectra, by its conversion efficiency and by a possible time delay between the absorption and re-emission of the photon. For TPB, we have assumed that all photons with

![Figure 6.6](image.png)
Figure 6.7. Total integrated fluorescence efficiency as a function of input photon wavelength, assuming a Lambertian angular distribution of reemitted light from both sides of the TPB film. Redrawn from Gehman et al. (2011) [218].

Figure 6.8. Visible re-emission spectrum (normalized to unit area) for a TPB coating illuminated with ultraviolet light. Redrawn from Gehman et al. (2011) [218].
a wavelength below 260 nm (this range covers the entire emission spectrum of xenon) are absorbed, and then re-emitted isotropically with the efficiency shown in Figure 6.7. The energy of the emitted photon is sampled from the spectrum shown in Figure 6.8. We assume that the TPB coatings (at least, for the thicknesses used in NEXT) are perfectly transparent to the emission wavelengths.

For each event, optical tracking ends with the probabilistic detection of photons arriving at each PMT photocathode. The quantum efficiency (QE) of a PMT — defined as the ratio of the number of photoelectrons emitted from the PMT’s photocathode to the number of incident photons — is typically measured in air or vacuum, and hence it includes the Fresnel reflection losses caused by the step change in refractive index at the PMT window. These measurements, therefore, cannot be directly used in NEXUS, since we are simulating explicitly the optical effects at the PMT windows. Consequently, using a dedicated simulation, we quantify the impact on the PMT’s detection efficiency of light losses due to optical effects, and increase then the measured QE by that amount.

6.6 Digitization

The digitization transforms the so-called photosensor hits — zero-suppressed time histograms that record the photons detected by each photosensor — into digital waveforms identical to those produced by the NEXT data acquisition system. The response of the photosensors is characterized by several input parameters that are set via a configuration file or read from the NEXT calibration database so that the simulated devices match real ones.

For a photomultiplier tube, the digitization starts with the smearing of the recorded signal according to the device’s charge resolution. Next, the time histogram is numerically convoluted with the single-photoelectron pulse shape, described by a Landau distribution with parameters extracted from real data. Finally, Gaussian noise is added to the resulting waveform. Figure 6.9 shows an example of a PMT signal pre and post-digitization.

In the case of silicon photomultipliers, the signal is integrated into time samples of 1 microsecond. No signal shaping is applied, since the large sampling time hides any possible effect. Lastly, random noise of intensity sampled from an input probability distribution is added to the waveform.
Figure 6.9. Digitization of the simulated signal of a photomultiplier tube. NEXUS records the photons detected by the PMT as a zero-suppressed time histogram (top panel). The digitization procedure modifies the histogram adding the effect of the PMT’s charge resolution, and the noise and pulse shaping of the electronics.
Figure 6.10. Primary-scintillation signal (SI) collected in the energy plane of NEXT-DEMO in terms of the reconstructed drift length for $\alpha$ decays from $^{222}$Rn in the xenon bulk. These data are fitted to simulated distributions in which the PTFE reflectivity has been varied in the range 90–100%. The curves for 92% (blue line) and 98% (red line) reflectivity are shown as an example. The best fit was obtained for the latter value.

6.7 Simulation results

6.7.1 Reflectivity of the NEXT-DEMO light tube

A precise knowledge of the reflectivity of PTFE is very important for estimating correctly the overall light collection of the detectors. Since scintillation and electroluminescence are isotropic emissions, an important fraction of the produced light will be reflected many times off the PTFE walls, meaning that a relatively small change in the reflectivity results in a large change in the detector response.

The reflectivity of the PTFE light tube of NEXT-DEMO was adjusted in NEXUS using primary-scintillation (SI) data from $\alpha$ decays in the bulk xenon. The total SI signal collected by the energy plane of the detector follows a linear dependence with the drift length (i.e. longitudinal position) reconstructed for the alpha particle, only deviating from this trend for positions close to the TPC cathode (see Figure 6.10, black points). Using NEXUS, we simulated $\alpha$ decays from $^{222}$Rn in the active volume of NEXT-DEMO, varying the reflectivity of PTFE in the range between 90% and 100%. The scintillation light yield was set to an arbitrary value in these simulation runs. The resulting distributions of SI
signal versus drift length were then fitted to the data, leaving free an overall scale factor, which corresponds to the ratio of the real, unknown scintillation yield for alphas to the arbitrary value chosen for the simulation runs. The best fit (minimum chi-square) corresponded to a reflectivity of 98%. Figure 6.10 shows as an illustration the comparison between the fitted distributions with 98% reflectivity (red line) and 92% (blue line).

### 6.7.2 Position of the tracking plane of NEXT-DEMO

During the commissioning of the tracking plane of NEXT-DEMO, the distance between the TPC anode and the face of the SiPMs was measured to be in the range of 5 to 10 mm (§5.2.1). The intensity of the light detected by a tracking-plane sensor is, essentially, a function of the solid angle subtended by the sensor’s active area at the light emission point, because reflected light, unlike in the case of the energy plane, does not contribute significantly to the signal. Therefore, a variation of up to a factor of 2 in the distance between anode and sensor plane should have a noticeable effect, especially for the sensors located not too close to the emission point.

Figure 6.11 compares the signal levels at the tracking-plane for xenon X-ray data and simulation. In the top panel, the signal recorded by the tracking-plane sensors, ordered in each event by signal intensity, and relative to the first sensor. Several simulation datasets were produced, varying the distance between anode and tracking plane. The best agreement was found for a distance of 7.5 mm. Nevertheless, data points are about 10% higher for SiPMs with order number higher than 5. The difference is attributed—at least, partially—to the behaviour of the optical model used to simulate the transmittance of the anode mesh, which absorbs in excess photons with high incidence angles (see Figure 6.6). The bottom panel of Fig. 6.11 shows the total charge collected by the tracking plane considering only the first \( n \) SiPMs, ordered according to their registered signals. Again, the best agreement between data and simulation was obtained for the dataset with 7.5 cm distance. The agreement, however, deteriorates if many SiPMs are included in the integral. This could indicate that a contribution to the background light levels at the tracking plane is not being simulated. Future work will address these differences between data and simulation.
Figure 6.11. Data-Monte Carlo comparison of the signal levels at the tracking plane of NEXT-DEMO. Top panel: Charge collected by the SiPMs (ordered by signal intensity) relative to the sensor with the highest signal. Bottom panel: Total charge (in photoelectrons) collected by the tracking plane considering only the first $n$ SiPMs (ordered by their signal intensity).
6.7. SIMULATION RESULTS

Figure 6.12. NEXT-DEMO $xy$-response map built using the energy depositions of simulated $K_{\alpha}$ X-rays emitted by xenon. The corrections are normalised to the value of an arbitrary bin. The differences in $xy$ response, of up to a factor of $\sim 7\%$, are smaller than those observed in real data. The pattern does not show any clear spatial, symmetric dependence. This is interpreted as a dominance of the sensor’s calibration residues on the detector response.

6.7.3 Energy resolution in NEXT-DEMO

The energy-resolution study presented in §5.2.1 was repeated for simulated datasets mimicking the conditions of the data. Figure 6.12 shows the $xy$-response map built using simulated xenon X-rays. The differences between $xy$ points, of up to 7%, are much smaller than those observed in real data. In addition, the pattern does not show any clear spatial, symmetric dependence. This is interpreted as a dominance of the sensor’s calibration residues on the detector response. The corrected energy spectrum is shown in Figure 6.13. From fits to the $K_{\alpha}$ xenon x-ray peak and the photopeak, the following energy resolutions were extracted: 4.68% FWHM at 30 keV and 1.17% FWHM at 667 keV. These results are better than those obtained with real data, but, still, somewhat worse than expected for the Fano factor of gaseous xenon. More detailed studies of this issue are underway.
Figure 6.13. Simulated energy spectrum from a $^{137}$Cs source after corrections.
Sensitivity of NEXT-100 to neutrinoless double beta decay

7.1 Introduction

In previous chapters we have seen that a xenon gas electroluminescent TPC can offer high-resolution calorimetry, charged-particle tracking and viable scalability to large detector masses. This unique combination of features, which will be exploited by the NEXT Collaboration with the NEXT-100 detector, has enormous potential for neutrinoless double beta decay searches. In this chapter we provide a quantitative assessment of that potential by estimating the sensitivity of NEXT-100 to the effective Majorana neutrino mass, $m_{\beta\beta}$.

