



HAL
open science

Event data model and reconstruction for direct dark matter search with DarkSide-20k

Timothée Hessel

► **To cite this version:**

Timothée Hessel. Event data model and reconstruction for direct dark matter search with DarkSide-20k. Physics [physics]. Université Paris-Cité, 2024. English. NNT : . tel-04886815

HAL Id: tel-04886815

<https://hal.science/tel-04886815v1>

Submitted on 14 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ PARIS CITÉ

École doctorale STEP'UP 560

Laboratoire Astro-Particules et Cosmologie

**Event data model and reconstruction
for direct dark matter search with
DarkSide-20k**

Par TIMOTHÉE HESSEL

Thèse de doctorat en PHYSIQUE DE L'UNIVERS

Dirigée par DAVIDE FRANCO

Présentée et soutenue publiquement le 24 septembre 2024

Devant un jury composé de :

CLAUDIA NONES, PHYSICIENNE HDR	Université Paris-Saclay	Rapporteuse
ANSELMO MEREGAGLIA, DR HDR	Université Bordeaux	Rapporteur
FABRICE HUBAUT, DR HDR	Université Aix-Marseille	Examinateur
JAIME DAWSON, CR	Université Paris-Cité	Examinatrice
DAVIDE FRANCO, CR HDR	Université Paris-Cité	Directeur de thèse

Abstract

Title: Event data model and reconstruction for direct dark matter search with DarkSide-20k.

Keywords: dark matter: direct detection, WIMP, time projection chambers, liquid argon, simulation, reconstruction

Abstract: DarkSide-20k is the next generation dual-phase TPC of the DarkSide programme with 50 ton underground argon target, currently under construction at LNGS (Italy). With data-taking to begin in 2027, DarkSide-20k will achieve cross-section discovery sensitivity of 10^{-47} cm² searching for interactions of WIMPs with 0.1 TeV/c² mass. The sensitivity projection relies on innovative technologies, such as low-noise high-efficiency Silicon Photomultipliers and argon extracted from deep underground, and on the powerful scintillation pulse shape discrimination.

In this thesis, the motivation for the direct dark matter search and

a comprehensive description of the DarkSide-20k detector will be first introduced. This is followed by a description of my contributions to the definition of the event data model, development and optimisation of on-line signal processing, and data reconstruction.

Then, the sensitivities to light dark matter particles are presented for both DarkSide-20k and its predecessor, DarkSide50. In view of improving the comprehension of argon response at the low energies at stakes in these analysis, I describe a potential future measurement. The design, commissioning, and data analysis of a neutron detector configuration for this purpose is also addressed.

Titre: Modèle de données et reconstruction des événements de DarkSide-20k pour la recherche directe de matière noire.

Mots clefs: Matière noire : détection directe; WIMP; chambre à projection temporelle; argon liquide; simulation; reconstruction

Résumé: DarkSide-20k est la chambre à projection temporelle à double phase de nouvelle génération. Ce détecteur, intégré au programme DarkSide, consiste en une cible de 50 tonnes d'argon sous-terrain actuellement en construction au LNGS, en Italie. À l'issue de la prise de donnée qui commencera en 2027, DarkSide-20k atteindra une sensibilité de découverte d'interaction de WIMP et nucléons jusqu'à une section efficace de 10^{-47} cm² à une masse de 0.1 TeV/c². La sensibilité projetée repose sur des technologies innovantes, telles que l'utilisation de photo-multiplicateur au silicium de bas bruit et haute efficacité et l'extraction d'argon en sous-terrain, ainsi que sur la puissante discrimination des formes d'impulsions de scintillation.

Dans cette thèse, les motivations pour la recherche directe de matière noire et une description générale de l'expérience DarkSide-20k seront présentées. Puis seront décrites mes contributions à la définition du modèle de données des événements, au développement et à l'optimisation du traitement de signal en ligne et à la reconstruction des données.

Ensuite, la sensibilité à des candidats de masses légères sera présentée pour DarkSide-20k ainsi que son prédécesseur, DarkSide-50. Dans le but d'améliorer notre compréhension de la réponse de l'argon aux énergies en jeu pour ces analyses, une potentielle future expérience est décrite. Le design, l'installation et l'analyse de données d'une configuration de détecteur à neutron nécessaire pour cette mesure est également présentée.

Acknowledgments

Some of the motivation to do a PhD in physics was to return to science community some of what I received. At University and during internships, I had the privilege to get from passionate people the knowledge that allows me to see the world as I see it today. Although I did to some extent give back during my thesis an inch of what science gave me, I feel today more thankful than ever.

I start this acknowledgments with the four members of my jury who took a huge amount of time to attentively read and question this manuscript. Their implications, the exchanges that we had, and their wide range of expertise have contributed to make this work something I am proud of.

My deepest thank goes to Davide who accepted me as his student and put me in the best situation to carry out this thesis. For your dedication, skills, patience, and most importantly, your humanism at work, you have all my recognition and admiration. I am looking forward to keep working and learning by your side.

I would like to acknowledge the whole DarkSide collaboration for its excellence and enthusiasm. Even if its members are spread around the world, I felt a real sense of belonging, every collaboration meeting providing a precious burst of motivation. I thank everyone I ever interacted with and in particular the young members with whom we shared our common yet unique experiences.

I want to thank my colleagues at APC as well, starting with my office mates and friends Evi and Yuhang for all the good times we have together. I thank Julie for having been an example of a brilliant hard-working student when I arrived, and all my friends of the particle and gravitation groups.

I cannot take a look back over these years without mentioning the music

which paced my life out of the lab. I believe that swapping my attention between music and physics is a healthy stretch that has kept my spirit blossomed for all those years. I thank Enya, Tom, Dario, Adrien, Filipe, Ema and all the current and former members of the Harmoniques Sphériques association for their passion and determination.

A special thank goes to my best friends Artur, my homonym Timothée, and their families who hosted me during the COVID pandemic, turning this dark period into a festive journey full of discoveries, music, and physics. In addition, this experience definitely put me on the rails to the NPAC master and to this PhD and for this, I am extremely grateful.

I thank both my parents for the value of knowledge that they transmitted to me. They always enlightened the path for me to make my own decisions. They have constituted an unconditional support and two examples of greatness. Making them proud is therefore the most precious reward.

I thank my brothers Arthur and Alexandre whose adventures have been an inexhaustible source of inspiration. I may be the doctor among the three of us, but you still are the big brothers I look up to with admiration.

Finally, I would like to thank Laurine for her love, kindness, and trust. You always had the reassuring words and ear I needed to overcome hard times and to make good ones even more beautiful.

Résumé

Depuis les années 1930, des mesures astrophysiques et cosmologiques supportent l'existence de la matière noire dans des échelles de taille s'étendant des galaxies au fond diffus cosmologique. Les premières preuves ont été obtenues en observant les amas de galaxies dont la masse est mesurée en exploitant le théorème du viriel qui relie l'énergie cinétique des galaxies à l'énergie potentielle gravitationnelle. Il apparaît alors que la masse visible, c'est-à-dire la masse des étoiles et du gaz qui constituent les galaxies, ne représente qu'environ un 5^{ème} de la masse totale du système. L'explication la plus communément acceptée est donc l'existence de matière massive et invisible qui contribue au reste de la masse de ces systèmes. De même, les courbes de rotation des galaxies ont servi à démontrer qu'un rapport similaire entre matière visible et « matière noire » existe à l'échelle d'une galaxie. Elles permettent par ailleurs de reconstruire la distribution de la masse dans le halo de matière noire de chaque galaxie. En cosmologie, les prédictions du modèle Λ CDM sont en excellent accord avec les observations, notamment celles des anisotropies du fond diffus cosmologique. Or, ce modèle repose sur l'hypothèse que la matière noire contribue à 85% de la masse de l'Univers. Le problème de la matière noire réside dans le fait que le modèle standard de la physique des particules est incapable de fournir un candidat pour expliquer les observations en astrophysique et en cosmologie.

Les observations permettent de dresser le portrait d'une particule qui aurait les propriétés requises pour constituer la matière noire. Cette particule est massive, électriquement neutre et interagissant faiblement avec la matière ordinaire, elle est appelée WIMP (Weakly Interacting Massive Particle). La détection de WIMPs constitue la direction la plus active pour la recherche de matière noire bien qu'il y ait d'autres types de candidats dont certaines théories justifient l'existence.

La recherche de matière noire se divise en trois axes. Premièrement, les candidats peuvent être produits lors de collisions en accélérateur de partic-

ules comme c'est le cas au LHC. Ce genre de recherche est cependant très dépendant du modèle théorique testé. La recherche indirecte est une autre approche qui consiste en l'observation de produits d'annihilation de matière noire dans des sources astrophysiques. Les données de télescopes au sol ou dans l'espace sont analysées à la recherche d'un excès dont l'origine serait la collision de deux particules de matière noire. Le défi majeur dans ce type de recherche réside dans la compréhension du bruit de fond observé dans ces données. La dernière méthode est le sujet de cette thèse et celle dans laquelle s'inscrit le programme DarkSide. Il s'agit de la recherche directe dont le but est d'observer une collision entre un atome et une particule de matière noire présente dans le halo galactique. L'observation de ce genre d'interaction confirmerait que la matière noire est une particule et permettrait une première estimation de sa masse ainsi que de la section efficace de l'interaction.

Le programme DarkSide traite de la détection directe de matière noire dans l'argon liquide dans une chambre à projection temporelle (TPC) à double phase. Cette technologie consiste en un volume actif liquide dans lequel l'interaction d'une particule incidente induit le recul d'un noyau d'argon ou d'un électron. Ce recul produit l'excitation et l'ionisation d'atomes voisins. Certains électrons se combinent avec les ions d'argon pour former d'autres atomes excités. Lors de la désexcitation, un photon d'une longueur d'onde de 128 nm est émis contribuant à un premier signal de scintillation nommé S1. Les électrons qui échappent à la recombinaison sont dérivés par un champ électrique jusqu'à la phase gazeuse en haut du détecteur. Lorsqu'ils atteignent cette région du détecteur, un champ électrique plus intense produit un signal lumineux d'électro-luminescence nommé S2. Les photons des deux signaux S1 et S2 sont détectés par deux matrices de photo-capteurs placés sur les faces supérieure et inférieure de la chambre. L'impulsion S2 est suffisamment proche de la matrice supérieure pour en déduire la projection dans le plan horizontal de la position de l'interaction primaire. La position verticale de l'interaction est quant à elle déduite du délai temporelle entre S1 et S2.

Les TPC à l'argon liquide exploitent les hauts taux d'ionisation et de scintillation de l'argon ainsi que l'excellente propriété d'identification de particules fournie par l'impulsion de scintillation. En effet, la forme du signal de scintillation est différente si l'interaction a lieu avec un noyau ou avec un électron. Cela est dû aux deux états d'excitation du dimère d'argon, l'état méta-stable responsable de la scintillation. Les états triplet et singulet ont des temps caractéristiques de dé-excitation de $1.5 \mu\text{s}$ et 7 ns respectivement. De plus, la population relative de singulets est plus élevée lors

d'un recul nucléaire dont le signal de scintillation sera par conséquent plus rapide. Or, l'interaction attendue pour une WIMP est un recul nucléaire tandis que la grande majorité des radiations qui constituent le bruit de fond de ces expériences produit un recul électronique. Le bruit de fond de recul électronique peut donc être rejeté avec une très grande efficacité en appliquant une coupure sur la fraction de photons détectés dans une courte fenêtre de l'impulsion. Cette propriété est au coeur des expériences basées sur l'argon liquide car elle permet d'atteindre des taux de bruits de fond instrumental quasi-nul.

DarkSide-20k est le prochain détecteur à grande échelle pour la recherche directe de matière noire avec l'argon liquide en cours de construction au LNGS en Italie. Il s'agit d'une TPC à double phase avec une masse active de 50 tonnes d'argon. La TPC est un prisme droit de 3.5 m de hauteur dont la base est un octogone dont le rayon inscrit à un diamètre de 3.5 m. L'argon utilisé est extrait en sous-terrain afin d'en réduire l'activité en ^{39}Ar , un émetteur β^- naturellement produit dans l'atmosphère par le rayonnement cosmique. La TPC est placée dans un plus grand blindage également rempli d'argon sous-terrain et utilisé comme un veto actif pour rejeter les événements γ en coincidence avec la TPC. La dernière couche du détecteur est le cryostat qui maintient l'argon liquide. Il s'agit d'une structure cubique de 10m de côté. Le volume entre le cryostat et le veto est rempli d'argon atmosphérique et sert de second veto extérieur pour identifier les muons qui déposent une grande quantité d'énergie (plusieurs dizaines de MeV) lorsqu'ils traversent le détecteur. La lumière produite dans les trois volumes actifs est détectée par des photo-multiplicateurs au silicium spécifiquement développés pour DarkSide-20k. C'est la première fois en physique des astroparticules que cette technologie est utilisée à si grande échelle. Le début de la prise de donnée de DarkSide-20k est programmé pour 2027 et la sensibilité projetée de découverte pour la section efficace d'interaction WIMP-nucléon est estimée à 10^{-47} cm pour une masse de 0.1 TeV/ c^2 .

Les données des 2720 canaux d'acquisition de DarkSide-20k seront stockées sans déclenchement global. Le signal produit par un photon sera donc écrit en mémoire indépendamment des autres canaux. Cela induit de nombreux défis liés au traitement en ligne de ce flux ininterrompu de données. Afin de décider au mieux de la stratégie à adopter en ligne, j'ai contribué au développement du logiciel d'émulation de la chaîne d'acquisition du détecteur. Ce travail m'a ensuite permis de développer le modèle de données des événements de DarkSide-20k, c'est-à-dire le format des données brutes qui seront sauvegardées pendant l'expérience. Ce modèle doit satisfaire les exigences liées aux

objectifs physiques de l'expérience telles que la résolution en énergie, la reconstruction de la position ou la discrimination des formes d'impulsions de scintillation. Le modèle de données doit également tenir compte des contraintes liés à la puissance de calcul en ligne et à l'espace de stockage disponible.

Les outils de simulation de DarkSide-20k ont été adaptés pour démontrer que le modèle de données proposé est viable en vue de toutes les contraintes énoncées. En particulier, j'ai participé au développement d'approches effectives pour simuler la propagation de la lumière dans la TPC et dans le veto interne. Cela a permis de réduire de plusieurs ordres de grandeur le temps requis pour la simulation optique et ainsi de produire un large échantillon d'événements de bruit de fond. Ces événements ont ensuite été combinés en tenant compte des activités de chaque isotope dans chaque élément du détecteur. Ce travail représente une étape importante pour l'ensemble de la collaboration car il permet une simulation efficace de tout le détecteur, depuis les processus de radioactivité des matériaux jusqu'à la simulation de l'électronique.

La validation du modèle de données a permis de commencer le développement du code de reconstruction des événements de DarkSide-20k. Puisque le détecteur fonctionne sans déclenchement, la première étape de la reconstruction est l'identification d'impulsion (S1 ou S2) dans les données brutes. Cela est fait par un algorithme de blocs bayésiens. Cette méthode est souvent utilisée en astrophysique mais le code a dû être spécifiquement optimisé pour convenir au flux de données de DarkSide-20k. Une fois ces impulsions identifiées, une première reconstruction est appliquée en vue de la classification entre S1 et S2. Lors de cette étape, un algorithme de recherche de pics est appliqué pour identifier d'éventuels empilements d'impulsions. Après la classification, une reconstruction plus avancée est appliquée, incluant des algorithmes spécifiques à la nature de l'impulsion. La dernière étape de la reconstruction n'est pas encore traitée, il s'agit de l'association des impulsions S1 et S2 en vue de former des événements physiques, c'est-à-dire un ensemble d'impulsions induites par la même particule incidente. Toute cette reconstruction est développée d'une façon modulaire grâce à laquelle chaque étape peut être indépendamment appliquée aux données. Le but de cette modularité est d'optimiser le retraitement des futures données et de faciliter le perfectionnement des algorithmes dédiés.