Some of the factors on which the sensitivity of NEXT-100 depends are fixed by design, such as the isotope chosen ($^{136}\text{Xe}$) or the available source mass (100 kg of xenon enriched to 91% in $^{136}\text{Xe}$). Other factors, like the energy resolution or the tracking performance, can be extrapolated from results of the prototypes (§5.2). Finally, the levels of the potential backgrounds and the discrimination power of the detector can be estimated with information on the radioactivity of the construction materials and the help of Monte Carlo simulations. These last two factors — the background model and the discrimination capabilities of NEXT — are the focus of this chapter.

The relevance of any potential background source in NEXT depends on its probability to generate a signal-like track in the active volume of the detector with energy around the $Q$ value of $^{136}\text{Xe}$. In principle, charged particles (muons, betas, etc.) entering the detector can be eliminated with essentially perfect efficiency defining a small veto region (of a few centimetres) around the boundaries of the active volume. Confined tracks generated by external neutral particles, such as high-energy gamma rays, or by internal contamination in the xenon gas can be suppressed taking advantage of the distinctive energy-deposition pattern of signal events, illustrated in Figure 7.1. Below the so-called critical energy (about 12 MeV in gaseous xenon [31]), electrons
CHAPTER 7. SENSITIVITY OF NEXT-100 TO NEUTRINOLESS DOUBLE BETA DECAY

Figure 7.1. Monte Carlo simulation of signal ($0\nu\beta\beta$ decay of $^{136}$Xe) and background (single electron of energy equal to the $Q$ value of $^{136}$Xe) events in gaseous xenon at 15 bar. The ionization tracks left by signal events feature large energy deposits, or blobs, at both ends.

Figure 7.2. Energy loss for electrons in argon and xenon as a function of energy. Note the steep increase in $dE/dx$ below 50 keV that results in the end-of-track blobs. Redrawn from Thomas et al. (1988) [224].
(and positrons) lose their energy at a relatively fixed rate until they become non-relativistic. At about that time, their effective dE/dx rises (see Figure 7.2), mostly due to the occurrence of strong multiple scattering, and the particles lose the remainder of their energy in a relatively short distance, generating a blob. Double beta decay events consist of two electrons emitted from a common vertex. Their reconstructed tracks, therefore, feature blobs at both ends. Background tracks, however, are generated by single electrons, thus having only one end-of-track blob.

In the remainder of this chapter, after discussing the various potential components of the background spectrum of NEXT-100 in the energy region around $Q_{\beta\beta}$, we will evaluate the signal detection efficiency and background rejection power achieved with a basic set of selection criteria for $0\nu\beta\beta$ events applied to simulation datasets. Finally, using these computed factors and measurements of the activity of the NEXT-100 materials and components, we will estimate the background rate of the experiment and its sensitivity to neutrinoless double beta decay.

### 7.2 Sources of background in the NEXT experiment

#### 7.2.1 High-energy gamma rays

Natural radioactivity in detector materials and surroundings is, as in most other $0\nu\beta\beta$-decay experiments, the main source of background in NEXT. In particular, the hypothetical $0\nu\beta\beta$ peak of $^{136}$Xe ($Q_{\beta\beta} = 2458.1 \pm 0.3$ keV \cite{119,150}) lies in between the photopeaks of two high-energy gammas emitted after the $\beta$ decays of $^{214}$Bi and $^{208}$Tl, intermediate products of the uranium and thorium series, respectively (see Figs. 3.8 and 3.9).

The daughter isotope of $^{214}$Bi, $^{214}$Po, emits a number of de-excitation gammas with energies around and above the $Q$ value of $^{136}$Xe \cite{225}. Most of these gamma lines have very low intensity, and hence their contribution to the background rate is negligible. The gamma of 2447 keV (1.57% intensity \cite{225}), however, is very close to $Q_{\beta\beta}$, and its photoelectric peak overlaps the signal peak even for energy resolutions as good as 0.5% FWHM. The decay product of $^{208}$Tl, $^{208}$Pb, emits a de-excitation photon of 2615 keV with an intensity of 99.75% \cite{225}. Electron tracks from its photopeak can lose energy via bremsstrahlung and fall in the region of interest (ROI) around $Q_{\beta\beta}$ defined by the energy resolution of the detector. Additionally, even though the Compton edge of the 2.6-MeV gamma is at 2382 keV, well below $Q_{\beta\beta}$, the Compton-
scattered photon can generate other electron tracks close enough to the initial Compton electron so that they are reconstructed as a single track with energy around $Q_{\beta\beta}$.

The NEXT Collaboration is carrying out a thorough campaign of material screening and selection using gamma-ray spectroscopy (with the assistance of the LSC Radiopurity Service) and mass spectrometry techniques (ICP-MS and GD-MS). Table 7.1 collects the measurements of the specific activity of $^{208}$Tl and $^{214}$Bi in the most relevant materials and components of NEXT-100, and Table 7.2 details the radioactivity budget of NEXT-100 separated by detector subsystems.

The rock walls of the underground laboratory are another rather intense source of high-energy gammas due to the presence of trace radioactive contaminants in their composition. The total gamma flux in Hall A at LSC is $1.06 \pm 0.24 \text{ cm}^{-2} \text{s}^{-1}$, with contributions from $^{40}$K ($0.52 \pm 0.23 \text{ cm}^{-2} \text{s}^{-1}$), $^{238}$U ($0.35 \pm 0.03 \text{ cm}^{-2} \text{s}^{-1}$) and $^{232}$Th ($0.19 \pm 0.04 \text{ cm}^{-2} \text{s}^{-1}$) [229]. Nevertheless, the external lead shield (§5.3.7) will attenuate this flux by more than 4 orders of magnitude, making its contribution to the final background rate negligible.

### 7.2.2 Radon

Airborne radon (§3.4.3) constitutes another dangerous source of high-energy gamma rays, mostly from the decay of $^{214}$Bi, since $^{222}$Rn is, in general, much more abundant than $^{220}$Rn. The measured activity of $^{222}$Rn at LSC (Hall A) varies between 60 and 80 Bq/m$^3$ [230]. Left at this level, radon would represent an intolerably high background for the NEXT experiment. Possible cures for the problem include the installation of a radon mitigation system at LSC such as those used in the NEMO-3 [231] and the DarkSide experiments [232], or flushing continuously the internal volume of the lead castle shield with a clean gas (nitrogen, typically). A reduction of a factor of 100 or more in the activity of radon in the vicinity of the detector should be possible with either of the two methods.

Radon can also emanate from detector components and be transported to the active volume through the gas circulation. The $\alpha$ decays of radon (either $^{220}$Rn or $^{222}$Rn) in the bulk xenon do not represent a background: they have energies higher than $Q_{\beta\beta}$, and their very short tracks are easily identified [155,197]. These alphas are useful, however, to monitor the activity of radon in the xenon gas [101,197]. The progeny of radon is positively charged and will drift toward the TPC cathode. Therefore, a majority of the subsequent
Table 7.1. Specific activity of $^{208}\text{Tl}$ and $^{214}\text{Bi}$ in the most relevant materials and detector components used in the NEXT experiment. Three materials (fused silica, sapphire and the field-cage resistors) have not been screened yet with sufficient precision; therefore, we use instead the measurements by the EXO Collaboration. The activities of fused silica, sapphire and the field-cage resistors were GDMS, ICPMS and NAA results were derived from Th and U concentrations. High-purity germanium (HPGe) $\gamma$-ray spectroscopy results correspond, whenever possible, to the lower parts of the natural decay chains. The limits are given at 95% CL. The figures in parentheses after the measurements give the 1-standard-deviation uncertainties in the last digits. The abbreviations used to refer to the NEXT-100 detector subsystems have the following meaning: EP: energy plane; TP: tracking plane; FC: electric-field cage; PV: pressure vessel; IS: inner shielding; OS: outer shielding.