Le potentiel de DarkSide-20k dans la recherche de candidats légers à la matière noire a été évalué exploitant le développement des outils de simulation. Cette analyse consiste à ne considérer que les impulsions S2 dont

l'efficacité de détection est plus élevée et permet de réduire le seuil en énergie à 0.4 keV en recul nucléaire contre 35 keV dans l'analyse standard. En l'absence d'impulsion S1, il est impossible d'appliquer de discrimination entre recul nucléaire et électronique. Un modèle du bruit de fond de recul électronique est donc nécessaire dans le cadre de cette analyse. Cette méthode avait déjà permis d'atteindre les meilleures limites en section efficace pour des WIMPs de masses inférieures à $4 \text{ GeV}/c^2$ avec les données de DarkSide-50, le prédécesseur de DarkSide-20k. De plus, cette approche permet de sonder des candidats dont l'interaction dans le détecteur se fait par recul électronique, tels que les particules de type axions, les photons sombres ou les neutrinos stériles. L'analyse a donc été transposée à DarkSide-20k afin d'obtenir les sensibilités projetées pour tous ces candidats.

Le succès de l'analyse de masse légère de DarkSide-50 et le potentiel de DarkSide-20k pour cette échelle de masse poussent à approfondir la compréhension de la réponse de l'argon liquide à des reculs nucléaires d'énergies inférieures au keV. Dans ce contexte, j'ai établi une première estimation de la précision qui pourrait être atteinte par une expérience utilisant un flux de neutron d'environ 100 keV. Ce travail repose sur les caractérisations d'un faisceau utilisant la réaction ${}^7\text{Li}(p, n){}^7\text{Be}$ et d'un détecteur de neutron sensible à ces énergies et ayant une résolution temporelle de l'ordre de quelques ns permettant une mesure du temps de vol. Nous avons effectué cette caractérisation pendant deux semaines en décembre 2023 sur la plateforme ALTO à Orsay. La future mesure consistera en une petite TPC dans l'axe du faisceau de neutron. Plusieurs détecteurs seront placés pour détecter les neutrons diffusés dans la TPC. La mesure de l'angle de diffusion permet de reconstruire l'énergie du recul nucléaire. Une telle mesure permettrait de lever certaines incertitudes sur la réponse de l'argon à basse énergie.

Contents

Abstract	3
Acknowledgments	5
Résumé	7
Introduction	17
1 Searching for dark matter	19
1.1 Astrophysical evidences	19
1.1.1 The Coma cluster	19
1.1.2 Galaxy rotation curves	21
1.1.3 Strong gravitational lensing	22
1.1.4 Modified gravity and Bullet cluster	23
1.1.5 Cosmological constraints	24
1.2 Dark Matter candidates	26
1.2.1 Halo model	26
1.2.2 Macroscopic candidates	28
1.2.3 Standard Model of particle physics	29
1.2.4 Axions	31
1.2.5 WIMP miracle	33
1.2.6 Sterile neutrinos	34
1.3 Dark matter searches	35
1.3.1 Indirect search	35
1.3.2 Collider search	37
1.3.3 Direct search	39
1.4 Signatures in direct searches	41
1.4.1 Event rate	42
1.4.2 Flux on Earth and annual modulation	42
1.4.3 Directionality	45

2	The DarkSide project	47
2.1	Noble liquid detectors	47
2.1.1	Scintillation and ionization in noble elements	48
2.1.2	Dual phase time projection chamber	49
2.1.3	Pulse shape discrimination in argon	51
2.2	From DarkSide-50 to DarkSide-20k	54
2.2.1	The DarkSide-50 performance	55
2.2.2	The DarkSide-20k TPC	58
2.2.3	Nuclear recoil and electronic recoil backgrounds	60
2.2.4	Outlier backgrounds	62
2.2.5	Vetoos	64
2.2.6	Photo-detection system	66
3	Electronic recoil background simulation	71
3.1	Tracking and optical simulation	71
3.1.1	The G4DS package	72
3.1.2	Time slice simulation	73
3.2	Effective light simulation	78
3.2.1	Liquid argon response	79
3.2.2	Light maps for S2 pulses	81
3.2.3	Light maps for inner veto pulses	84
4	DAQ emulation and Event Data Model	89
4.1	DAQ emulation	89
4.1.1	DarkSide-20k data acquisition	89
4.1.2	Zero-length encoding	90
4.1.3	DAQ emulation and fast waveform generation	92
4.2	Hit finder	92
4.2.1	Grouping and response templates	93
4.2.2	Electronic noise simulation	93
4.2.3	Matched filtering algorithm	95
4.2.4	Configuration optimisation	97
4.2.5	Laser calibration	100
4.2.6	Noise characterisation	102
4.3	Event data model	106
4.3.1	Physics requirements	106
4.3.2	Summed waveforms	107
4.3.3	DAQ output data structure	109
4.3.4	Digitizer transfer rate and data storage	111
4.4	Physics validation	113
4.4.1	PE detection efficiency	113

4.4.2	Energy reconstruction	114
4.4.3	Pulse shape discrimination	115
4.4.4	Position reconstruction	116
5	Data reconstruction	119
5.1	Clustering	119
5.1.1	Bayesian Blocks	120
5.1.2	Fast Bayesian Blocks algorithm	122
5.2	Pulse finder and classification	124
5.2.1	Pulse finder	125
5.2.2	Peak finder	127
5.2.3	Low-level pulse reconstruction and classification	131
5.3	High-level pulse reconstruction	134
5.3.1	S1 position reconstruction	135
5.3.2	z from S2 pulse shape	135
5.4	Modular framework	138
5.4.1	Reprocessing constraints	138
5.4.2	Reconstruction level formats	139
6	Sensitivity to light dark matter particles	141
6.1	Ionization response	142
6.1.1	Electronic recoil response	142
6.1.2	Nuclear recoil response	143
6.2	Light dark matter signal models	144
6.2.1	Low-mass WIMPs	145
6.2.2	Leptophilic signal	146
6.2.3	Migdal effect	147
6.3	DarkSide-50 low-mass analysis	148
6.3.1	Data selection	148
6.3.2	Background fit and systematics	149
6.3.3	Limits	152
6.4	DarkSide-20k low-mass projected sensitivity	153
6.4.1	Expected background	154
6.4.2	Single electron resolution	157
6.4.3	Live time estimation	158
6.4.4	Expected limits	159
7	Argon response to low energy nuclear recoil	163
7.1	Setup for beam and detector characterisations	165
7.1.1	Low energy neutron beam	165
7.1.2	^{10}B -based neutron detector	166

7.1.3	^6Li -Glass neutron detector	167
7.1.4	Position and angle reconstruction	168
7.2	Characterisation results	169
7.2.1	Time-of-flight resolution	169
7.2.2	Beam characterisation	171
7.3	Nuclear recoil low energy sensitivity	172
7.3.1	Experimental setup	173
7.3.2	NR energy uncertainty	174
7.3.3	Quantum yield uncertainty	175
Conclusion		179
References		181
List of Figures		189
List of Tables		199

Introduction

The existence of dark matter is supported by many evidences from astrophysical and cosmological objects whose scales range from individual galaxies to the largest structures of the Universe. The most favoured description of dark matter, from an epistemic point of view, is in the form of an elementary particle accounting for about 85% of the mass of the Universe. As the Standard Model of particle physics is unable to provide a viable candidate, the nature of dark matter remains unknown. This thesis covers the search for Weakly Interacting Massive Particles (WIMPs), one of the preferred candidates to dark matter, with the DarkSide-20k experiment, a dual-phase liquid argon Time Projection Chamber.

Nowadays, the search for very rare interactions of dark matter with ordinary matter is one of the highest priority topic of particle physics. Noble liquid detectors are leading the sensitivity to dark matter interaction with nucleon thanks to their radio-purity, scalability, and high quantum yields. In particular, the particle identification properties of liquid argon were exploited in the DarkSide-50 experiment to provide a first instrumentally background-free search for WIMP-induced nuclear recoil interactions. DarkSide-20k is the next generation detector that will be operational in 2027. The target mass of the detector will be increased by 3 orders of magnitude with respect to its predecessor and a novel SiPM-based photo-detection system is currently under production. The many challenges related to the design of this new detector are set to reach an unprecedented sensitivity and a strong discovery potential. The first two chapters of this thesis describe the phenomenology of dark matter, the experimental techniques to detect it, and the DarkSide-50 and DarkSide-20k detectors.

When I joined the DarkSide collaboration in 2021, some questions on the design of DarkSide-20k were pending. The simulation tools needed to tackle these technical aspects were still under development. It is in this context that I had the opportunity to participate to the development of the whole

simulation framework of DarkSide-20k and its application to the studies presented in this thesis. The Monte Carlo simulation and the custom response model of argon are described in chapter 3.

The online processing and storage of the raw data from ~ 2500 read-out channels is among the most crucial aspects of the experiment. I developed the Event Data Model, namely the definition of the raw data format of DarkSide-20k, which takes into consideration all the computing constraints and the trigger-less strategy adopted for the data acquisition system (DAQ). In this study, I verified that the strong data reduction still ensures the complete reconstruction of all physics observables needed for dark matter search. This work, along with the development of the DAQ simulation code and the algorithm treating online the signal from the SiPMs, is presented in chapter 4.

The full simulation chain and the Event Data Model allow to start the development of the reconstruction software of DarkSide-20k. Its role is to provide, from the raw data, the physical events and all the information required for the analysis, as presented in chapter 5. The impact of the reconstruction is also an input for the sensitivity study of DarkSide-20k.

A recent analysis of DarkSide-50 data demonstrated strong sensitivity to light dark matter candidates, namely low-mass WIMPs, axion-like particles, leptophilic dark matter. This search for low energy signal motivated the study of the projected sensitivity of DarkSide-20k to this class of candidates. In chapter 6, I detail the DarkSide-50 analysis and the study of the projected sensitivity of DarkSide-20k.

The light dark matter search relies on a model of the response of liquid argon to low energy nuclear and electronic recoils. I discuss in the last chapter a future test experiment aiming at improving the precision of this model down to the sub-keV range. A first series of test were carried out in December 2023, in order to assess the beam and neutron detectors that will be involved. From this, I evaluated the sensitivity that will be achieved during this measurement.

Chapter 1

Searching for dark matter

Dark matter is one of the most intriguing question of fundamental physics and the DarkSide-20k experiment will probe for interaction with ordinary matter with unprecedented sensitivity. Here, we introduce the evidences for dark matter provided by astrophysics and cosmology. The different approaches to search for dark matter are then detailed with some examples from major experimental results.

1.1 Astrophysical evidences

Observations at many different scales from galaxies to large structures suggest unequivocally that the composition of matter in the universe is dominated by something unknown and invisible. Beyond astrophysics, dark matter is a key ingredient of contemporary cosmology and became an extreme experimental challenge as no detection other than gravitational has ever been provided.

1.1.1 The Coma cluster

The first evidence of dark matter dates back to the 1930s. Fritz Zwicky studied the Coma cluster while he was looking for hints of antimatter annihilation. The Coma cluster is a group of more than 9000 galaxies 100 Mpc away from Earth. In 1933, the dynamics of the cluster allowed him to infer some invisible matter located between the galaxies. It is not the first measurement to suggest that most of the matter is something unknown, but it is the first time that the existence of dark matter is suggested.

Assuming the cluster to be bound and stable, a relation can be established

between the relative velocities of the galaxies that compose it and its gravitational potential. The projections of the velocities in the radial direction were derived by measuring the redshifts of the galaxies. Considering only the masses of the stars and gas, the relative speed of the galaxies is too high for the system to be bound. Without any additional mass, those velocities would be enough for the galaxies to escape the system [1]. This is quantified by the virial theorem which relates the kinetic and potential energies of a bound system as follow:

$$\frac{1}{2} \langle v^2 \rangle = \frac{3}{10} \frac{GM}{R} \quad (1.1)$$

Here, $\langle v^2 \rangle$ is the measured mean squared dispersion velocity, G is the gravitational constant, R is the radius of the cluster, and M its total mass. The virial theorem is a consequence of the law of equipartition applied to a system with a gravitational potential and a temperature defined as the kinetic energy of galaxies. The left-hand term is the kinetic energy and the right-hand term is half of the gravitational potential energy, both expressed in units of mass.

The mass obtained from the measured dispersion velocity around 1000 km/s is much higher than the one predicted only from the luminosity of the galaxies. If the mass of the cluster was only carried by the stars, this value would be 80 km/s. The discrepancy is solved by suggesting some additional invisible matter to increase by a factor 500 the total mass of the system.

The Zwicky's measurement however, was highly overestimated because of the limited knowledge of the Universe expansion at that time. In particular, the Hubble constant describing the state of the expansion at our epoch has a strong impact on the derivation of the velocities from the redshift. With more recent data, the same measurement gives a mass-to-light ratio around 8 which is in agreement with the results from other astrophysical systems.

By the time, the solution proposed by Zwicky was not very well received and the idea of missing mass only resurfaced in the 1970s. Still, it is the first time the virial theorem was applied to measure the mass of a galaxy cluster [2].

1.1.2 Galaxy rotation curves

Historically, the evidence that contributed the most to convince the scientific community is the flatness of galaxy rotation curves. The dynamics of isolated spiral galaxies suggest the existence of missing mass in a similar manner than the Coma cluster. Once again, assuming the galaxy to be stable, the gravitational force applied by the galaxy on a single star has to be exactly compensated by the centrifugal force. The latter can be measured by building the rotation curve of the galaxy, namely the rotation speed as a function of the distance to the galaxy centre r . This rotation curves are obtained by spectrometry, exploiting the continuous distribution of stars and gas in the galaxy. Then, the following equation can be established by equaling centrifugal and gravitational forces:

$$\frac{m_* v(r)^2}{r} = \frac{GM(r)m_*}{r^2} \Rightarrow v(r)^2 = \frac{GM(r)}{r} \quad (1.2)$$

Here, $M(r)$ is the enclosed mass at the radius r . m_* is the canceling out mass of the considered star. Observations show that beyond the galactic disks, $v(r)$ remains close to constant which implies $M(r) \propto r$ (figure 1.1). Since $M(r)$ is the integral of the mass distribution, it is proportional to r if the mass distribution is constant. That is the evidence of some mass almost uniformly distributed in a halo much bigger than the galactic disk.

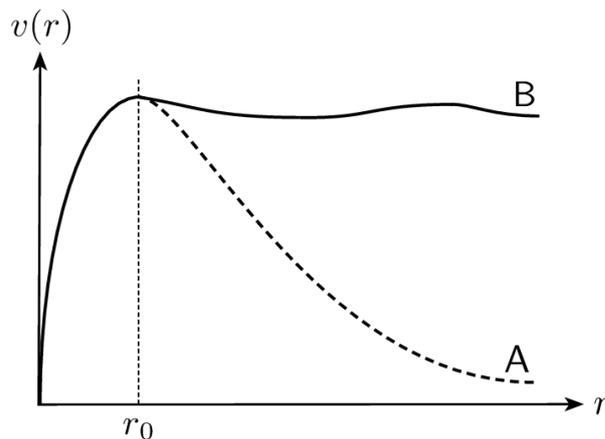


Figure 1.1: Example of rotation curves. Dashed curve A describes a situation in which all matter is contained in the disk of radius r_0 . Curve B is what is typically measured.

Preliminary measurements were firstly made in 1930 by Knut Lundmark who measured the mass-to-light ratios of 5 galaxies [3]. The high values he found were not related to some missing mass by that time. It is in 1970 that Vera Rubin worked on high quality observations of the Andromeda Galaxy, highlighting the flatness of the rotation curves, which allowed to invoke additional mass in galaxy disks [4]. In the following decade, the number of measurements increased and in 1989 the rotation curve of the Milky Way was revealed [5]. From then, the compatible results in different galaxies, in clusters, and even in our own galaxy made the dark matter hypothesis the most accepted solution in the astrophysics sector.

1.1.3 Strong gravitational lensing

Dense objects curve space-time inducing light rays to bend when passing nearby. The curvature of the light depends on the mass of the object. If the light between the observer and a luminous source passes close to a very massive object, it is deflected by an angle depending on the mass of the object. For an extended object, the formation of arcs can appear while for point-like sources, it is possible that the light reaches the observer from multiple directions in the sky. Two examples of this effect are shown on figure 1.2.

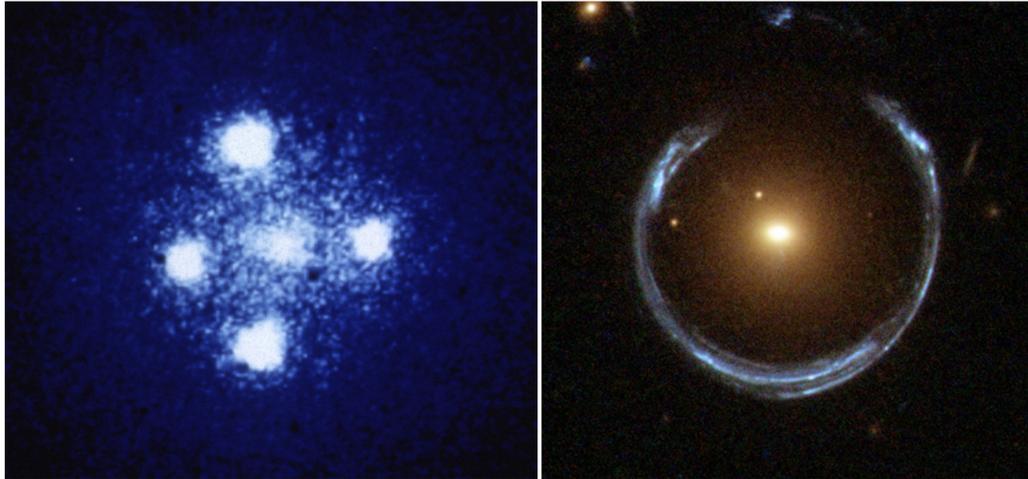


Figure 1.2: Strong lensing from point-like (left) and extended (right) sources seen by the Hubble telescope. The left image shows a quasar seen 4 times as it is distorted by the central galaxy. The right image shows the arc image of a galaxy formed through gravitational lensing.

This effect is the strong gravitational lensing and it provides a precise and independent measurement of the mass of the intermediate object. Starting from 2004, this method was applied on many systems, giving further confirmation on the discrepancy between “gravitational” and “luminous” masses, *id est*, estimated from gravitational considerations and from the empirical luminous to matter ratio law [6].

1.1.4 Modified gravity and Bullet cluster

As the evidences for a discrepancy in the mass of astrophysical objects accumulate, an alternative to dark matter was proposed in 1982, initially focused on explaining Vera Rubin’s rotation curves. By implementing a modification of the Newton’s law of gravity, Mordehai Milgrom described the dynamics of galaxies without requiring additional mass [7]. By scaling the force not with the acceleration but with the square of it, this Modified Newtonian Dynamics (MoND) approach was first proposed as an effective solution for the dark matter problem without any strong theoretical framework.

The extension of MoND to theoretically supported modified gravity models is today a very active field and the scope is not only to modify Newton’s equation of motion but also general relativity. It is indeed quite natural when a tension is raised between observation and theory, here the mass-to-light ratio and the dynamical tools to describe it, not only to question the content of the studied system but also the theory itself. Then, the solution could appear by fixing our understanding of gravitation interaction.

A very constraining observation made in 2006 is the X-Ray and gravitational lens map of the Bullet cluster. In a paper named "A Direct Empirical Proof of the Existence of Dark Matter", Douglas Clowe and his colleagues reveal a map obtained from data of the Giant Magellan Telescope of the collision remnants of two galaxy clusters [8]. The collision induced an increase of the gas temperature, leading to X-ray de-excitation emissions correlated to the mass distribution. These emissions allowed to reconstruct the distribution of ordinary matter while a gravitational mass distribution was obtained by strong gravitational lensing effect. The difference between the two distributions is shown in figure 1.3.

The disparity between the two contents reveal that during the collision of the two galaxy clusters, the baryonic mass was slowed and scattered because of interactions while most of the mass was not affected. The authors refer to

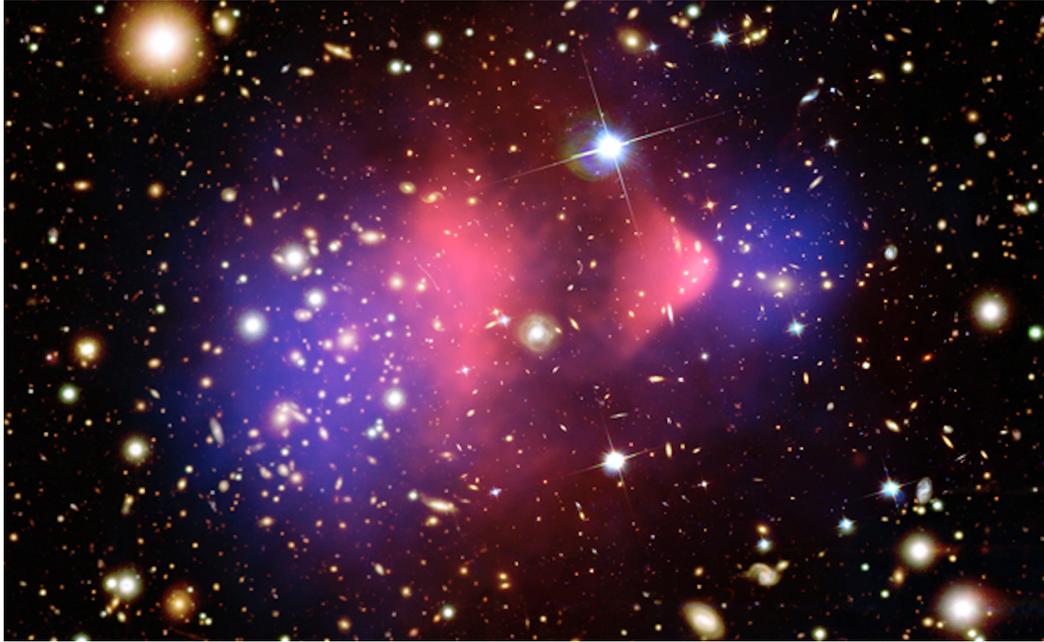


Figure 1.3: Composite image of the galaxy cluster 1E 0657-56, known as Bullet cluster. The pink area is the X-Ray distribution corresponding to baryonic (ordinary) matter. The purple area is the mass distribution reconstructed with gravitational lensing.

that as a proof of the existence of a dark matter. However, it is the gravitational lensing effect that is an evidence of dark matter and is observed in many other systems. The Bullet cluster simply informs on the collision-less property of dark matter and consists in a hard obstacle for modified gravity.

This dynamical effect cannot be directly explained by modified gravity theories. It is not yet excluding them since the Bullet cluster dynamic could be described by setting a lower limit of $2 \text{ eV}/c^2$ on the mass of neutrinos. Thus the mass distribution can be carried by the neutrinos that are indeed collision-less. However, recent cosmological measurement already reject total neutrino masses above $0.3 \text{ eV}/c^2$ maintaining the tension between modified gravity and the Bullet cluster [9].

1.1.5 Cosmological constraints

Cosmology has also provided during the last decades some crucial constraints about dark matter phenomenology. The best model of cosmology nowadays is the Λ CDM model in which dark matter is described as a non-relativistic

neutral long-lived or stable particle. Information on the origin and faith of the universe that are gathered in agreement with the Λ CDM model can therefore be related to the properties of dark matter.

Λ CDM stands for Λ the cosmological constant, the cause of dark energy, and Cold Dark Matter. It describes the mass-energy content of the universe being largely dominated by dark energy (72%), the remaining are 23% of dark matter and around 5% of baryonic matter. The constraining behaviour of the universe on those values is its expansion observed to be accelerating. If the universe content was dominated by matter or by radiation, as it was the case in some earlier stages, the expansion would be decelerating hence the need of introducing another unknown source of energy, the dark energy, described by a different equation of state [10].

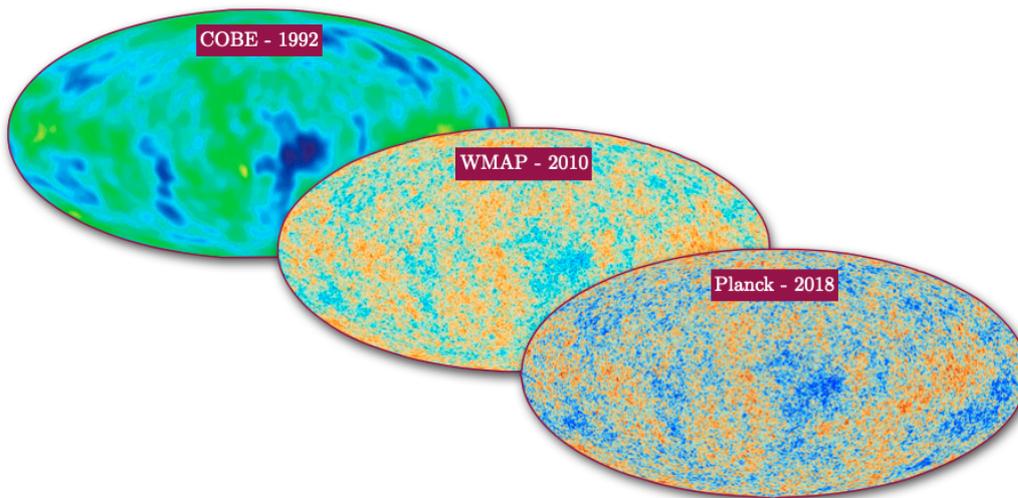


Figure 1.4: Maps of the anisotropies in the CMB

A result of cosmology is the bottom-up mechanism at the origin of large scale structures formations. This is inferred from the cosmic microwave background (CMB), the most ancient image we have of the universe (figure 1.4). The further we observe, the longer the light took to reach us, and the further object we can see is the CMB. Up to 380,000 years after the Big Bang, the mean free path of photons in the universe was too small, making the later opaque. After that time, the universe cooled down and expanded. It eventually became transparent and photons could reach us nowadays. This event is called the recombination. CMB is made of the last photons of the opaque universe and provides priceless information for cosmology.

Observations of the CMB, and especially its anisotropy, allowed to understand that small structures had to appear before larger objects could form, this is the bottom-up mechanism. A consequence of that is a strong preference for cold dark matter, where dark matter would be composed of high mass particles with low velocity. If dark matter was hot, made of low mass particles with high velocity, it would escape the gravitational potential of large scale structure, preventing their formation. The extreme compatibility of the CMB anisotropies with the Λ CDM predictions makes it the strongest evidence of dark matter.

These observations also indicate that dark matter and ordinary matter were in thermal equilibrium at the recombination. When density and cross-section are high enough, two particle species will exchange energy and reach the same temperature through elastic collisions, this is the thermal equilibrium. For instance, photons and ordinary matter were at thermal equilibrium until recombination. The broken equilibrium left a relic photon density that is today of 411 photons/cm^3 with a temperature of 2.7 K and this is the CMB. The thermal equilibrium of dark matter with ordinary matter is the evidence of some interaction between them.

1.2 Dark Matter candidates

The previously described observations give many constraints on the general properties of dark matter. The absence of light produced or absorbed by the dark matter halo implies that it does not interact through electromagnetism, therefore it cannot be electrically charged. The dynamics of systems such as the bullet cluster suggest that dark matter is not or very weakly interacting, besides gravitational interaction, while the CMB anisotropies exclude the possibility of strictly non-interacting dark matter. It has to be stable and obviously massive. These four conditions constitute the common particle dark matter paradigm which is the most actively studied and in which fits the DarkSide program.

1.2.1 Halo model

The local measurements of dark matter are of the most crucial since they guide direct search experiments such as DarkSide. Direct search experiments

aim at detecting the dark matter located around the Earth and eventually crossing it. Therefore, the information given by local measurements is an essential piece to interpret and compare their results with the expected flow.

There are two main methods to determine the local dark matter density (ρ_{DM}), the first is by looking at stars close to the Sun. Similarly to the Coma cluster study, the mass inferred by gravitational dynamics can be compared to the visible mass density. Another possibility is to build a dark matter density for the whole Milky Way by measuring its rotation curve. The latter is obviously made harder by the fact that the Milky Way can only be observed from inside. These two kinds of study started to give compatible results in the 2010s [11].

While the Milky way disk has a radius of 15 kpc with the solar system located 8 kpc away from the centre, the dark matter halo extends up to 100 kpc. The dark matter halo of the Milky Way can be described according to several models. In the Standard Halo Model (SHM), it is assume spherical and following a Maxwell-Boltzmann distribution which are the most basic hypotheses to consider. Latest measurement gives the following parameters for the SHM:

$$\left\{ \begin{array}{l} v_{escape} = 544 \text{ km/s} \\ v_0 = 220 \text{ km/s} \\ v_{Earth} = 232 \text{ km/s} \\ \rho_{DM} = 0.3 \text{ GeV}/c^2/\text{cm}^3 \end{array} \right. \quad (1.3)$$

Here v_{escape} is the escape velocity beyond which a particle would have enough mechanical energy to escape the gravitational potential of the halo. It settles, on the velocity distribution, a strict limit that is a priori not considered in the Maxwell-Boltzmann distribution. v_0 is the mean speed and v_{Earth} is the mean speed of the Earth in the Milky Way frame. Since the halo properties are rather unconstrained, there are many well-motivated alternative models with for instance oblate halos. Different halo models lead to different local parametrisation and the signatures expected in direct searches would also be affected [12]. However, at this stage where no dark matter has been observed, the impact of varying the dark matter halo model is very small.

1.2.2 Macroscopic candidates

It was assumed so far that dark matter is made of microscopic particles, relaxing this hypothesis would affect the previously discussed properties. If dark matter is made of heavy macroscopic elements, a same mass density is achieved with a lower number density. Thus the lack of observed interaction can be explained by the rarefaction of those elements. Many models are developed introducing this kind of objects, denoted as MACHOs (MASSive Compact Halo Objects).

One common example of MACHO is the primordial black hole. Primordial black holes are stable objects that could have formed in the early universe. The masses are spread in very wide ranges between 0.01 to 300 solar masses. A method to probe for this kind of objects is by looking for microlensing events. Similarly to strong lensing, small dense objects should bend nearby light when passing in front of the source, inducing a small and short increase of the luminosity while the object aligns with the source and the Earth. A scheme of this effect is presented in figure 1.5. By looking for these increases of intensity in far away sources, one can constrain the abundance of primordial black holes. These observations also consider different MACHOs candidates such as neutron stars or brown dwarfs.

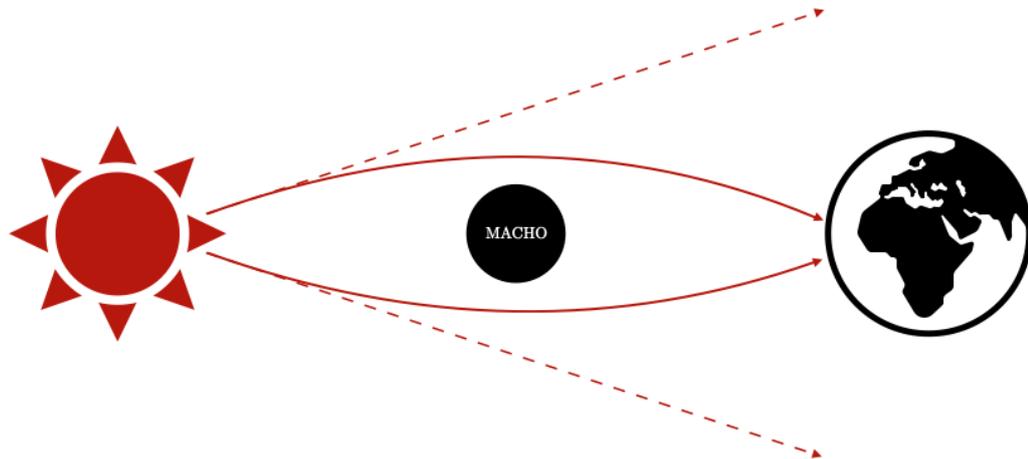


Figure 1.5: By microlensing effect, light rays emitted are converged by a massive object between the source and the observer.

No signal was observed through microlensing effect so far, excluding a wide region in the MACHO mass-halo-fraction phase space [13]. Although

searches for MACHOs excluded the possibility that they compose all the dark matter, it is still possible that they contribute to a fraction of it as shown in figure 1.6. Recently, a large amount of black holes with masses of the order of $40 M_{\odot}$ was observed in gravitational waves detector, reviving the motivation toward black holes dark matter. However, with this mass, the density should be extremely high in order to contribute significantly to the dark matter halo. With the current observations, the halo fraction upper limit of $40 M_{\odot}$ primordial black holes was set at 1% [14]. As it remains the most commonly accepted hypothesis, dark matter will be assumed to be of particle nature in the following.