<table>
<thead>
<tr>
<th>Material</th>
<th>Subsystem</th>
<th>Technique</th>
<th>Units</th>
<th>$^{208}\text{Tl}$</th>
<th>$^{214}\text{Bi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copper (CuA1)</td>
<td>IS, EP, FC</td>
<td>GDMS</td>
<td>mBq/kg</td>
<td>$&lt;0.0014$</td>
<td>$&lt;0.012$</td>
</tr>
<tr>
<td>Fused silica</td>
<td>FC</td>
<td>NAA</td>
<td>mBq/kg</td>
<td>0.0097(18)</td>
<td>0.07(3)</td>
</tr>
<tr>
<td>Kapton board</td>
<td>TP</td>
<td>HPGe</td>
<td>mBq/unit</td>
<td>0.014(2)</td>
<td>0.03(1)</td>
</tr>
<tr>
<td>Lead</td>
<td>OS</td>
<td>GDMS</td>
<td>mBq/kg</td>
<td>0.034(7)</td>
<td>0.35(7)</td>
</tr>
<tr>
<td>PMT R11410-10</td>
<td>EP</td>
<td>HPGe</td>
<td>mBq/PMT</td>
<td>0.25(13)</td>
<td>$&lt;0.96$</td>
</tr>
<tr>
<td>Polyethelene</td>
<td>FC</td>
<td>ICPMS</td>
<td>mBq/kg</td>
<td>$&lt;0.0076$</td>
<td>$&lt;0.062$</td>
</tr>
<tr>
<td>Resistor (1 GΩ)</td>
<td>FC</td>
<td>HPGe</td>
<td>mBq/unit</td>
<td>0.000011(6)</td>
<td>0.00009(4)</td>
</tr>
<tr>
<td>Sapphire</td>
<td>EP</td>
<td>NAA</td>
<td>mBq/unit</td>
<td>0.04(1)</td>
<td>$&lt;0.31$</td>
</tr>
<tr>
<td>Steel (Type 316Ti)</td>
<td>PV</td>
<td>GDMS, HPGe</td>
<td>mBq/kg</td>
<td>$&lt;0.15$</td>
<td>$&lt;0.46$</td>
</tr>
<tr>
<td>SiPM SensL</td>
<td>TP</td>
<td>HPGe</td>
<td>mBq/unit</td>
<td>$&lt;0.0007$</td>
<td>$&lt;0.0027$</td>
</tr>
</tbody>
</table>
Table 7.2. Radioactivity budget of the NEXT-100 detector. For most subsystems, we only have upper limits (at 95% CL) to their activity. The figures in parentheses correspond to the 1-sigma uncertainty in the last digit. See Table 7.1 for the specific activities of the materials and components listed here.

<table>
<thead>
<tr>
<th>Detector subsystem</th>
<th>Material</th>
<th>Quantity</th>
<th>$^{208}\text{Tl}$ (mBq)</th>
<th>$^{214}\text{Bi}$ (mBq)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pressure vessel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Steel 316Ti</td>
<td>1310 kg</td>
<td>&lt; 197</td>
<td>&lt; 603</td>
</tr>
<tr>
<td><strong>Energy plane</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMTs</td>
<td>R11410-I0</td>
<td>60 units</td>
<td>12(3)</td>
<td>&lt; 56</td>
</tr>
<tr>
<td>PMT enclosures</td>
<td>Copper CuA1</td>
<td>60×4.3 kg</td>
<td>&lt; 0.36</td>
<td>&lt; 3.1</td>
</tr>
<tr>
<td>Enclosure windows</td>
<td>Sapphire</td>
<td>60×0.14 kg</td>
<td>0.34(8)</td>
<td>&lt; 2.6</td>
</tr>
<tr>
<td>Support plate</td>
<td>Copper CuA1</td>
<td>408 kg</td>
<td>&lt; 0.6</td>
<td>&lt; 5</td>
</tr>
<tr>
<td><strong>Tracking plane</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiPMs</td>
<td>SENSIL 1 mm$^2$</td>
<td>107×64 units</td>
<td>&lt; 5</td>
<td>&lt; 18</td>
</tr>
<tr>
<td>Boards</td>
<td>Kapton FPC</td>
<td>107 units</td>
<td>1.5(2)</td>
<td>3.2(1.1)</td>
</tr>
<tr>
<td><strong>Field cage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrel</td>
<td>Polyethylene</td>
<td>128 kg</td>
<td>&lt; 1</td>
<td>&lt; 8</td>
</tr>
<tr>
<td>Shaping rings</td>
<td>Copper CuA1</td>
<td>120×3 kg</td>
<td>&lt; 0.5</td>
<td>&lt; 4</td>
</tr>
<tr>
<td>Electrode rings</td>
<td>Steel 316Ti</td>
<td>2×5 kg</td>
<td>1.5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Anode plate</td>
<td>Fused silica</td>
<td>9.5 kg</td>
<td>0.092(17)</td>
<td>0.7(3)</td>
</tr>
<tr>
<td>Resistor chain</td>
<td>1-GΩ resistors</td>
<td>240 units</td>
<td>&lt; 0.0026</td>
<td>&lt; 0.020</td>
</tr>
<tr>
<td><strong>Shielding</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner shield</td>
<td>Copper CuA1</td>
<td>9210 kg</td>
<td>&lt; 13</td>
<td>&lt; 111</td>
</tr>
<tr>
<td>Outer shield</td>
<td>Lead</td>
<td>60700 kg</td>
<td>2060(430)</td>
<td>21300(4300)</td>
</tr>
</tbody>
</table>

$^{214}\text{Bi}$ and $^{208}\text{Tl}$ beta decays will occur on the cathode rather than in the active volume. These cathode events are equivalent to other background sources close to the active volume: if the $\beta$ particle enters the active volume, the event can then be vetoed; otherwise, the de-excitation gamma rays interacting in the xenon can generate background tracks. In addition, a small fraction (0.2%) of the $^{214}\text{Bi}$ beta decays occurring in the xenon bulk will produce an electron track with energy around $Q_{\beta\beta}$. Luckily, the disintegration of $^{214}\text{Bi}$ is followed shortly after by the $\alpha$ decay of $^{214}\text{Po}$ ($T_{1/2} = 164 \mu s$). The detection of this so-called Bi-Po coincidence can be used to identify and suppress with high efficiency these background events. In the design of NEXT-100 we have tried to avoid or minimize the use of materials and components known to emanate...
radon in high rates, such as plastics, cables or certain seals. Nevertheless, estimating a priori the radon emanation rate and consequent activity in the xenon is difficult, since the available data are scant and have been acquired in conditions sometimes very different to those of NEXT-100. The Collaboration is designing a radon trap for the gas circulation system in case the activity of radon in the xenon proves to be too high.

7.2.3 **Muons and neutrons**

Neutrons can activate the enriched xenon and produce $^{137}$Xe, a $\beta$ emitter with a Q value of 4173 keV and a half-life of 3.8 minutes [225]. The fraction of $\beta$ tracks with energy around $Q_{\beta\beta}$ is, therefore, a potential background. In the case of the EXO-200 experiment, for instance, $^{137}$Xe is responsible for about one quarter of the background events [62]. The production rate of $^{137}$Xe at the Laboratorio Subterráneo de Canfranc is given by

$$r = \Phi \cdot \sigma \cdot \frac{\rho N_A}{M}, \quad (7.1)$$

where $\Phi = (3.44 \pm 0.35) \times 10^{-6}$ cm$^{-2}$ s$^{-1}$ [233] is the neutron flux at Hall A of LSC, $\sigma \approx 1$ mb $= 10^{-27}$ cm$^2$ [234] is the neutron-capture cross-section of $^{136}$Xe, $\rho = 0.08663$ g cm$^{-3}$ is the density of xenon at 15 bar and 25 °C, $M \approx 136$ g mol$^{-1}$ is the atomic weight of $^{136}$Xe, and $N_A$ is the Avogadro constant. Substituting all these quantities in the above equation, we obtain that $r = 1.35 \times 10^{-6}$ m$^{-2}$ s$^{-1}$. Therefore, in a year ($3.16 \times 10^7$ seconds) approximately 50 atoms of $^{137}$Xe will be produced in the active volume of NEXT-100 (1.15 m$^3$) by neutron activation. However, only 0.5% of them will produce a $\beta$ electron with energy in the region of interest around the Q value of $^{136}$Xe.

The backgrounds associated to the muon flux at LSC ($\sim 5 \times 10^{-3}$ m$^{-2}$ s$^{-1}$ [126]) appear to be largely negligible, according to our estimations. Nevertheless, if needed, they could be further suppressed installing active veto detectors around the lead castle shield.

7.3 **Signal detection efficiency and background rejection power**

7.3.1 **Simulation**

The study presented here makes use of large datasets — of the order of $10^{10}$ events of signal and background — produced with the fast-simulation mode
of NEXUS (§6.2.2), which provides as output for each event the total energy deposited in the active volume, smeared by the detector’s energy resolution function, and a collection of cubic voxels that represent the reconstructed ionization tracks. The size of the voxels (1 × 1 × 1 cm$^3$) and the energy resolution (0.75% FWHM at the Q value of $^{136}$Xe) are extrapolated from the results of the R&D phase (§5.2).

The Decay0 event generator [209], described in §6.4, was used for the production of millions of events of both 2νββ and 0νββ decay. These events were read afterwards by the simulation program and given a random initial position within the xenon gas volume. The decays of $^{208}$Tl and $^{214}$Bi were simulated using the radioactive-decay module included in Geant4. For each isotope and each detector subsystem considered as source of background in the detector, one billion events were simulated with initial positions uniformly distributed within the corresponding volume. In all cases, events with total energy deposited in the active volume below 2.3 MeV were discarded already at the simulation stage.

7.3.2 Voxelization and graph-search algorithms

The collection of voxels of each simulated event can be regarded as a graph, that is, a set of nodes and the links that connect them. A graph of n nodes is characterized by its distance matrix, a square ($n \times n$) matrix that contains the distance between any pair of nodes. Specifically, the matrix element $d_{ij}$, which indicates the distance between nodes $i$ and $j$, is defined in the following way:

- $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$ if the cubic voxels corresponding to nodes $i$ and $j$ are in contact (i.e. they share one face, one edge or one vertex). For cubic voxels of 1 cm$^3$, like in our case, there are only three possible values: 1, $\sqrt{2}$ and $\sqrt{3}$ cm.

- $d_{ij} = \infty$ (or an arbitrary, large number) if the voxels are not in contact.

- $d_{ij} = 0$ if $i = j$.

Once the distance matrix of a graph has been built, it can be used to identify connected subsets of nodes via graph-search algorithms such as breadth-first search (BFS) [235]. In our case, these connected components correspond, in general, to the continuous ionization tracks left by charged particles in the active volume of the detector. The BFS algorithm can also be used to calculate
the length of the (shortest) path between any pair of voxels in a connected subset. The longest of such paths should connect the ends of a track, that is, the voxels that in a $0\nu\beta\beta$-decay track correspond to the blobs. Figure 7.3 shows, for simulated electron tracks of energy equal to the $Q$ value of $^{136}\text{Xe}$, the Euclidean distance between the true ends of a track and those found with the BFS algorithm. The most frequent distance in this histogram corresponds to about 6 mm, that is, less than the size of a voxel (1 cm). In other words, in most cases, we are able to locate the ends of a track with an accuracy of one or two voxels.

### 7.3.3 Event selection

In what follows, we describe a basic set of selection criteria for $0\nu\beta\beta$ events detected and reconstructed in NEXT-100. Applying this event selection to the simulation datasets discussed earlier in this section (§7.3.1), we can obtain an estimate of the signal detection efficiency and background rejection power of the NEXT-100 detector. This selection also serves as a benchmark for more sophisticated algorithms for the discrimination of signal and background that are currently under development.
The criteria to accept a reconstructed event as a 0νββ-decay candidate are the following:

1. The event consists of one single reconstructed track confined within the fiducial volume of the detector — defined by excluding a region of 2 cm around the boundaries of the active volume — and with energy between 2.4 and 2.5 MeV.

2. The reconstructed track features a blob at both ends.

3. The energy of the event is within the region of interest (ROI) around \( Q_{\beta\beta} \).

The definition of a fiducial volume has two purposes: it rejects all charged backgrounds entering the detector, and it discards those events in which the tracked particles may have left the active volume, depositing part of their energy in passive materials. The size of the excluded region, 2 centimetres around the boundaries of the active volume, takes into account the voxel size (which, in turn, depends on the spatial resolution of the detector) and the higher inhomogeneity of the electric field near the edges of the field cage (which may affect the quality of the track reconstruction in that region). In practical terms, this fiducial cut is implemented demanding that none of the voxels located in the vetoed region contain energy above the detection threshold of the tracking plane (set, conservatively, to 10 keV).

The requirement for the accepted events to have one and only one reconstructed track takes advantage of the very different track multiplicities — that is, the number of reconstructed tracks per event — of signal and background (see Figure 7.4). Approximately 70% of the signal events satisfy the single-track condition, whereas only 10% of \(^{208}\text{Tl}\) and \(^{214}\text{Bi}\) events do so. Taking into account that, according to Eq. (3.21), maximizing the ratio \( \epsilon / \sqrt{b} \) — where \( \epsilon \) and \( b \) are, respectively, the acceptances of signal and background — optimizes the experimental sensitivity to \( m_{\beta\beta} \), it would only be worth accepting events with more than one track if the fraction of background events passing the cut were such that

\[
b' \leq (\epsilon' / \epsilon)^2 b,
\]

where the unprimed and primed quantities are, respectively, the acceptances of the default (that is, track multiplicity equal to 1) and the alternative (i.e. higher track multiplicities) selection cuts. For the values shown in Fig. 7.4 (\( \epsilon = 0.71 \),...
The second selection criterion exploits the characteristic energy-deposition pattern of $0\nu\beta\beta$-decay tracks (see Fig. 7.1), which feature a blob at both ends due to the effective rise in the $dE/dx$ of electrons with low momentum. We define the energy of a blob as the total energy contained in all the voxels whose center is at a maximum distance of $\sqrt{2}$ cm with respect to the one reconstructed as track end. From the point of view of the discrimination power of the cut, this definition proved to be the optimal among those considered (integration radii between 0 and $\sqrt{3}$ cm). Figure 7.5 shows the probability distributions of signal and background events in terms of the energies of the end-of-track blobs. The populations of signal and background are clearly separated. Additionally, the distributions of $^{208}$Tl and $^{214}$Bi are very similar, indicating that they correspond to the same type of events (single-electron tracks with energy around $Q_{\beta\beta}$). A simple and reasonably clean selection cut could be established with a threshold around 0.2 MeV on the energy of the less energetic blob. According to the Neyman-Pearson lemma [236], however, the most efficient selection criterion
Figure 7.5. Probability distribution of $0\nu\beta\beta$ (top panel), $^{208}$Tl (centre) and $^{214}$Bi (bottom) events in terms of the energies of the end-of-track blobs. The blob labelled as ‘1’ corresponds to the more energetic one, whereas ‘blob 2’ corresponds to the less energetic of the two.
Figure 7.6. Likelihood-ratio distributions for signal (red, solid histogram) and background ($^{208}$Tl: grey, solid histogram; $^{214}$Bi: grey, dotted histogram).