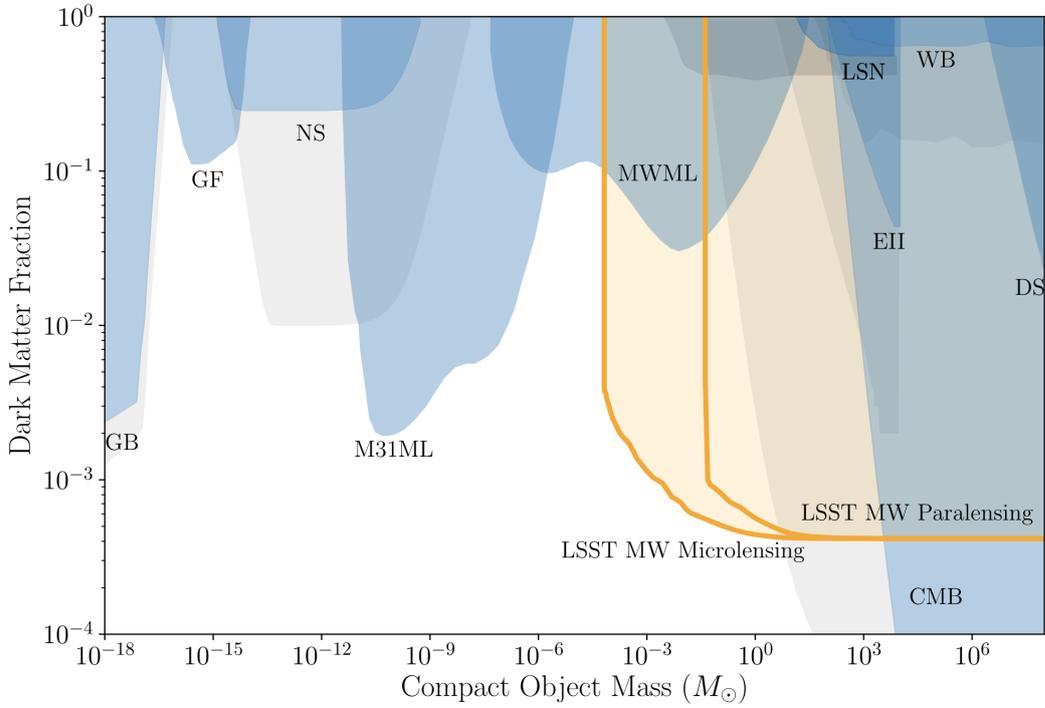


Figure 1.6: Current 95% confidence level limits on MACHO contribution to halo as a function of its mass along with projected sensitivity of the Large Synoptic Survey Telescope [15].

1.2.3 Standard Model of particle physics

The Standard Model of particle physics provides the best description of the elementary constituents of our universe and their interactions. It is a quantum field theory developed in the 1970s which had many great successes, the

most major one being the discovery of the predicted Higgs boson in 2012 [16]. It is natural when looking for dark matter to investigate the capability of every known particle to describe it.

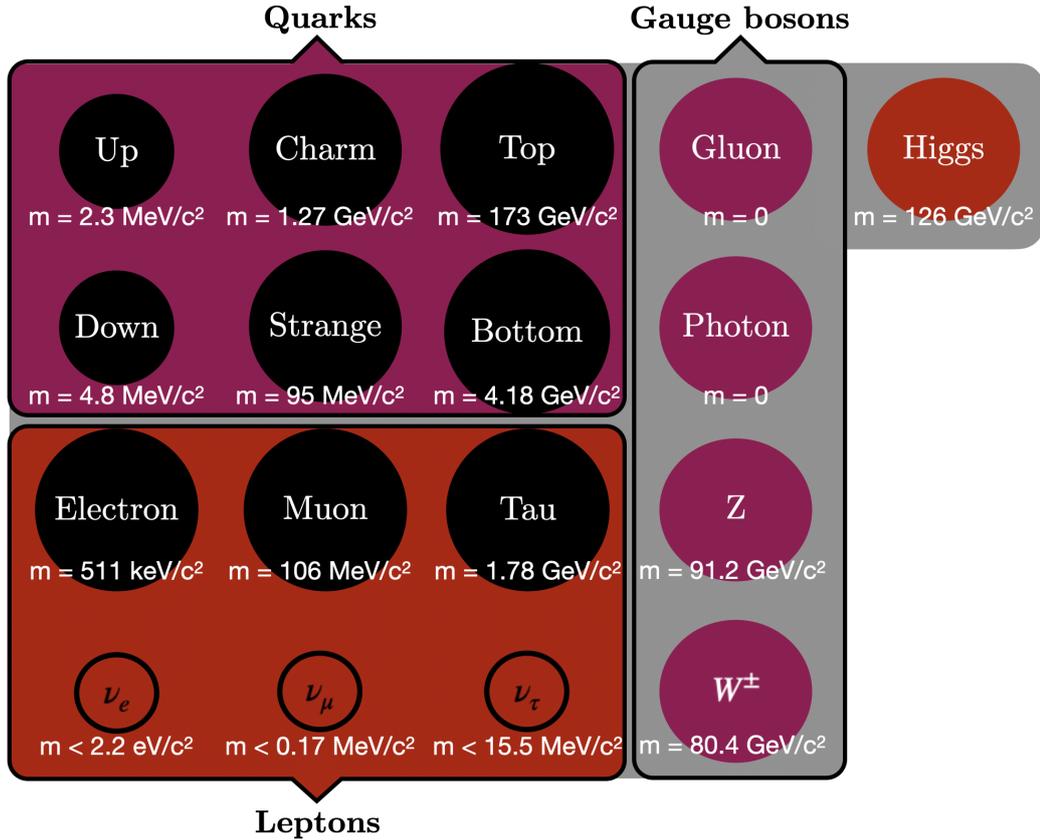


Figure 1.7: The Standard Model of particle physics. Indicated neutrino masses are those of mass eigenstates which differ from the represented flavour eigenstates.

The Standard Model includes 12 fermions, 4 interaction vector bosons plus the Higgs scalar boson (figure 1.7). The fermions are split in two categories, the leptons and the quarks. There are 6 quarks, with the up and down quarks being the lightest and the only stable ones. They are electrically charged, interacts through the strong interaction, and couples with the gluon boson to form composite particles such as proton and neutron.

On the other side are the charged and neutral leptons. The electron is the lightest and the only stable charged lepton. The neutral leptons are the three

neutrinos, each sharing a common leptonic quantum number with a charged lepton. Since they are not electrically charged, they do not interact through electromagnetism, in other words, they do not couple with the photons. The only coupling they have is with the bosons of the weak interaction, the Z and W^\pm .

The neutrino, as described in the Standard Model, is massless. However, it was demonstrated that neutrinos oscillate between their three flavours when propagating in space [17]. Upon being detected, the neutrino flavour can differ from the one it had when emitted, with a probability depending on the travelled distance. This means that the neutrino does not strictly travel at the speed of light. Otherwise, a change of state in the rest frame of the neutrino would not be possible in a finite time observed in the lab frame. Therefore, flavour eigenstates differ from the mass eigenstates. The oscillation parameters are related to the mass differences between the three neutrino mass states [18].

The only could-be candidates for dark matter within the Standard Model are the stable particles, namely the electron, the neutrinos and quark composite particles. Electrons are charged so would interact with light. In the quark sector, the neutron is the only stable neutral candidate, but it would interact through the strong force to bind with protons and form atoms that would in turn interact with light. Light particle dark matter being excluded by the Λ CDM model, neutrinos are not good candidates either, leaving no possibility for dark matter but to exist beyond Standard Model. It is to bear in mind that there exist alternative models for cosmology with warm dark matter in which neutrinos could fit, although they are less favoured by observations.

1.2.4 Axions

One very popular candidate for dark matter appears from the strong sector. The best description of quark and gluon physics is provided by the quantum chromodynamics (QCD). Within QCD, charge-parity (CP) symmetry of the quarks states, namely the invariance through charge conjugation (C) and parity inversion (P), should not be conserved. It is broken by the weak interaction and has no reason to behave differently with the strong interaction. This symmetry breaking is described in the Standard Model Lagrangian, the quantity that governs the dynamics and interaction of every particle field. Among the QCD terms of the Lagrangian, the one responsible for the cou-

pling between quarks and gluons is not invariant under CP transformation [19]:

$$\mathcal{L}_{QCD} \supset \theta \frac{g_s^2}{32\pi} G^{\alpha\mu\nu} \tilde{G}_{\mu\nu}^\alpha \quad (1.4)$$

Here, g_s is the strong force coupling constant. $G^{\alpha\mu\nu} \tilde{G}_{\mu\nu}^\alpha$ is the field strength tensor characterising gluon interaction with quarks, this operator violates P but not C, therefore it violates CP. It is proportional to a parameter θ which should be non-zero if CP is violated. However this symmetry breaking has not been observed yet. Experimental measurements, especially bounds on the neutron electric dipole moment, set strong upper limits ($\theta < 10^{-9}$). The very small value of θ is known as the strong CP problem [20].

A common proposed solution to the CP problem is to introduce a spontaneously broken $U(1)$ symmetry. It is a symmetry with one degree of freedom which is preserved by the equations of motion except around the vacuum solution. For perturbations around the vacuum state, this new symmetry is broken. In that case, θ describes a scalar field and its low value is a direct consequence of the dynamics.

CP violation in the strong sector also appears in the quark mass term and can also be parametrised by θ intervening as a phase in the quark mass matrix:

$$\mathcal{L}_{QCD} \supset - \sum_q m_q \bar{q} e^{i\theta_q} q \quad (1.5)$$

Here, q denotes the quark flavours which are not mass eigenstates if θ_q is non-zero. The corresponding particle of the new field is called axion and the expansion of the Lagrangian at the minimum of the effective potential introduces a mass term. Axion weakly couples with the other Standard Model fields in diverse way, thus it is massive, neutral and weakly interacting, therefore a good candidate for dark matter. Proposed in 1977, the axion is a very well motivated candidate because it is a natural solution in QCD which happens to provide an answer to the problem of dark matter. Beside the original axion, the searches extend to a whole phase space of axion-like particles, not motivated by QCD but still potential candidates of dark matter.

1.2.5 WIMP miracle

In cosmology, the origin of the dark matter population observed today is described by the interruption of the thermal equilibrium with baryonic matter in the early universe. The breaking of the thermal equilibrium leaving relic density and temperature is called freeze-out.

An alternative scenario is the interruption of the species production. It is possible in the case where the species abundance is initially zero or very small. If its abundance can be increased by annihilation of an abundant other species, then it will be produced as long as the annihilation rate is high enough. The cooling and expansion will lower the rate and eventually stop the production. This is called a freeze-in and also leaves a relic abundance [21]. In that context, the decoupling happens much earlier and the final abundance value depends on both the cross-section and the initial abundance (figure 1.8).

In the case of dark matter and baryonic matter, the decoupling is expected to occur much sooner than with photons because of the low cross-section between the two species. Assuming a mass in the weak scale and an interaction rate in an order compatible with the weak interaction, the freeze-out leaves a relic abundance matching the order of magnitude of the observed one. This feature is known as the WIMP (Weakly Interacting Massive Particle) miracle. WIMP is for this reason the favorite kind of dark matter candidates nowadays with mass range from the GeV and the tens of TeV. The WIMP miracle can also be described by a freeze-in mechanism.

WIMP candidates are supported by many beyond Standard Model theories. Among them is the supersymmetry, well-motivated by particle physics, which extends the Standard Model by including a new spin symmetry. The consequence is that for any Standard Model particle exists one or more supersymmetric partner. If the original particle is a fermion, the partner will be a boson and vice versa. This partner are not necessary mass eigenstate, but the neutralino is a predicted stable massive neutral fermion state and therefore a good candidate for dark matter [22]. Moreover, the expected neutralino-nucleon cross-section region which depends on many parameters will be partially probed by the next generation of direct dark matter search experiment. Although it has not been observed in accelerators, supersymmetry is still a very attractive theory in particular from the dark matter point of view.

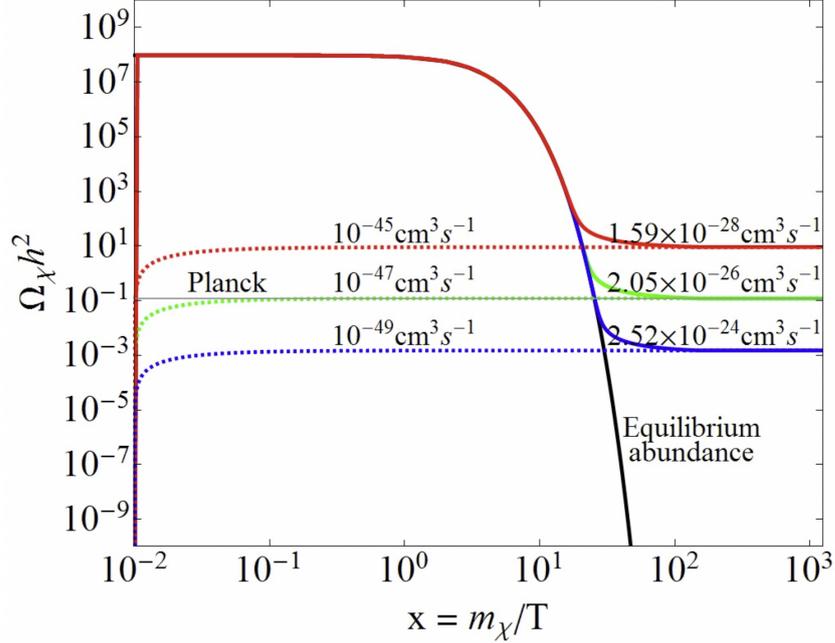


Figure 1.8: Abundance evolution of dark matter in case of freeze-out (continuous lines) and in case of freeze-in (dotted lines) for a mass $m_\chi = 100$ GeV/ c^2 . The horizontal line is the relic density observed by the Planck telescope [21].

1.2.6 Sterile neutrinos

Neutrinos interact only via the weak interaction which only couples left-handed particle or right-handed anti-particles. These are therefore the two only parity/helicity states allowed for neutrino to interact. The LEP experiment gave a constrain on the number of neutrino families that is in agreement with the 3 lepton flavours in the Standard Model [23]. However this number concerns neutrinos subject to the weak interaction so there could be a fourth right-handed state which does not couples with the weak gauge bosons W^\pm and Z , it is for this reason called a sterile neutrino and is a candidate to dark matter. For sterile neutrinos to exist, they need to have been produced in the early universe by mixing with the interacting states [24]. Neutrino flavour states mixing with mass states is done through the Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix).

Sterile neutrinos need masses above the keV to be proper cold dark matter candidates. At lower masses, they could also be part of warm dark matter model [25]. One of the main motivation for the sterile neutrinos is the seesaw

mechanism. The idea is that the active neutrinos are massless in the leading order but acquire very small masses by loop effect from the sterile state. The seesaw mechanism provides an explanation for the very low neutrino masses and gives a natural candidate to dark matter.

The search for sterile neutrinos is centred around neutrino oscillation experiments. Extra disappearance of electronic neutrinos can be a hint of an additional flavour state and some short baseline experiments have recently shown some anomalies in that direction [26]. The anomaly is not solved yet and is under close investigation by many efforts, with on-going and incoming accelerator, reactor, and atmospheric neutrino experiments.

Many other class of candidates are of interest relaxing some hypothesis that lead to the WIMP paradigm. For instance, considering non-thermal relic allows to develop dark matter model with very high masses called WIMPZILLA [27]. Different freeze-out scenario involving three body collision in the dark sector allows to propose candidate with very high cross-section denoted SIMP (Strongly Interacting Massive Particles) [28]. In the end, dark matter search extend across several tens of orders of magnitude both in cross-section and mass. A summary of the main candidates to dark matter is shown on figure 1.9. As the field widens and despite the lack of discovery, the WIMP remains the most motivated candidate but one can only appreciate the diversity of the theoretical and experimental efforts.

1.3 Dark matter searches

In this section, the three experimental approaches for dark matter physics are described. They can be represented using simplified Feynman diagrams as in figure 1.10. Within this depiction, the standard model particle the dark matter interacts with is not specified. The precise nature of the interaction is not described as well since it varies with the theoretical model.