... is based on the likelihood ratio test statistic:

\[ \mathcal{L} = \frac{P(E_1, E_2 \mid 0\nu\beta\beta)}{P(208\text{Tl}) \cdot P(E_1, E_2 \mid 208\text{Tl}) + P(214\text{Bi}) \cdot P(E_1, E_2 \mid 214\text{Bi})}, \]  

(7.3)

where \( P(E_1, E_2 \mid H) \) is the probability for an event with blob energies \( E_1 \) and \( E_2 \) to be signal (\( H \equiv 0\nu\beta\beta \)) or background (\( H \equiv 208\text{Tl} \) or \( H \equiv 214\text{Bi} \)), and \( P(208\text{Tl}) \) and \( P(214\text{Bi}) \) are the \textit{a priori} probabilities for a background event to be either \( 208\text{Tl} \) or \( 214\text{Bi} \). In other words, \( P(E_1, E_2 \mid H) \) is the probability given in Fig. 7.5 and \( P(208\text{Tl}) \) and \( P(214\text{Bi}) \) are the relative initial abundances of each background source. Once the likelihood ratio (or the natural logarithm of the likelihood function, the so-called log-likelihood, which is, in general, more convenient to work with) is computed for all values of \( E_1 \) and \( E_2 \) (see Figure 7.6), we choose as selection threshold the value of \( \mathcal{L} \) that maximizes the figure of merit \( \epsilon/\sqrt{b} \).

A similar procedure can be followed with the third selection criterion in order to decide on the optimal region of interest in the energy spectrum. In this case, the likelihood ratio is defined as follows:

\[ \mathcal{L} = \frac{P(E \mid 0\nu\beta\beta)}{P(208\text{Tl}) \cdot P(E \mid 208\text{Tl}) + P(214\text{Bi}) \cdot P(E \mid 214\text{Bi})}, \]  

(7.4)
Figure 7.7. Energy spectra of signal (red, solid curve) and background (\(^{208}\)Tl: grey, dashed distribution; \(^{214}\)Bi: grey, dotted distribution; total: grey, solid distribution) in the region of interest (ROI) around \(Q_{\beta\beta}\). The optimal ROI — the energy range that maximizes the ratio of the signal efficiency over the square root of the background rate — is indicated by the shaded, blue region. The signal strength represented here corresponds to a neutrino Majorana mass of 200 meV, while the backgrounds are scaled to their expected values in NEXT-100 (\(6 \times 10^{-4}\) cts keV\(^{-1}\) kg\(^{-1}\) yr\(^{-1}\)), assuming an exposure of 91 kg yr.

where \(P(E|H)\) is the probability of an event of energy \(E\) of being signal \((H \equiv 0\nu\beta\beta)\) or background \((H \equiv {^{208}\text{Tl or } 214}\text{Bi})\). Figure 7.7 shows the distribution of signal and background around \(Q_{\beta\beta}\) and the region of interest that maximizes the quantity \(\varepsilon/\sqrt{b}\), selected using the likelihood ratio defined above.

Table 7.3 lists the computed acceptances for signal and background of the selection criteria described above. The natural radioactive backgrounds, \(^{208}\)Tl and \(^{214}\)Bi, are suppressed by more than 6 orders of magnitude, and the contribution of the two-neutrino double beta decay to the background rate is completely negligible. The cuts yield a signal efficiency of 28%. Note, however, that approximately half of the events are lost already in the first selection cut: 88% of the events are contained within the fiducial volume of the detector, 71% have one single track, and 76% of them have reconstructed energy above 2.4 MeV (the \(0\nu\beta\beta\) spectrum has a tail extending to low energies composed of
Table 7.3. Acceptance of the selection criteria for 0νββ-decay events described in the text. The values for 208Tl and 214Bi correspond to one of the dominant sources of background in the detector.

<table>
<thead>
<tr>
<th>Selection criterion</th>
<th>0νββ</th>
<th>2νββ</th>
<th>208Tl</th>
<th>214Bi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiducial, single track</td>
<td>0.4759</td>
<td>8.06 × 10^{-9}</td>
<td>2.83 × 10^{-5}</td>
<td>1.04 × 10^{-5}</td>
</tr>
<tr>
<td>E ∈ [2.4, 2.5] MeV</td>
<td>0.6851</td>
<td>0.6851</td>
<td>0.1141</td>
<td>0.105</td>
</tr>
<tr>
<td>Track with 2 blobs</td>
<td>0.8661</td>
<td>3.89 × 10^{-5}</td>
<td>0.150</td>
<td>0.457</td>
</tr>
<tr>
<td>Energy ROI</td>
<td>0.2824</td>
<td>2.15 × 10^{-13}</td>
<td>4.9 × 10^{-7}</td>
<td>4.9 × 10^{-7}</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

events with missing energy in the form of bremsstrahlung radiation).

7.4 Predicted background rate and sensitivity of NEXT-100

The contribution of each detector subsystem to the overall background rate of NEXT-100 is shown in Table 7.4. These rates are obtained dividing the initial activities of 208Tl and 214Bi by the corresponding background rejection factors (defined as the inverse of the background acceptance resulting from the 0νββ-decay event selection described in the previous section). They are also represented graphically in Figure 7.8. The photosensors are, by far, the dominant source of background in NEXT-100. Notice, however, that our knowledge is, in any case, quite uncertain, given that for most background sources we only have at present a limit to their activity. This is, in fact, a problem common to all 0νββ-decay experiments, and it will be even more serious for the experiments of the tonne scale, which will require materials and components of higher radiopurity.

Table 7.5 shows the contributions grouped into six major subsystems. The background from 214Bi is 4.3 times more abundant than the background from 208Tl. The overall background rate estimated for NEXT-100 is

\[ < 5.86 \times 10^{-4} \text{ counts/(keV kg year).} \] (7.5)

This rate includes only radioactive backgrounds from detector materials and components. All other sources of background are expected to contribute at the level of 10^{-5} keV^{-1} kg^{-1} yr^{-1} or below:
<table>
<thead>
<tr>
<th>Detector subsystem (key (1))</th>
<th>Activity (MBq)</th>
<th>Rejection factor (2)</th>
<th>Total (10^3 cts/keV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer shield</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner shield</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The fourth column contains the contribution to the background rate of NEXT-100 predicted for each subsystem of the detector considered. The last two columns show the background rate estimated for each subsystem (i.e., the ratio of the previous quantities) expressed in 10^3 cts/keV. For most subsystems, we only have upper limits to their induced background rate in those cases where we have a positive measurement.

Table 7.4 Contribution to the background rate of NEXT-100 predicted for each subsystem of the detector considered in our simulation and the columns contain the rejection factors computed with the detector simulation. The last two columns show the background rate estimated for each subsystem (i.e., the ratio of the previous quantities) expressed in 10^3 cts/keV. For most subsystems, we only have upper limits to their induced background rate in those cases where we have a positive measurement.
Figure 7.8. Contribution to the background rate of NEXT-100 of the different detector subsystems considered in our background model. An asterisk (*) next to a bar indicates that the contribution corresponds to a positive measurement of the activity of the material.
Table 7.5. Contribution of major subsystems to the expected background rate of NEXT-100, expressed in $10^{-4}$ counts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$.

<table>
<thead>
<tr>
<th>Detector subsystem</th>
<th>$^{208}$Tl</th>
<th>$^{214}$Bi</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure vessel</td>
<td>$&lt; 0.23$</td>
<td>$&lt; 0.07$</td>
<td>$&lt; 0.31$</td>
</tr>
<tr>
<td>Energy plane</td>
<td>$&lt; 0.38$</td>
<td>$&lt; 1.31$</td>
<td>$&lt; 1.69$</td>
</tr>
<tr>
<td>Tracking plane</td>
<td>$&lt; 0.38$</td>
<td>$&lt; 1.27$</td>
<td>$&lt; 1.65$</td>
</tr>
<tr>
<td>Electric-field cage</td>
<td>$&lt; 0.14$</td>
<td>$&lt; 0.93$</td>
<td>$&lt; 1.07$</td>
</tr>
<tr>
<td>Inner shield</td>
<td>$&lt; 0.17$</td>
<td>$&lt; 0.70$</td>
<td>$&lt; 0.87$</td>
</tr>
<tr>
<td>Outer shield</td>
<td>$0.025(13)$</td>
<td>$0.25(14)$</td>
<td>$0.28(14)$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$&lt; 1.33$</td>
<td>$&lt; 4.53$</td>
<td>$&lt; 5.86$</td>
</tr>
</tbody>
</table>

- The activity of airborne radon in the vicinity of the detector — which translates, ultimately, into $^{214}$Bi activity on the internal surface of the lead shield and on the external surface of the vessel — will be reduced by at least two orders of magnitude with respect to the activity in the experimental hall of LSC ($\sim 80$ Bq/m$^3$) thanks to the use of a radon mitigation machine or the removal of the air inside the lead shield (3.2 m$^3$) with clean nitrogen. The computed rejection factor for this source of background is $2 \times 10^9$, resulting in a background rate of about $10^{-5}$ keV$^{-1}$ kg$^{-1}$ yr$^{-1}$ for a $^{222}$Rn activity of 0.5 Bq/m$^3$ (see Figure 7.9 for other values of the specific activity of $^{222}$Rn in the range between $10^{-4}$ and $10^2$ Bq/m$^3$).

- Radon contamination in the xenon gas causes two different types of background events (§7.2.2): $\beta$ tracks from the decay of $^{214}$Bi in the active volume, and photoelectrons generated by gamma rays emitted, for the most part, from the TPC cathode following the decay of $^{214}$Bi. In the EXO-200 TPC, the latter type of events constitute about 80% of the measured activity of $^{222}$Rn in the liquid xenon, while the former make up the remaining 20% [101]. The rejection power against both types of background events is similar, approximately $2.5 \times 10^6$. In the case of the $\beta$ decays of $^{214}$Bi in the xenon bulk, we have assumed that Bi-Po tagging — i.e. the coincident detection in an event of the $\beta$ emitted in the decay of $^{214}$Bi and the alpha emitted by $^{214}$Po shortly after — can be done with high efficiency ($\gtrsim 99\%$). Figure 7.9 (red lines) shows the background rate generated in NEXT-100 by this internal contamination of radon in
Figure 7.9. Background rate induced in NEXT-100 by airborne radon and radon contamination in the xenon gas (labelled as internal) in terms of the activity of $^{222}$Rn. In order for this background to contribute, at most, at the level of $10^{-5}$ keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, radon activities in the xenon gas below a few mBq per cubic metre will be required. The EXO-200 detector, which has been operating without a radon suppression system, has measured, for instance, an activity of $^{222}$Rn of that order in their xenon volume: $(3.65 \pm 0.37) \mu$Bq/kg \cite{101}. Similarly, the radon activity of the NEMO-3 tracking gas was measured to be about 5 mBq/m$^3$ \cite{114}.

- Out of the 50 atoms of $^{137}$Xe produced on average every year by neutron activation of $^{136}$Xe (§7.2.3), 0.25 of them will decay emitting a $\beta$ track with energy within our region of interest. These events will be suppressed by about a factor of 10 by the 2-blobs selection criterion, yielding a background rate of approximately $9 \times 10^{-6}$ keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. Nevertheless, the neutron flux can be attenuated by several orders of magnitude using polyethylene shielding of a few tens of centimetres in thickness.

The sensitivity of NEXT-100 to neutrinoless double beta decay — calculated following the Feldman-Cousins prescription for the construction of confidence intervals described in §3.2 — is shown in Figure 7.10. In the top panel, the sensitivity (at 90% CL) to the half-life (red, solid curve) and the corre-
sponding sensitivity to $m_{\beta \beta}$ (blue, dashed curves) for the largest and smallest NME calculations (see Table 3.3) are represented in terms of the exposure, assuming a signal detection efficiency of 28% and a background rate of $6 \times 10^{-4}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. The bottom panel shows, for an exposure of 100 kg-year the variation of the sensitivity with respect to the background rate in the range between $10^{-5}$ and $10^{-4}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. NEXT-100 will reach, after a 3-years run, a half-life sensitivity of approximately $4.6 \times 10^{25}$ years, corresponding to the $m_{\beta \beta}$ range 84–182 meV.
7.4. PREDICTED BACKGROUND RATE AND SENSITIVITY OF NEXT-100

Figure 7.10. Sensitivity (at 90% CL) of NEXT-100 to neutrinoless double beta decay. The (crimson) solid curves represent the half-life sensitivity, while the (blue) dashed curves correspond to the $m_{\beta\beta}$ sensitivity for three different NME calculations (from top to bottom: EDF [4], IBM-2 [72], ISM [69]). Top: Sensitivity of NEXT-100 as a function of the exposure. Bottom: Sensitivity of NEXT-100 after a 3-years run (equivalent to an exposure of 273 kg yr) as a function of the achieved background rate.
Towards the tonne scale

8.1 Reach of the current generation of experiments

The importance of the worldwide experimental program searching for neutrinoless double beta decay can hardly be overstated: the discovery of the radioactive process is the most promising way — perhaps the only way — of determining the nature of neutrino mass and proving the violation of total lepton number. The current generation of experiments consists of projects using source masses of tens to hundreds of kilograms and a variety of detection techniques. Three of these experiments (EXO-200, KamLAND-Zen and GERDA) are operating already and have released physics results (§3.3), and at least five other (CUORE, MAJORANA, NEXT-100, SNO+ and SuperNEMO) plan to start taking data in the next few years.

The basic operational parameters of all the experiments of the current generation are listed in Table 8.1. The size of the uncertainties associated to the parameters varies according to the state of development of each project. Naturally, the smallest uncertainties correspond to the running experiments, which have measured their operational parameters. CUORE and GERDA-II have assessed their expected performance with setups operating under conditions similar to those of the final experiment. The remaining experiments base their expectations on results obtained with R&D prototypes, ancillary measurements and Monte Carlo simulations.

Figure 8.1 shows the half-life sensitivity (at 90% CL) of the current generation of $0\nu\beta\beta$-decay experiments, calculated using the parameters of Table 8.1 and the method described in §3.2. The experiments are grouped according to their chosen source isotope. In Figure 8.2 (top panel), we have translated the half-life sensitivities into the more physically-relevant effective neutrino Majorana mass, $m_{\beta\beta}$, using the IBM-2 nuclear matrix elements [72]. For the same exposure, the $^{76}$Ge-based experiments — GERDA-II and MAJORANA—
Figure 8.1. Half-life sensitivity (at 90% CL) of the current-generation experiments in terms of the exposure. The experiments are grouped according to their $\beta\beta$ source isotope. From left to right and top to bottom: $^{76}$Ge (GERDA and Majorana), $^{82}$Se (SuperNEMO), $^{136}$Xe (EXO-200, KamLAND-Zen and NEXT-100) and $^{130}$Te (CUORE and SNO+).
Figure 8.2. Sensitivity to $m_{\beta\beta}$ (at 90% CL) of the current generation of $0\nu\beta\beta$-decay experiments. Top: $m_{\beta\beta}$ sensitivity as a function of the exposure. Bottom: $m_{\beta\beta}$ sensitivity as a function of the experiment’s live time.
Table 8.1. Basic operational parameters of the $0\nu\beta\beta$-decay experiments of the current generation: $\beta\beta$ source isotope, energy resolution (FWHM), $\Delta E$; background rate in the region of interest around $Q_{\beta\beta}$; signal detection efficiency, $\varepsilon$; and source mass.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Isotope</th>
<th>$\Delta E$ (keV)</th>
<th>Bkgd. rate (keV$^{-1}$ kg$^{-1}$ yr$^{-1}$)</th>
<th>$\varepsilon$ (%)</th>
<th>Mass (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUORE-0$^a$</td>
<td>$^{130}\text{Te}$</td>
<td>5</td>
<td>0.23</td>
<td>78</td>
<td>11</td>
</tr>
<tr>
<td>CUORE$^b$</td>
<td>$^{130}\text{Te}$</td>
<td>5</td>
<td>0.04</td>
<td>87</td>
<td>206</td>
</tr>
<tr>
<td>GERDA-I$^a$</td>
<td>$^{76}\text{Ge}$</td>
<td>5</td>
<td>0.013</td>
<td>62</td>
<td>15</td>
</tr>
<tr>
<td>GERDA-II$^b$</td>
<td>$^{76}\text{Ge}$</td>
<td>3</td>
<td>0.001</td>
<td>66</td>
<td>33</td>
</tr>
<tr>
<td>EXO-200$^a$</td>
<td>$^{136}\text{Xe}$</td>
<td>88</td>
<td>0.002</td>
<td>85</td>
<td>76</td>
</tr>
<tr>
<td>KamLAND-Zen$^a$</td>
<td>$^{136}\text{Xe}$</td>
<td>243</td>
<td>0.00014</td>
<td>25</td>
<td>348</td>
</tr>
<tr>
<td>MAJORANA$^c$</td>
<td>$^{76}\text{Ge}$</td>
<td>4</td>
<td>0.0009</td>
<td>70</td>
<td>25</td>
</tr>
<tr>
<td>NEXT-100$^c$ (Chaps.$^a$)</td>
<td>$^{136}\text{Xe}$</td>
<td>18</td>
<td>0.0006</td>
<td>28</td>
<td>91</td>
</tr>
<tr>
<td>SNO+</td>
<td>$^{130}\text{Te}$</td>
<td>264</td>
<td>0.0001</td>
<td>15</td>
<td>800</td>
</tr>
<tr>
<td>SuperNEMO-D$^a$</td>
<td>$^{82}\text{Se}$</td>
<td>120</td>
<td>0.0005</td>
<td>30</td>
<td>7</td>
</tr>
<tr>
<td>SuperNEMO$^c$</td>
<td>$^{82}\text{Se}$</td>
<td>120</td>
<td>0.00005</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