1.3.1 Indirect search

Indirect search consists in observing astrophysical messengers that are evidences of dark matter in space. Those messengers are mainly photons or cosmic rays produced by the annihilation of dark matter that induce an excess in the energy spectrum. The main challenge of this approach is the

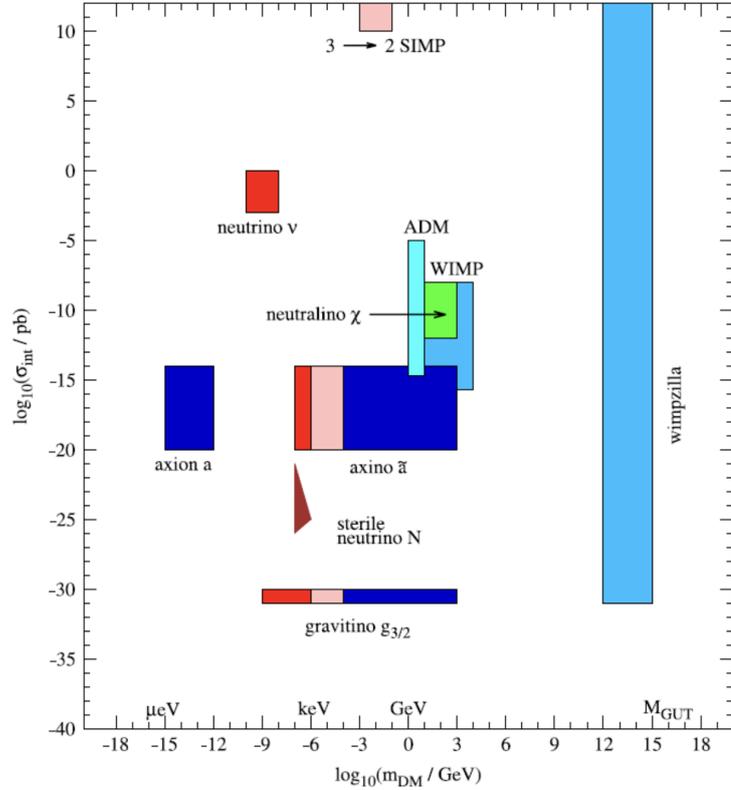


Figure 1.9: Summary of the main dark matter particle candidates in the cross-section-mass phase space[29].

requirement for a very good understanding of the background [30].

On the side of the photon probes, gamma signal is looked at by indirect search experiments in all energy ranges. In the MeV regime, the signature is Compton scattering as exploited in satellites like COMPTEL (Imaging Compton Telescope) or COSI (Compton Spectrometer and Imager), the latter planned to be launched in 2025. In the GeV and TeV ranges, Fermi Gamma-ray Space Telescope, HAWC (High Altitude Water Cherenkov Observatory), LHAASO (Large High Altitude Air Shower Observatory) are some of the main currently running experiments relying on electron-positron pair creation in space or radiative shower for ground telescopes. This detection channel is often considered preferential because γ rays travel in straight line for a very long distance. In particular, it allows to look at regions with high dark matter density, since the expected signal scales with the square of the latter [29]. In other channels like protons and positrons, one can only have access to nearby interactions.

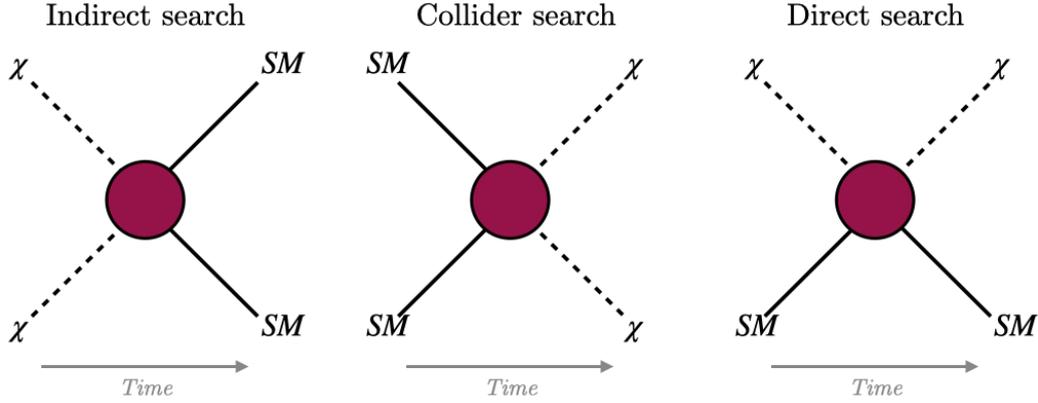


Figure 1.10: Schematic diagrams for the three methods of dark matter search.

In the X-ray sector, experiments like XMM-Newton, Suzaku, Chandra or NuSTAR lead the limits in $\mathcal{O}(0.1\text{-}10\text{ keV})$ masses. Those telescopes are especially focused on sterile neutrinos searches, in particular the previously mentioned unexplained line around 3.5 keV.

The study of cosmic-ray probes as dark matter annihilation product is centred around anti-particles. Since standard astrophysical sources emit mostly particles, the anti-particles fluxes from background sources are much lower. Processes involving dark-matter should not favour one over the other, therefore, it is the best strategy to look at anti-particle spectra, such as positrons or antiprotons. This is the case of BESS (Ballon-borne Experiment with Superconducting Spectrometer), PAMELA (Payload for Anti-matter Matter Exploration and Light-nuclei Astrophysics) or AMS-02 (Alpha Magnetic Spectrometer).

1.3.2 Collider search

The nature of dark matter can be probed in colliders experiment such as the LHC by producing the particle itself. At the LHC, the collision of protons with an energy of 13 TeV in the center of mass could result to the creation of particles beyond standard model. Unlike the standard particles produced by the collision, a dark matter particle will not be detected but will cross the detector and escape with a fraction of the initial kinetic energy of the collision. The signature of such product is therefore a difference between the total energy of the detected particles and the 13 TeV of the initial protons.

This is however made more complicated by missing energy coming also, in standard model processes, from neutrinos which carry some kinetic energy and are not detected similarly to dark matter. That is why, to detect dark matter in collider, it is needed to search for a specific theoretical model which predicts not only missing energy but also the behaviours of the standard particles that must be detected in the process. The expected process has then to be compared to standard model processes with the same particles in the final states and often with many orders of magnitude larger cross-sections.

As an example of dark matter search in collider, the theory of supersymmetry can be tested to some extent by looking for the predicted neutralino, which is a WIMP candidate naturally appearing from the theory. The stop quark, which is the supersymmetric partner of the top quark is expected to be the lightest supersymmetric quark or squark [31]. It is therefore the easiest to produce and can hypothetically decay into a set of jets, leptons, neutrinos and one neutralino $\tilde{\chi}_1^0$. The Feynman diagram of the process is in figure 1.11. This signal must be distinguished from all the standard background processes with similar particles at stakes and in which the missing energy is carried by neutrinos.

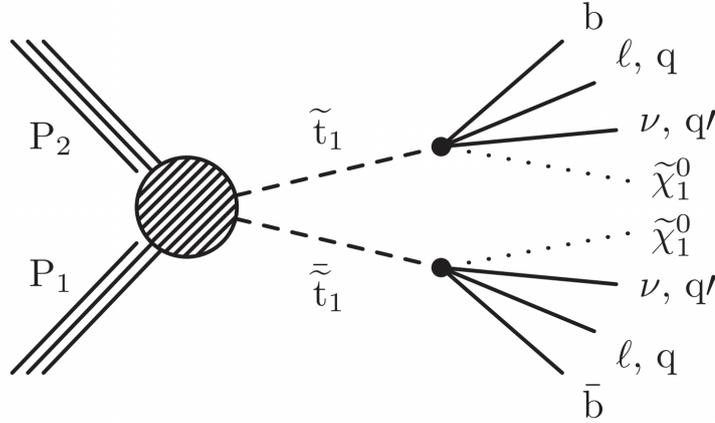


Figure 1.11: Stop pair production at the LHC with four-body decays

A less model-dependent search was carried out both in ATLAS and CMS, the two most important experiments of the LHC. It is a comparison of the missing transverse energy in mono-photon final state and the expected value from the standard model. This provides limits on a generic dark-matter particle which would induce an excess of events with missing energy [32]. Here

the main background is the decay of the Z boson into 2 neutrinos. This analysis and similar simple final state study were competitive with direct search measurements only in a restrained mass region and for spin dependent cross-section as shown in figure 1.12.

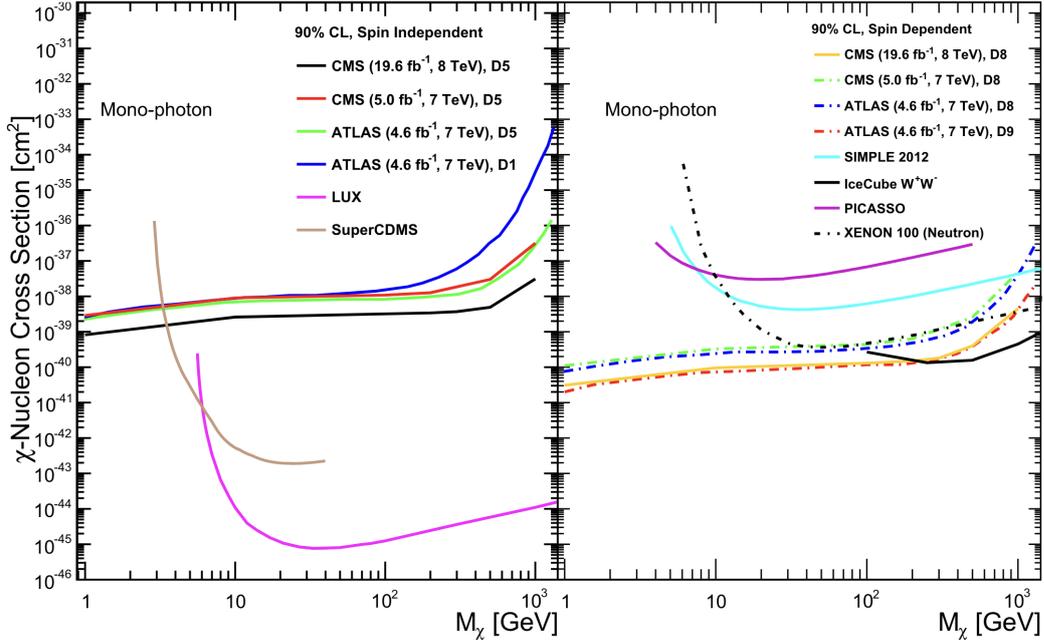


Figure 1.12: 90% confidence level limit on spin-independent (left) and spin-dependent (right) WIMP-Nucleon cross-section in ATLAS and CMS from the mono-photon channel compared to direct search experiments in 2014.

This kind of study exploits the very advanced technologies provided by the LHC and the great understanding of its detectors but it requires to focus on theoretical models which individually are not necessary well motivated. However, in the case of a discovery, it would not only give an evidence of the particle nature of dark matter but would also provide informations of the model in which it fits.

1.3.3 Direct search

The direct search for dark matter is the topic of this thesis. It regroups the full DarkSide program and many other experiments aiming to detect the interaction of a dark matter particle with an ordinary particle inside a

dedicated detector [33]. It is therefore a search for a local dark matter density signal from halo particles crossing the Earth and more precisely crossing the detector. The interaction in question is an elastic collision inducing a deposit of energy in the detector. This energy is distributed and can be seen through three different channels:

- Scintillation is the induce of light that can be seen by some photodetector such as photo-multiplier tubes (PMT) or silicon photo-multiplier.
- Ionisation of the medium produces free charges that can be collected with some electric field and detected.
- Heat is released during the interaction in the form of phonons. The increase of temperature can be detected in bolometers by using thermistors whose resistance is proportional to temperature.

Each detector is designed to have access to one or more of those channels providing different physics potential and background mitigation techniques. For instance, the CRESST (Cryogenic Rare Event Search with Superconducting Thermometers) experiment is focused on both heat and light detection. Calcium tungstate crystals are cooled down to 10 mK and read by superconducting thermometers and silicon light absorber.

The CDMS (Cryogenic Dark Matter Search) collaboration is using germanium and silicon detectors to read phonons and charge carrier from particle interactions. The charge and phonons are collected by aluminum and tungsten sensors. This kind of solid state detectors are sensitive to very low energy deposits (few tens of eV), eventually induced by low-mass dark matter.

Lastly, liquid scintillator detectors using noble gas liquid are the leading experiments above several of GeV/c^2 WIMP mass. This technology relies on the high scintillation and ionization yields of noble elements, mostly xenon and argon. Xenon is used in the XENON, LUX or PANDAX collaborations and argon is used in DEAP and DarkSide.

The common point with all these experiments is the constant struggle to reduce and control background being interaction in the detector from other sources such as radioactivity or cosmogenic particles. All of them are installed in underground laboratories to limit the amount of cosmogenic particles reaching the detector. The Gran Sasso underground laboratory (LNGS) built in Italy under 1400 meters of rock (3800 water equivalent meters) is shielding from cosmic ray cascades reducing the muon flux from

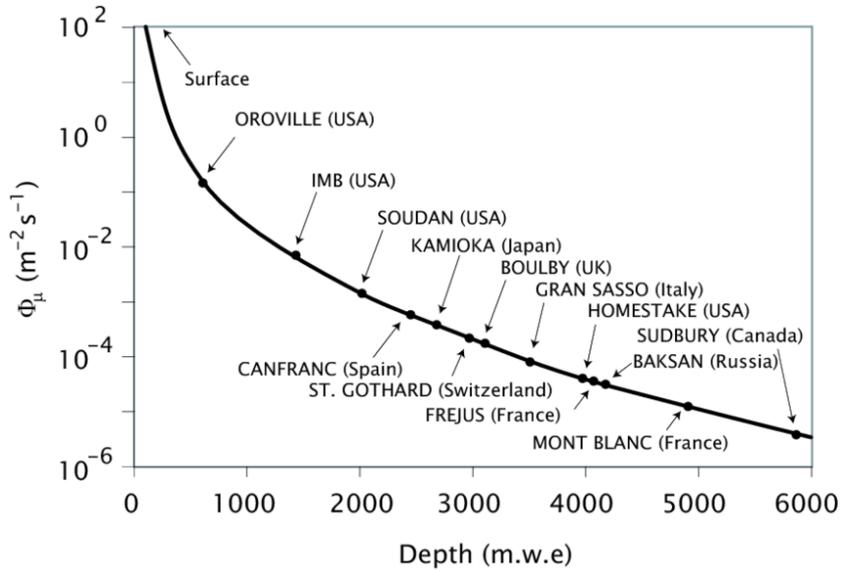


Figure 1.13: Muon flux and depth of underground laboratories.

$10^4 \text{ min}^{-1} \cdot \text{m}^{-2}$ at surface to $0.02 \text{ min}^{-1} \cdot \text{m}^{-2}$. It is the largest underground laboratory in the world and is the one which hosted DarkSide-50 and which will host DarkSide-20k.

Background is reduced by building these detectors choosing materials with extreme radio-purity. Still there is an ultimate irreducible background that are the solar and atmospheric neutrinos. Due to their low interaction rate, they cannot be shielded and have similar behaviour as dark matter. This implies a fundamental limit in the sensitivity of direct search experiment first described as a neutrino floor. The next generation of direct dark matter search detectors willing to tackle the neutrino floor is leading to a renaming of this limit as a neutrino fog, something that can be crossed not without difficulties.

1.4 Signatures in direct searches

To relate an excess of events to a dark matter signal, the good comprehension of the background is not enough. The astrophysics properties of dark matter must come into play to predict the shape of this signal. Some dependency can be derived either in energy, in time, or in direction to validate the nature of the excess. These different signatures allow to build not only counting

experiments, but comprehensive and complementary analysis.