$^a$ The experiment is running and has measured its operational parameters.  
$^b$ The experiment has proven its feasibility with a demonstrator.  
$^c$ The operational parameters are estimations based on R&D results and simulations.

would be the most sensitive detectors of the current generation if they attained their target operational parameters. In its first phase, GERDA has achieved a background rate of about $10^{-2}$ keV$^{-1}$ kg$^{-1}$ [116]. A reduction of a factor of 10 is expected for the second phase of the experiment thanks to the use of broad-end germanium detectors — which present improved pulse shape discrimination and energy resolution — and the instrumentation of the liquid argon bath to veto external backgrounds [140]. The MAJORANA Collaboration pursues a slightly better background rate than that of GERDA-II with a more conventional design for the detector, paying particular attention to the selection of radiopure materials. For instance, the required purity levels for the electroformed copper used in the inner layers of shielding and in the detector holders are extremely stringent, at the level of 0.3 $\mu$Bq/kg of $^{232}$Th or $^{238}$U or below [141].

The experiments of the current generation deploy very different source masses, and hence they cannot reach the same exposures with equal ease. To take this into account, we have represented in the bottom panel of Fig. 8.2 the $m_{\beta\beta}$ sensitivity as a function of the experiments’ live time. In this case, the experiments reaching the best sensitivities are those with the highest effective
8.1. REACH OF THE CURRENT GENERATION OF EXPERIMENTS

Figure 8.3. Effective neutrino Majorana mass ($m_{\beta\beta}$) sensitivity ranges (at 90% CL) defined by the largest and smallest NME calculations for a 3-year run of the $0\nu\beta\beta$-decay experiments of the current generation.

source masses (i.e. mass times detection efficiency): SNO+ and CUORE. These $^{130}$Te-based experiments would achieve a sensitivity to $m_{\beta\beta}$ below 100 meV after a 3-year run, with the $^{76}$Ge and $^{136}$Xe-based experiments reaching $m_{\beta\beta}$ sensitivities about a 30% higher. Nevertheless, the uncertainties deriving from the NMEs ($\S$ 2.5) blur significantly this picture, as shown in Figure 8.3, where we represent the $m_{\beta\beta}$ sensitivity ranges defined for each experiment by the largest and smallest NME calculations after a run of 3 years.

The reliability of the sensitivity estimates given above depends, of course, on how realistic the operational parameters of each experiment are. Besides, in our sensitivity computation, we have neglected systematic uncertainties and any energy-shape information that may be present in the energy distribution of events. Systematic uncertainties may possibly affect the parameters listed above, especially the knowledge of the backgrounds, and deteriorate the sensitivity. On the other hand, use of additional information beyond the overall
count rate of 0νββ-decay candidates within the ROI may yield some sensitivity improvement. While important, both effects would be extremely difficult to incorporate in such a sensitivity comparison, given that most current-generation experiments discussed here have not even started their commissioning phase yet.

In summary, most of the experiments appear to have a a very good chance to reach a sensitivity of 100 meV or better after a few years of effective live time. Given the uncertainties, we cannot predict which among the different experiments will provide the best sensitivity after, say, a 3-year run. To this end, a better knowledge of the actual values for the background rates, of the systematic uncertainties affecting the measurement, and of the NMEs would be necessary for all experiments. The 76 Ge claim should be unambiguously solved by current-generation experiments using different isotopes, but it appears that it will be almost impossible for this generation to discover 0νββ-decay if the neutrino mass spectrum is hierarchical rather than degenerate, as favoured by the cosmological constrains on the lightest neutrino mass.

8.2 Exploring the inverted-hierarchy region of neutrino masses

We turn now our attention to the future prospects of the field. The goal of the next generation of 0νββ-decay experiments is to probe effective Majorana neutrino masses down to 10–20 meV, fully covering the inverted hierarchy region. Such future searches will involve detectors at the tonne or multi-tonne scale in ββ isotope mass. Most collaborations are discussing already possible designs for a tonne-scale version of their experiments. Nevertheless, the diversity of experimental approaches we are currently witnessing will not be viable in the next generation given the cost and difficulty associated; only a few approaches — most likely based on different isotopes — are going to be retained.

Figure 8.4 shows the $m_{\beta\beta}$ sensitivity of perfectly-efficient experiments based on six different ββ isotopes ($^{76}$Ge, $^{82}$Se, $^{100}$Mo, $^{116}$Cd, $^{130}$Te and $^{136}$Xe) and for four different assumptions for the background rate in the ROI. These graphs let us estimate the requirements that next-generation experiments would have to fulfil in order to meet their physics goal. We discuss them briefly in the following:

- GERDA and MAJORANA plan to merge their efforts towards a tonne-scale germanium experiment. In order to reach in a reasonable time
Figure 8.4. Exposure dependence of the $m_{\beta\beta}$ sensitivity (at 90% CL) of perfectly-efficient experiments based on 6 different isotopes and with 4 different assumptions for the background rate in the ROI. The grey band represents the inverted-hierarchy region of neutrino masses.
$m_{\beta\beta}$ sensitivities in the range 10–20 meV, their successor experiment would require at least a tonne of enriched material, i.e. 20–25 times more mass than that deployed at present by the two experiments. Besides, the background rate would have to be reduced by a factor of 10 with respect to the target rate of the current-generation experiments, reaching $10^{-4}$ counts/(keV kg year).

- A tonne-scale version of CUORE would require isotopically-enriched TeO$_2$ bolometers and a background rate two orders of magnitude smaller than the one expected in its present version. The detection of the Cherenkov light produced by signal events in the crystals could be used to suppress the main source of background in CUORE (α activity on the surface of the bolometers) [240][241]. Alternatively, the LUCIFER [242], LUMINEU [243] and AMoRE [244] projects are exploring the use of scintillating bolometers (Zn$^{100}$MoO$_4$, $^{40}$Ca$^{100}$MoO$_4$, $^{116}$CdWO$_4$ or Zn$^{82}$Se), which offer an additional experimental signature for background suppression, but a slightly worse energy resolution. The procurement and enrichment of these crystals would also be notably more expensive than for the $^{130}$Te-based ones. The ultimate sensitivity of bolometers may be limited by the accidental pile-up between multiple $2\nu\beta\beta$ events in the same crystal due to their slow response [245].