1.4.1 Event rate

The dark matter signal in direct search experiments is the collision of a dark matter particle with an electron or a nucleus of the target. The transmitted kinetic energy is converted into a recoil of the target, either a nuclear or an electronic recoil. The number of expected events depends on the target mass m_A and on astrophysical parameters that are the local density ρ_0 and the speed distribution of WIMPs $f(\vec{v})$ in the laboratory referential frame. Finally it depends on the mass of the WIMP, or the considered candidate, m_χ and on its cross-section σ with the target. The differential flux can be expressed as follow:

$$\frac{dR}{dE} = \frac{\rho_0}{m_\chi m_A} \int \vec{v} f(\vec{v}, t) \frac{d\sigma}{dE}(E, \vec{v}) d^3\vec{v} \quad (1.6)$$

Additional form factors from nuclear physics describing high energy suppression can appear depending on the nature of the target. Although the cross-section depends on the very nature of the interaction which is unknown, it is known to scale with the square of the number of nucleons which favours the use of heavy nucleus. Giving the various inputs, this equation gives an exponential decrease for the rate as a function of the energy. The lowest is the particle mass, the steepest is the curve, favouring low energy recoils.

From this, dark matter can be search directly by looking for an excess in the energy spectrum compatible with an exponential decay on top of the background spectrum. The understanding of background is therefore the first challenge of such experiments. The energy range probed by the experiment will translate into a sensitivity to different mass, while the quality of background mitigation and rejection will lead to sensitivity to lower cross-sections.

1.4.2 Flux on Earth and annual modulation

Direct search consists in the measurement of the dark matter flux that is crossing the Earth. The average dark matter velocity is zero in the halo frame but Earth is not at rest in this frame. Then in the lab frame, the average velocity we expect is the one of the Earth in the halo frame which is around 220 km/s. Figure 1.14 shows the expected velocity distribution

provided by different simulations.

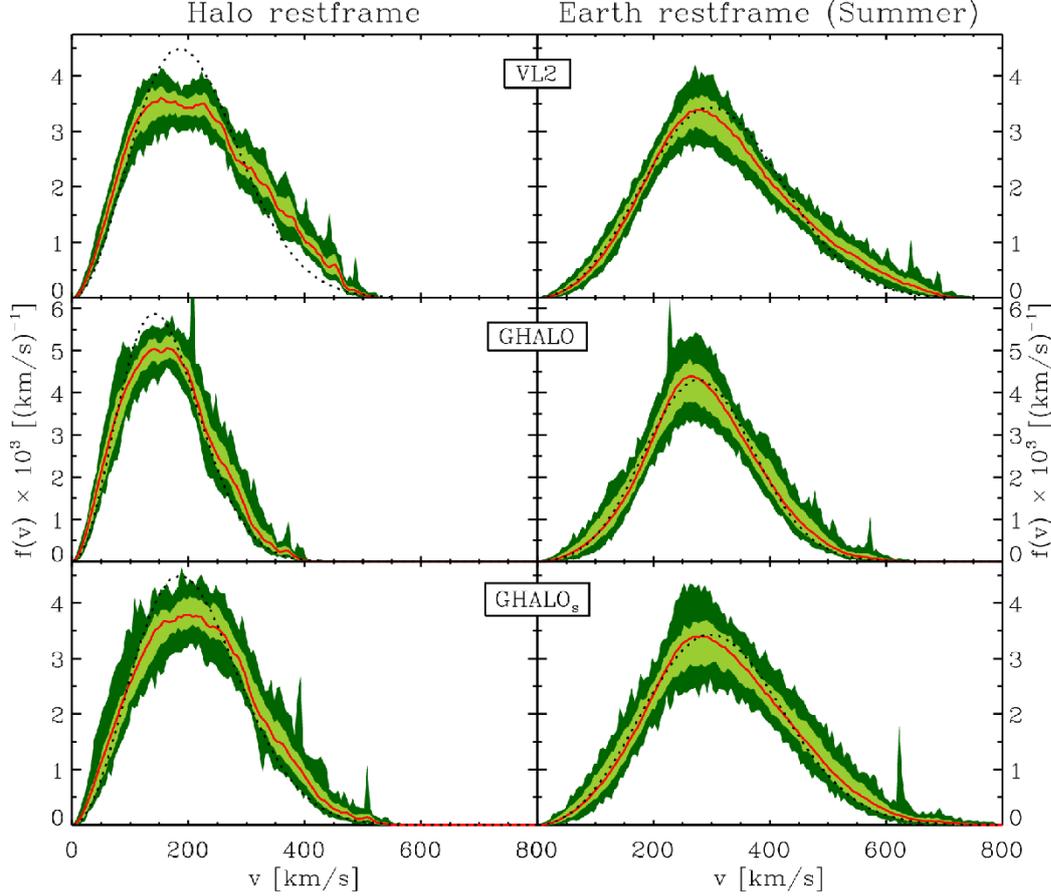


Figure 1.14: Local velocity distribution of dark matter in the halo and Earth restframes from three different simulations (Via Lactea 2, GHALO and GHALO_s). Dotted lines are the expected distribution from Maxwell-Boltzmann distribution. [12]

The latter is not constant along the year as the Earth is rotating around the Sun with a speed of 30 km/s. The orbital plane is tilted by 60° with respect to the velocity of the Sun in the Milky Way frame. The Earth velocity in this frame is maximum in June when it is most aligned with the Sun velocity. It is minimum in December when the two velocities are in most opposed direction. Therefore, the increase of the dark matter wind speed in June is of the order of $30 \text{ km/s} \times \cos(60^\circ) = 15 \text{ km/h}$. This implies a relative difference of the flux between June and December of $\sim 15\%$.

Some experiments aim to detect annual modulation which directly propagates in the expected event rate as a sinusoidal variation with a 1-year period. This is the case of the DAMA/LIBRA (Large sodium Iodide Bulk for RAre processes) experiment, a 250 kg solid crystal detector using NaI(Tl) coupled to low background photo-multiplier tubes. In 2010, after 4 years of data-taking, a periodic signal matching the expected signature from annual modulation was observed with a significance of 8.2σ (figure 1.15) [34]. However this discovery is rejected by the other experiments sensitive to the same parameter space. This suggests that the DAMA/LIBRA signal is likely due to an uncontrolled background. A common explanation is that this periodic signal is caused by a change of rate in cosmogenic background due to the atmospheric density variation along the year. To test this hypothesis, other experiments are trying to reproduce this result, notably a very similar setup being installed in the Stawell Underground Physics Lab in Australia. The idea is that if the DAMA/LIBRA signal is due to the weather change, the modulation should be out of phase in the other hemisphere. On the other hand, if it is due to dark matter, it should not change.

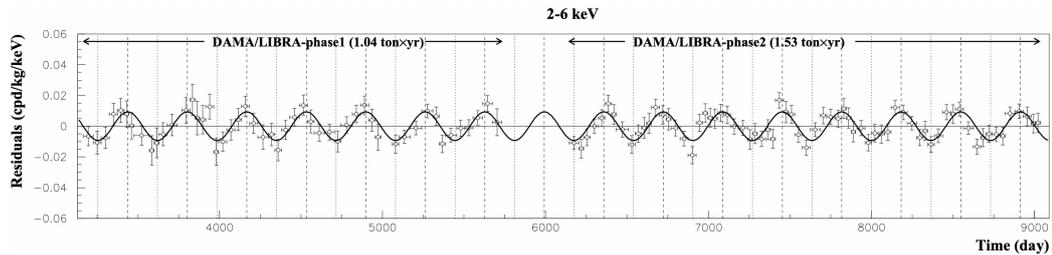


Figure 1.15: DAMA/LIBRA residual rate of single-hit scintillation events between 2 and 6 keV [34].

Similarly, daily modulations should also affect the expected rate with a one day period but the amplitude of this effect is negligible. What is however predicted is that when travelling through Earth, if cross-section is high enough, dark matter can be deviated or slowed by scattering, reducing the flux in the detector. This induces a shielding effect occurring at maximum intensity when the detector zenith is opposed to the incoming direction of the WIMP wind. As a result, a modulation of the rate with a 1-day period is expected. This provides especially strong sensitivity improvements for experiments where the background is hard to describe and for which the cross-section limits are high enough so that the expected earth-scattering induced modulation is stronger. In the DAMIC-M experiment, it improved

by 2 orders of magnitude the limit set on cross-section for masses between 1 and 2.5 MeV [35].

1.4.3 Directionality

Since the speed of Earth in the halo frame is close enough to the average speed of WIMPs, it is expected that in the Earth frame, they must have a preferential direction. In the halo frame, WIMPs have an average speed of 220 km/s and Earth velocity is 232 km/s. It must be taken into account that in the Earth frame, particles with lower speed are less likely to cross the detector. This effect should even further increase the mean velocity of WIMPs crossing the Earth. This is a feature that is used to discriminate WIMPs from radioactive cosmogenic backgrounds. In particular, it is the only known method to reject solar neutrinos for which the incoming direction is very well known. That is why several experiments focus on the directionality of the energy deposit to search for dark matter. The idea is that the recoil of light nuclei at low pressure could carry the directional signature [36]. The scalability of this technology is a challenge as the volume required for a gaseous phase detector to reach competitive target mass would be huge. Also it has potential only to low mass candidates due to the small energies at stakes.

In argon, which is a much heavier target, the displacement is smaller and the directionality harder to infer. Within the DarkSide collaboration, the ReD setup (Recoil Directionality) was designed to study if such measurement can be done in argon. A dual phase TPC is shot with a neutron beam to look for the signature of a track. With the current precision, it appears that such signal cannot be seen in argon [37].

Annual modulation and directionality analysis exploit astrophysical properties of dark matter to substitute background modelling but they require more statistics to claim for discovery. Only in a quite distant future when the neutrino fog is reached and if dark matter is not found by then, they could be among the last option to explore further phase space.

Chapter 2

The DarkSide project

The DarkSide-20k experiment is embedded in the DarkSide program which exploits the potential of liquid argon for direct dark matter search. In this chapter, will be first detailed the particle detection mechanisms of noble liquids. The DarkSide program will be presented starting from a description of DarkSide-50, the predecessor of DarkSide-20k, followed by a summary of the standard analysis for WIMP search in the $\mathcal{O}(100 \text{ GeV}/c^2)$ mass region. Then, we will discuss the design of DarkSide-20k with an exhaustive list of the different backgrounds.

2.1 Noble liquid detectors

Noble element technologies are leading the sensitivity in the search for high mass ($>10 \text{ GeV}/c^2$) dark matter particles. This is thanks to the two high yield detection channels that they provide, ionization and scintillation, along with the scalability and high chemical purity that can be reached [38]. Those properties are common to every noble gas but some crucial characteristics, such as photon yields, ionization yields, or scintillation times differ from one to another, as reported in table 2.1.

As things stand now, two noble elements are used as target for direct dark matter search, argon and xenon. XENONnT, LUX-ZEPLIN, and PANDAX-4T experiments are dual phase Time Projection Chambers (TPC) using liquid xenon. Xenon has a higher atomic number than argon and the cross-section with WIMPs increases with the square of this number. On the other hand, a lighter target will make detectable recoils at lower energy, allowing lower threshold in argon experiments. In addition to that, xenon is an expensive material, limiting the scalability of xenon-based experiments while

	Ar	Xe	Ne
Atomic number	18	54	10
Atomic mass	40.0	131.3	20.2
Boiling temperature (K)	87.3	165.0	27.1
Liquid density (g/cm ³)	1.4	2.9	1.2
Scintillation wavelength (nm)	128	178	78
w (eV)	23.6	15.6	20
Fast scintillation component (ns)	7	4.3	<10
Slow scintillation component (ns)	1500	22	10000

Table 2.1: Physical and scintillation properties of noble liquids [39].

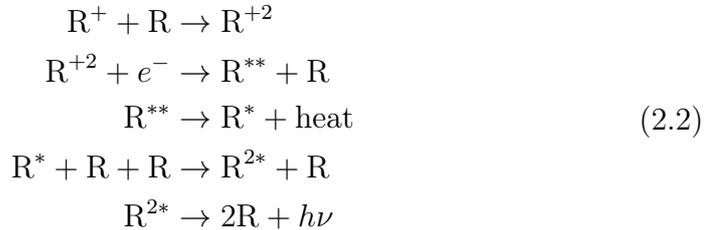
argon is very abundant (0.93% of the atmosphere) [40].

2.1.1 Scintillation and ionization in noble elements

Upon the recoil of the nucleus or electron, nearby atoms will either be ionized or excited [41]. The excitation process can be described as follow:



Here, R is the noble atom and R^{*2} is the excimer state, an unstable molecule responsible for the scintillation light. Measurements demonstrated that the excimer is formed through a three-body collision [42]. $h\nu$ is the energy of the emitted photon which, in liquid argon, corresponds to a wavelength of 128 nm. In liquid xenon, the emission peak is broader, centred around 175 nm with a 14 nm width [43]. The ionization process is described as:



The $R^{**} \rightarrow R^* + \text{heat}$ reaction is a non-radiative decay of the R^{**} excited state leading to the same situation as in 2.1. $R^{+2} + e^- \rightarrow R^{**} + R$ is the recombination process which happens between a free electron and the ionized atom. Under an electric field, the probability of recombination lowers as

electrons and ions are drifted toward opposite direction. A charge collection system can be used to detect the drifted electrons, providing observables of the ionization channel.

The combined detection of scintillation and ionization gives access to the deposited energy of electronic recoil E_{er} and nuclear recoil E_{nr} through the following relations:

$$\begin{aligned} E_{er} &= w (n_\gamma + n_{e^-}) \\ E_{nr} &= w (n_\gamma + n_{e^-}) \frac{1}{\mathcal{L}} \end{aligned} \tag{2.3}$$

Here, w is the average energy required to produce an excited or an ionised atom and n_γ and n_{e^-} are the number of photons and free electrons produced. Some of the deposited energy is lost due to quenching during the interaction. The quenching is the reduction of the measured energy due to undetectable energy loss in the medium. Quenching can correspond to heat, vibrations, or undetected radiations. For argon and without any electric field, it was observed that the quenched energy is a constant fraction of the electron recoil energy (figure 2.1). Therefore, it can be ignored as the accessed energy always remains proportional to the deposited energy. This is not true for nuclear recoil. \mathcal{L} is the Lindhard factor which describes, as a function of the energy, the relative difference in quenching between electronic and nuclear recoils.

2.1.2 Dual phase time projection chamber

Noble liquid TPC is a technology aiming at detecting both scintillation and ionization channels. The DarkSide project exploits the dual phase technology of TPC with liquid argon. On the top of the detector, the noble element is in a thin gaseous phase layer, with a thickness around 1 cm. This gas pocket is the charge collection system. The scintillation pulse in the liquid phase is detected immediately by multiple photosensors providing time and energy observables. This first pulse is called S1. The ionised electrons are then drifted by an electric field to be detected in the gas pocket. The speed of the electron increases with the intensity of the field. After some drift time, the electrons produce a second signal denoted S2 when reaching the gas. The drift time is proportional to the distance travelled by the electron, allowing depth reconstruction of the event with complementary information on the energy.

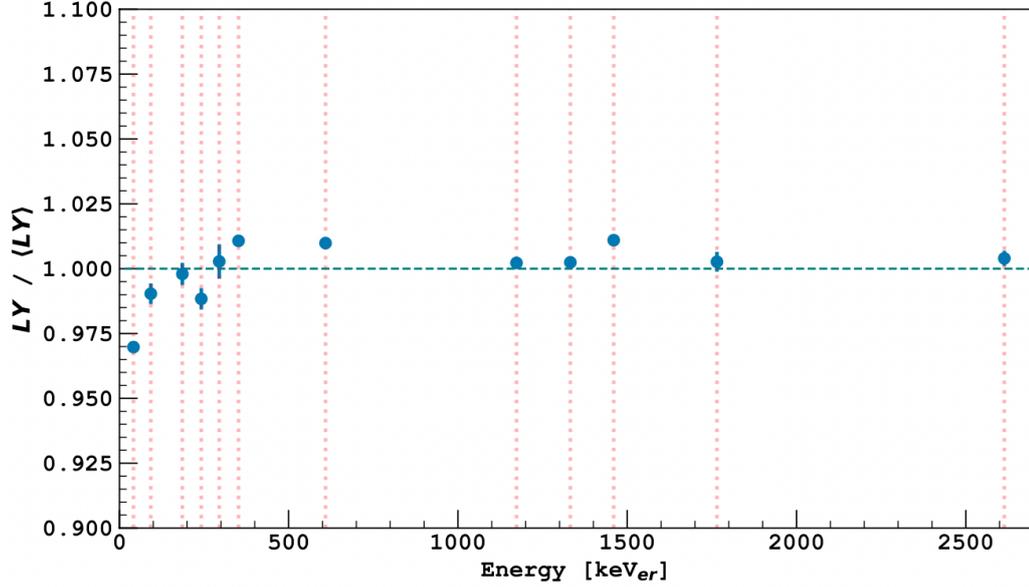


Figure 2.1: Light yields, normalised to the mean value, from each radioactive background gamma peak identified in the DarkSide-50 null-field spectrum. The light yield, namely the number of detected photons divided by the deposited energy, is constant within 2%, so it is with the quenching. The deviation at low energy may be explained by the increase of the fraction of energy deposits that are lower than w .

In the gas pocket, a higher electric field is applied inducing electrons to produce an electro-luminescence pulse when they arrive (figure 2.2). The increase of the electric field is achieved by placing an additional conductive grid under high voltage in the horizontal plane just below the gas pocket. The electro-luminescence mechanism differs from the one of scintillation. In addition to the excimer emission, the drifting electrons induce neutral bremsstrahlung, namely scattering on atoms [44]. That leads to emission below the excitation threshold of the atom. The time profile of the electro-luminescence pulse is very different than the S1 pulse, allowing to easily distinguish between them. The electro-luminescence process is characterised by the amplification factor g_2 which is the average number of photo-electrons detected for each electron reaching to the gas pocket. g_2 is higher than the detection efficiency g_1 , hence the S2 pulse counts much more photo-electron than S1. The kinetic energy of the event can therefore be reconstructed as:

$$\begin{aligned}
E_{er} &= w \left(\frac{S1}{g_1} + \frac{S2}{g_2} \right) \\
E_{nr} &= w \left(\frac{S1}{g_1} + \frac{S2}{g_2} \right) \frac{1}{\mathcal{L}}
\end{aligned}
\tag{2.4}$$

Here, S1 and S2 are expressed in number of detected photo-electrons (PE). It is sometimes suggested that the electric field induces some extra quenching, changing the value of w and \mathcal{L} with respect to null-field. In a first approximation, this effect can be neglected.

Since the electro-luminescence signal happens in the gas pocket and therefore very close to the top photosensor array, the detection channels close to the interaction thus observe most of the light. This is exploited to reconstruct the projection of the interaction position in the horizontal plane. Combined with the drift time measurement, it allows the full reconstruction of the event position in space. In general, position reconstruction is crucial for fiducialisation, namely the rejection of events coming from the edges of the detector, those events being more likely to come from material radioactivity. It also allows to separate events occurring at the same time but at different location in the TPC.

2.1.3 Pulse shape discrimination in argon

One of the main advantage of argon over other noble gases is the excellent discrimination power between electron and nuclear recoil. The S1 pulse shape strongly varies whether the interaction of the incoming particles occur with an electron or a nucleus. This allows to reject very effectively the electronic recoil background, bearing in mind that the signal expected from dark matter is a nuclear recoil. This feature is particularly crucial in liquid argon as it also provides rejection of ^{39}Ar background, a beta emitter naturally present when argon is extracted from the atmosphere. The power of pulse shape discrimination in argon was recently demonstrated by DarkSide-50 and DEAP [55].

The difference between the nuclear and electronic recoil pulse shapes comes from the argon excimers responsible for the scintillation light. The excited argon atom can be described as a Rydberg state considering an Ar^+ ion plus an electron with an orbital radius very large compared to the atom radius, it then effectively behaves like an hydrogen atom [45]. This allows a

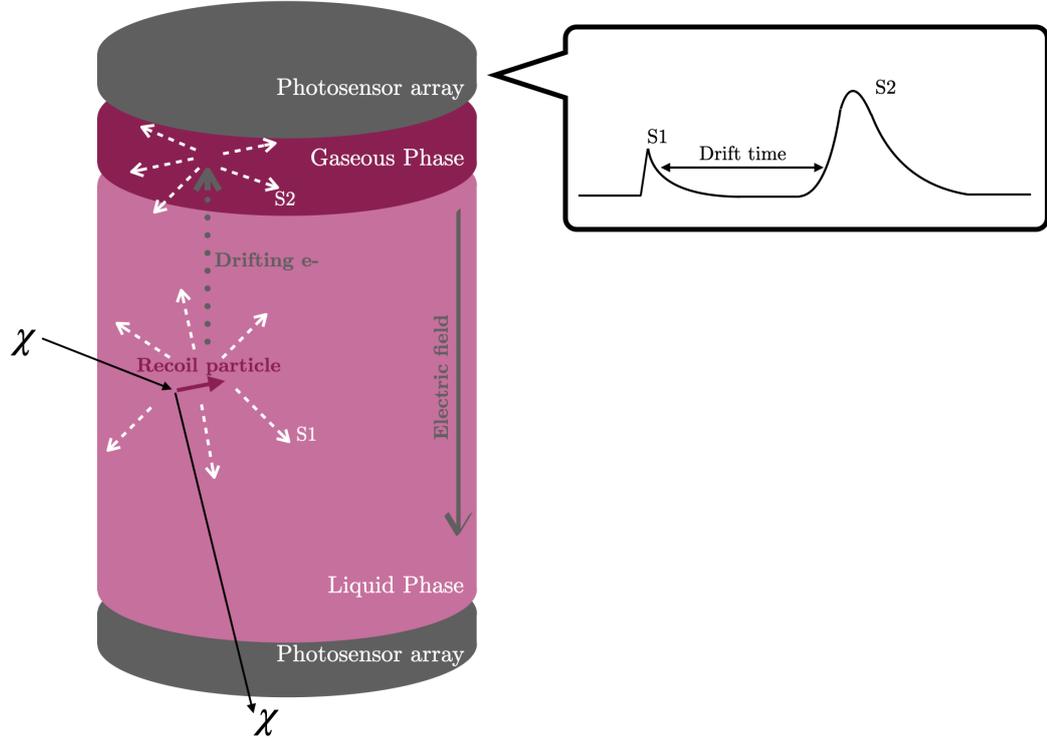


Figure 2.2: Detection principle of a dual phase time-projection chamber.

covalent bound between the two argon atoms to form an entangled system whose state is the tensor product of the spin states of the two electrons and their orbital angular momentum:

$$|\psi\rangle = |\sigma_{12}\rangle \otimes |o_{12}\rangle \quad (2.5)$$

As an entanglement of two identical particles with spin 1/2, the system has to be anti-symmetric, namely the eigenvalue of the exchange operator must be -1 . This leaves two possibilities, either the spin state $|\sigma_{12}\rangle$ is anti-symmetric and the orbital state $|o_{12}\rangle$ is symmetric or the opposite. From the four possible spin associations ($|\uparrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$, $|\uparrow\downarrow\rangle$, and $|\downarrow\uparrow\rangle$), only one linear combination is anti-symmetric, forming a singlet state, and 3 are symmetric, forming a triplet state:

$$|s_{singlet}\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (2.6)$$

$$|s_{triplet}\rangle = c_{+1} |\uparrow\uparrow\rangle + c_{-1} |\downarrow\downarrow\rangle + c_0 \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}$$

The total angular momentum of the combination of two spin 1/2 particles

is 0 or 1. Since one state is accessible for the spin anti-symmetric configuration, it implies that the total angular momentum is 0 with only 0 as a possible eigenvalue for the projection. Similarly, the three accessible states for the spin-symmetric case indicates that the total angular momentum is 1 with projection taking values among 1, 0, and -1. The spin of ^{40}Ar being 0, so would be the spin of the dissociative ground state of the excimer. This means that only the singlet decay is allowed by momentum conservation. Therefore the triplet state is strongly suppressed and decays in $1.5 \mu\text{s}$, while the de-excitation of the singlet state is the fastest (6 ns). As a matter of comparison, in xenon nuclei whose spin are $1/2$, the singlet state and triplet state lifetimes are 4.3 and 22 ns [46].

The relative population of excimer states between singlet and triplet depends on the linear energy transfer (LET), namely the amount of energy deposited by unit of distance travelled by the recoiling particle. The higher the LET, the lower the triplet-to-singlet ration will be. This can be explained considering that the exchange symmetry or anti-symmetry slightly affects the energy of the singlet and triplet states. In addition to the repulsive Coulomb interaction between the two electrons, an exchange term appears in the Hamiltonian to favour for the singlet or penalise for the triplets overlapping orbitals [47]:

$$\begin{cases} E_{singlet} = E_{o,1} + E_{o,2} + E_C + E_{ex} \\ E_{triplet} = E_{o,1} + E_{o,2} + E_C - E_{ex} \end{cases} \quad (2.7)$$

Here, $E_{o,1}$ and $E_{o,2}$ are the orbital energies of the first and second electrons respectively. E_C is the energy from Coulomb interaction and E_{ex} is the exchange energy term, which is positive. So the triplet energy is slightly lower than the singlet and is favoured by low LET.

As nuclear and electronic recoils have a very different LET, the relative singlet and triplet population are significantly different, more triplets states are populated for electronic recoils [48]. Regardless of the nature of the interaction, the pulse shape will be a combination of two exponential decays, with 6 ns and $1.5 \mu\text{s}$ lifetimes. However the ratio between the fast and the slow components will strongly differ allowing to discriminate between electronic recoil and nuclear recoil from the shape of the pulse.

To exploit the pulse shape discrimination (PSD) power of liquid argon, a dedicated observable is defined on S1 pulses as the fraction of prompt light over the total light. The duration of the prompt window has to be optimised for each detector as it depends on the photo-detection system. In the case

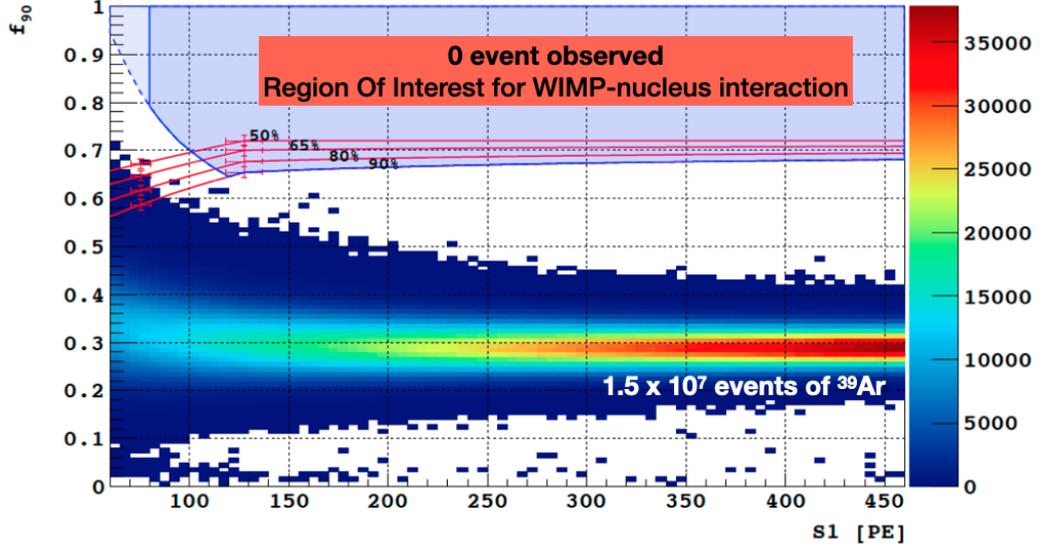


Figure 2.3: Pulse shape discrimination in DS-50 from atmospheric argon run showing a rejection power higher than $1 : 10^7$. The red lines are different acceptances of the NR band.

of DarkSide-50, it was set to 90 ns. This fraction is thus denoted f_{90} and takes higher value if the pulse is fast. Therefore the NR and ER are expected to populate two different bands in the $(f_{90}, S1)$ space where $S1$ is the total light in number of PE. The observed ER band of ^{39}Ar and the expected NR band is shown from DarkSide-50 data in figure 2.3. NR are populated by more singlet states so they induce a faster signal, so a higher f_{90} value. The bands are only overlapping at low energy so the region in which the strong rejection is possible defines the region of interest for WIMP search. This is the key feature of liquid argon for dark matter search, made possible by the 3 orders of magnitude between the two lifetimes.

2.2 From DarkSide-50 to DarkSide-20k

The DarkSide project started with DarkSide-10, a dual-phase prototype filled with 10 kg of underground argon that operated in LNGS from 2011 to 2012 [49]. The use of underground argon is at the core of the DarkSide project since its beginning and aims at reducing the activity of ^{39}Ar which is produced by spallation of cosmic ray in the atmosphere and has a half-life of 269 years. Argon extracted from underground sources is protected from cosmic

rays and therefore depleted of ^{39}Ar .

The predecessor of DarkSide-20k is DarkSide-50, a smaller scale detector which was the first liquid argon dual phase TPC with physics results in the WIMP search. Also located in the hall C of LNGS, it operated between 2013 and 2019 and demonstrated the great physics potential that can be reached with low-radioactivity argon.

DarkSide-20k is under construction since 2022, and the data-taking is scheduled to start in 2027. It is once again located in the hall C of LNGS. A complete technical description of the detector is provided in the DarkSide-20k Technical Design Report [50]. After DarkSide-20k, the possibility for the construction of a next generation experiment is under discussion. This project, named ARGO, would be carried out by the wide Global Argon Dark Matter Collaboration (GADMC) which is the union of 400 researchers working on dark matter search experiments with liquid argon. The goal would be to build a ~ 300 tons detector with a sensitivity reaching the neutrino floor.

2.2.1 The DarkSide-50 performance

DarkSide-50 had an active mass of 46 kg of underground argon and operated for 532 live days. The 35 cm high TPC was placed inside a liquid scintillator veto for neutron and γ rejection [51]. This veto consists in 30 tons of borated scintillator, boron having high neutron capture cross-section. The liquid scintillator was itself placed under a water Cherenkov muon veto. The kiloton of water is also a good passive shielding against neutrons. The TPC is seen by two arrays of 19 photomultiplier tubes (PMTs), one on top and one on the bottom. A scheme of the DarkSide-50 is shown figure 2.4. The gas pocket is 1 cm thick and the maximum drift time was $376 \mu\text{s}$ while it will be 3.7 ms in DarkSide-20k.

The full 532.4 live day data-taking of DarkSide-50 allowed to reach an exposure of $16\,660 \pm 270$ kg.d. During this period, the average trigger rate was around 1.2 Hz with major contribution from γ rays and internal contamination of ^{85}Kr and ^{39}Ar . The standard analysis is based on the S1 energy reconstructed through the number of PE, and the PSD estimator f_{90} defined as the fraction of light detected within 90 ns of the S1 pulse [52].

The data selection starts by requiring that for each event, data are present for all channel, baselines of the waveforms are successfully computed, veto

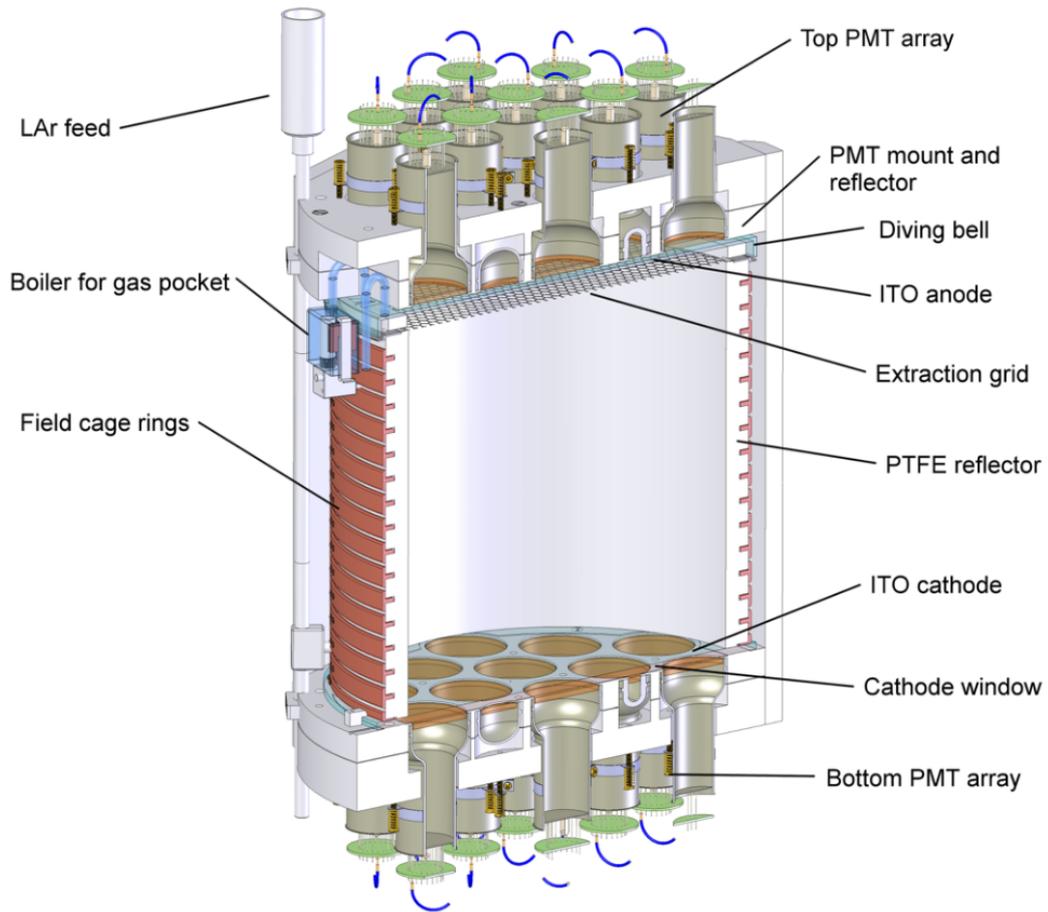


Figure 2.4: DarkSide-50 TPC.

data are synchronised and accessible at this time, and the previous event occurred more than $400 \mu\text{s}$ ago.