- The two large liquid-scintillator calorimeters, SNO+ and KamLAND-Zen, are planning future phases for the exploration of the inverse neutrino hierarchy. KamLAND-Zen intends to dissolve about 1000 kg of enriched xenon in the liquid scintillator. SNO+ can, in principle, increase the concentration of telluric acid in the liquid scintillator by a factor of 10 without affecting the detection and optical properties of the mixture [146]. The energy resolution of both detectors may be improved from the current 10% FWHM to values close to 6% FWHM with the use of a brighter liquid scintillator and an enhancement of the light collection efficiency. As regards the background level, SNO+ would have to suppress by some means its present target rate, $10^{-4}$ counts/(keV kg year), by an order of magnitude. KamLAND-Zen would require an even higher suppression factor to compensate for the lower deployed mass and the differences between the $\beta\beta$ sources. The sensitivity of these experiments will be limited, ultimately, by irreducible backgrounds such as the $2\nu\beta\beta$ spectrum or the solar neutrino flux.
• The EXO Collaboration has started the design of a 5-tonne liquid xenon TPC called nEXO. In order to fully probe the inverted-hierarchy region, nEXO would require a background rate approximately two orders of magnitude better than that achieved in EXO-200. Self-shielding would help suppressing the external backgrounds, even though at the cost of reducing the fiducial mass. In addition, the detector will require neutron shielding to mitigate the activation of $^{136}$Xe.

• The NEXT Collaboration is considering a xenon gas detector with a mass in the range of 1 to 3 tonnes, energy resolution close to 0.5% FWHM at 2.5 MeV, and a background rate of the order of $5 \times 10^{-5}$ counts/(keV kg yr). Besides possible improvements in the radioactivity budget of the detector (more radiopure photosensors, for example), the target background rate can be attained enhancing the discrimination power of the tracking signature with a moderate magnetic field. A single energetic electron should produce a clear single spiral with radius indicative of its momentum, and a double-electron track with the same energy will produce two spirals each with much less momentum and originating from a common vertex. This information provides an additional way of separating single-electrons arising from background processes from double electrons produced in $0\nu\beta\beta$ decays, in spite of the large multiple scattering that the electrons suffer in the dense gaseous xenon.

• SuperNEMO has, by far, the worst mass-to-volume ratio among the techniques considered for $0\nu\beta\beta$-searches: a SuperNEMO module houses 5–10 kg of $\beta\beta$ isotope in approximately 53 m$^3$, whereas, for example, the NEXT-100 pressure vessel, with a volume of about 3 m$^3$, contains 100 kg of enriched xenon in its active volume. Considering this and the limited underground space available, a tonne-scale version of the SuperNEMO experiment seems very unlikely, if not impossible.
At the turn of the 21st century, the observation of neutrino oscillations, which implies that neutrinos are massive particles, and the possible evidence of $0\nu\beta\beta$-decay in the Heidelberg-Moscow experiment boosted the interest in neutrinoless double beta decay searches, prompting a new generation of experiments characterized by source masses in the range of tens to hundreds of kilograms and based on a variety of detection techniques. Indeed, searching for neutrinoless double beta decay is well motivated: first, there is no fundamental reason why total lepton number should be conserved; and second, Majorana neutrinos provide natural explanations for both the smallness of neutrino masses and the baryon number asymmetry observed in the Universe. Consequently, theoretical prejudice in favour of a Majorana nature for neutrinos has gained widespread consensus.

In this work, we have discussed in depth one of the new-generation experiments: the Neutrino Experiment with a Xenon Tpc (NEXT), which will search for the neutrinoless double beta decay of $^{136}$Xe at the Laboratorio Subterráneo de Canfranc (LSC). NEXT possesses two features of great value in $0\nu\beta\beta$-decay searches: excellent energy resolution and an extra experimental signature, charged-particle tracking, for the active suppression of backgrounds. Xenon is a suitable detection medium with strong scintillation and ionization primary signals. In its gaseous phase, xenon can offer an energy resolution as good as 0.3% FWHM at 2.5 MeV. In NEXT, in order to achieve optimal resolution, the ionization signal is amplified using the electroluminescence of xenon: the electrons liberated by ionizing particles passing through the gas are first drifted towards the TPC anode by a moderate electric field (0.3–0.5 kV cm$^{-1}$), entering then into another region where they are accelerated by a stronger electric field (2–3 kV cm$^{-1}$ bar$^{-1}$), intense enough so that the electrons can excite the xenon atoms but not enough to ionize them. This excitation energy is ultimately released in the form of proportional
secondary scintillation light. An array of photomultiplier tubes (PMTs), the so-called energy plane, located behind the TPC cathode detects a fraction of the secondary scintillation (S2) light to provide a precise measurement of the total energy deposited in the gas. These PMTs detect as well the primary scintillation (S1), used to signal the start of the event. The forward-going S2 light is detected by a dense array of silicon photomultipliers (SiPMs), known as the tracking plane, located behind the anode, very close to the EL region, and is used for track reconstruction. We call this detector concept the Separated, Optimized Functions Tpc (§4.3).

The initial phase of the NEXT experiment was dedicated to the demonstration of the detector concept described above using prototypes that contained approximately 1 kg of natural xenon at 10–15 bar (§5.2). An energy resolution of 1% FWHM at 662 keV, which extrapolates to 0.5% FWHM at the Q value of $^{136}$Xe, was measured in the NEXT-DBDM prototype. The best resolution measured in NEXT-DEMO, 1.62% FWHM at 511 keV, extrapolates to 0.74% FWHM. In addition, the DEMO prototype has shown that track reconstruction is possible with an EL-based amplification scheme.

The ultimate goal of the NEXT project is the construction, commissioning and operation of the NEXT-100 detector, a 100-kg xenon gas TPC built with radiopure materials that will start taking low-background data at LSC in 2017. Prior to that, the NEXT Collaboration will operate the NEXT-White (NEW) detector, a technology demonstrator that implements in a 1:2 scale the design chosen for NEXT-100 using the same materials and photosensors. The NEW data will make possible the validation of the NEXT-100 background model, currently based on detailed Monte Carlo detector simulations that predict a background rate for NEXT-100 of at most $6 \times 10^{-4}$ keV$^{-1}$ kg$^{-1}$ yr$^{-1}$. With this background rate, the NEXT-100 detector will reach a sensitivity to the $0\nu\beta\beta$-decay half-life of $T_{0\nu} > 4.6 \times 10^{25}$ years after running for 3 (effective) years. This translates into the $m_{\beta\beta}$ sensitivity range $84–182$ meV.

In the final part of this work, we made an attempt at a quantitative comparison of the physics case of the different $0\nu\beta\beta$-decay experiments of the current generation. There is an intense competition among these experiments to establish themselves as the best approach for neutrinoless double beta decay searches. Since the exploration of the inverted-hierarchy region will unavoidably need to involve experiments at the tonne or multi-tonne scale in isotope mass, the current experiments do not only have to proof their performance for source masses of the order of 100 kg, but also show their scalability to the tonne scale, including the ability to improve their current background rate by,
at least, a factor of 10. In the case of NEXT, we believe that there is ample room for improvement with respect to the baseline detector performance described in this work. With the use of more sophisticated reconstruction algorithms, currently under development, it should be possible to reach an energy resolution close to 0.5% FWHM at 2.5 MeV and fully exploit the potential of the tracking signature. Furthermore, the addition of a magnetic field is a clear path to improve the discrimination power of the technique. With a resolution of 0.5% FWHM and a background rate in the range of $5 \times 10^{-4}$ cts keV$^{-1}$ kg$^{-1}$ yr$^{-1}$, it would be possible to reach sensitivities close to 15 meV, fully exploring the inverse-hierarchy region of neutrino masses.
Bibliography


[38] P. Minkowski, $\mu \rightarrow e\gamma$ at a rate of one out of 1-billion muon decays?, Phys. Lett. B 67 (1977) 421.