The next set of cuts is designed to reject multiple scatter. In DarkSide-50, the maximum drift time is small ($376 \mu\text{s}$) so an event is defined as a triggered time-window that must contain S1 and S2. The selection criteria are the following: the pulse must occur at trigger time; there must be a second pulse identified as S2; there must be no saturation in the first pulse; the second pulse must be larger than 200 PEs; the second pulse must have $f_{90} < 0.20$; transversal position must be successfully derived. Then a low-threshold S2/S1 cut is applied to further remove unphysical events.

Another set of cuts is applied for the rejection of surface background mostly induced by α decays. It includes rejection of long tail S1, a high-

threshold $S2/S1$ cut, and a pulse shape cut in $S2$.

The neutron background and muon rejection cuts are based on the liquid scintillator and the water Cherenkov vetoes. In the liquid scintillator veto, 3 windows are defined around the event time with different threshold for rejection. In the water veto, in addition to a coincidence cut to reject muons, a delayed veto window is applied to account for neutrons produced by cosmic-ray-activated isotopes.

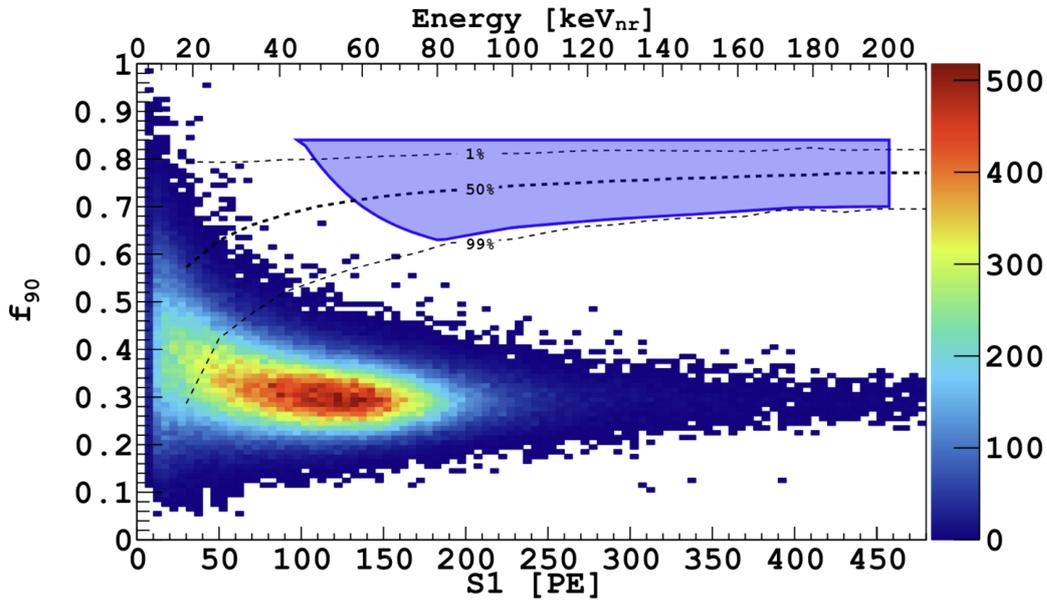


Figure 2.5: Event distribution in the $(S1, f_{90})$ space after all analysis cut. Blue area indicates the WIMP search region. the dashed line are the acceptance contours for NR derived from $^{241}\text{ArBe}$ calibration.

The last set of cuts aims at ER background rejection. It is at this stage that are applied radial and vertical fiducialisation and an $S1$ prompt fraction cut to remove events with Cherenkov light produced in the PMT or TPC windows.

The total acceptance of those cuts in the WIMP search region of the $(S1, f_{90})$ space, excluding the data quality set, is of 0.725 ± 0.005 . As shown in figure 2.5, no events were found in the WIMP ROI after applying all the analysis cuts. This result demonstrated the underground argon potential for background free WIMP search. The derived limits on the dark-matter-nucleon cross-section are presented in figure 2.6. This analysis set an upper

limit on WIMP-nucleon cross section at $1.14 \times 10^{-44} \text{ cm}^2$ for $100 \text{ GeV}/c^2$ dark-matter particle.

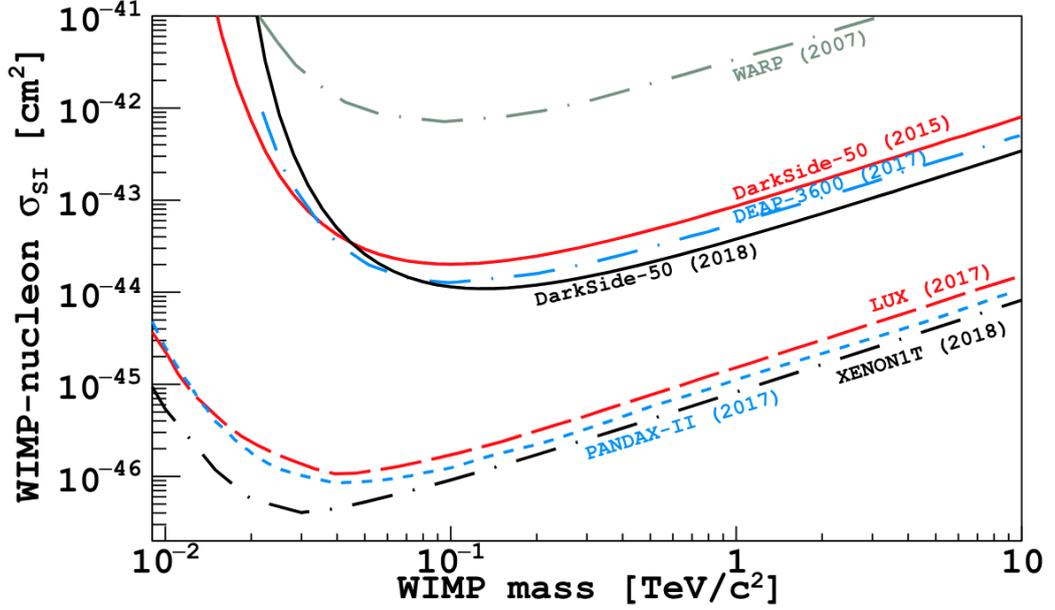


Figure 2.6: Spin-independent DM-nucleon cross-section 90 % confidence level in 2018.

DarkSide-50 data were also exploited for a low-mass analysis detailed in section 6.3.2. The success of DarkSide-50 led down the path for DarkSide-20k. As 3 orders of magnitude separate the target mass of the two experiments, the full exposure of DarkSide-50 will be obtained in one day of data with DarkSide-20k. Still, the large scale and the new technology involved for the latter portend challenges at the design, simulations, construction, data management, resources management, and analysis levels.

2.2.2 The DarkSide-20k TPC

DarkSide-20k is a dual-phase TPC filled with 50 tons of underground liquid argon under construction at LNGS. After fiducialisation, the active target mass in the centre of the TPC will be around 20 tons. The TPC is a right prism whose base is a regular octagon. The diameter of the octagon inscribed circle and the height of the TPC are both 350 cm. The octagons on top and on bottom of the detector are fully covered with silicon photomultipliers

(SiPMs) over a total surface of $2 \times 21 \text{ m}^2$. The optical planes are designed to provide a light yield of approximately 8 PE/keV. The top plate also features a diving bell to contain the gas pocket.

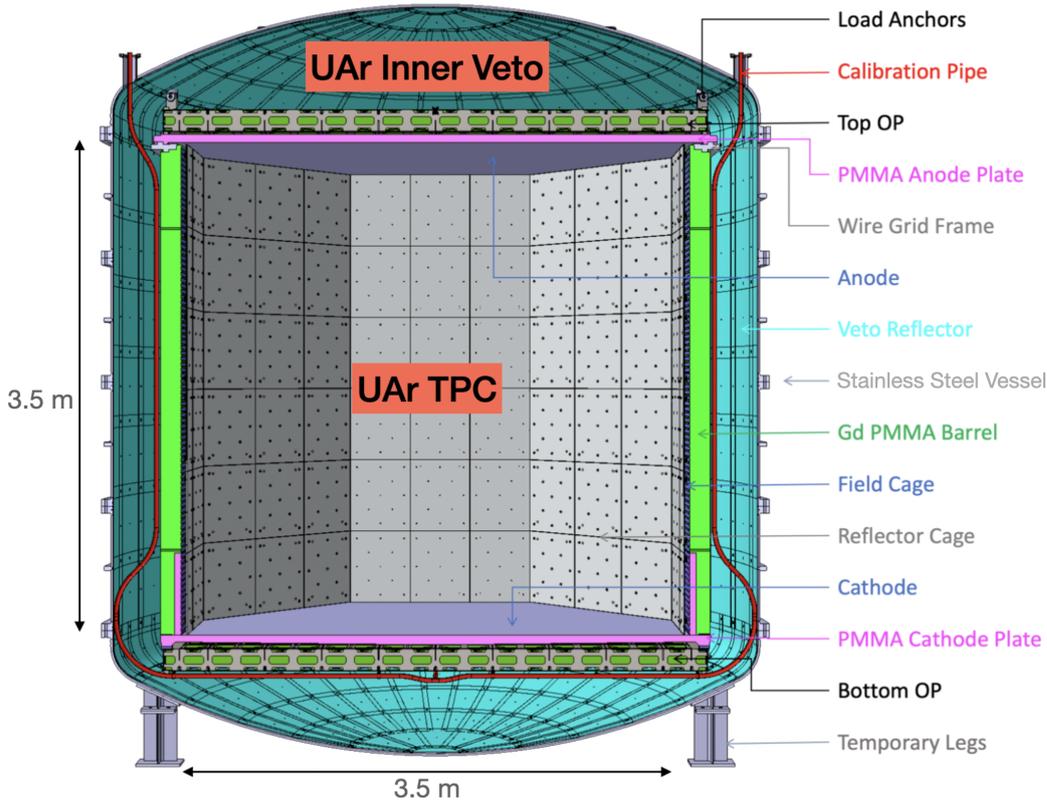


Figure 2.7: DarkSide-20k TPC, inner veto, and stainless steel vessel.

The optical planes (OP) and the lateral walls of the TPC are coated with Tetra-phenyl Butadiene (TPB) which is a wavelength shifter converting the 128 nm light from argon scintillation into 420 nm light that can be detected by the SiPMs. An electric field is applied from the top plane (the anode) to the grid, and from the grid to the bottom plane (the cathode) to drift electrons to the top of the detector.

The walls are made of Poly(Methyl Methacrylate) (PMMA), a transparent thermoplastic, doped with gadolinium which has a very high neutron capture cross-section. In addition to the PMMA, a $50 \mu\text{m}$ layer of Enhanced Specular Reflector (ESR) is added in order that light hitting the walls will eventually reach the OP after reflections. The geometry of the walls is shown

on figure 2.8. The reflectivity of ESR was measured to be around 98%. ESR is coated with TPB and rings of Clevios, a conductive polymer, are added on the walls forming a field cage to guide the electric field. The Clevios rings are spaced by 4 cm and will ensure uniformity of the 200 V/cm drift field.

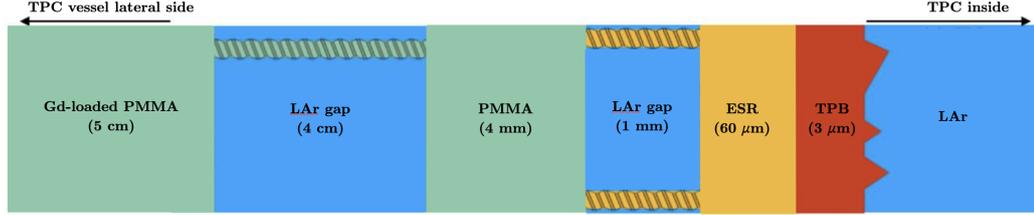


Figure 2.8: Geometry of the TPC walls including ESR and TPB coating. The shape of the TPB layer indicates expected non-uniformity of the coating.

The stainless steel vessel contains both the TPC and a liquid argon buffer used as a veto around it. The TPC and the inner veto form the so-called inner detector. An extra plastic layer will be placed on the outside of the vessel to moderate external neutrons. The other main elements of the inner detector are the PMMA placed over the OP, calibration pipes used to circulate radioactive sources to calibrate the detector, and structural parts. The geometry of the DarkSide-20k inner detector is summarised figure 2.7.

2.2.3 Nuclear recoil and electronic recoil backgrounds

The interactions expected in the TPC forming the background of the experiment can have cosmogenic or radiogenic origins. They can be internal backgrounds if the source is the detector itself or external if they come from the environment or are produced by cosmic-ray interaction in the atmosphere. It is important to distinguish backgrounds inducing ER that can be efficiently rejected by PSD from those inducing NR that could fall in the WIMP search region. The physical backgrounds expected in DarkSide-20k can be listed as follow:

- ^{39}Ar : Internal contamination in liquid argon, naturally produced by cosmic-ray spallation on atmospheric argon. It decays through β^- emission with a Q-value of 565 keV, inducing ER events. The underground argon extracted for DarkSide-50 showed a depletion factor of ^{39}Ar of 1400 with respect to atmospheric argon for a residual activity of 0.73 mBq/kg.

- **^{85}Kr :** Radioactive contaminant in liquid argon. Also decays through β^- emission with a Q-value of 687 keV. Given its anthropogenic nature, it is not yet confirmed that the underground argon extracted for DarkSide-20k will be contaminated with ^{85}Kr after purification. Indeed, ^{85}Kr is in majority produced by nuclear fuel reprocessing, as it is a volatile medium-lived fission product that can escape nuclear waste and diffuse in the atmosphere. Since the presence of ^{85}Kr was not expected in DarkSide-50, it is not excluded that is due to some leakage. In that case, the activity of ^{39}Ar would also have been increased. This suggests that the activity of ^{39}Ar might be lower in DarkSide-20k than what was observed in DarkSide-50.
- **γ :** Emitted by natural radioactivity in a very wide energy range. The SiPMs, the PMMA walls, and the stainless steel vessel are the most important sources of γ activity. They come from ^{238}U , ^{235}U , and ^{232}Th decay chains and radioactive isotopes naturally present in all materials such as ^{137}Cs , ^{60}Co , or ^{40}K . γ interaction produces ER. At high energy ($\mathcal{O}(\text{MeV})$), γ interacts multiple times in the detector which is a way to discriminate them from a WIMP signal.
- **α :** Emitted by radioactivity. They can also occur in liquid argon due to radon contamination. They induce NR but α particle have a small radiation length in matter, thus most of them do not reach the TPC. In addition, the interaction mechanisms in argon are different than with respect to other particles. The energy partition between scintillation and ionization is different enough to discriminates α events with an S1/S2 cut.
- **Muons:** Coming from cosmic-ray shower in the high atmosphere. Despite the shielding from the Gran Sasso mountain, muons will reach the detector, producing a big amount of light along their tracks.
- **Neutrons:** Originating from (α, n) reactions, spontaneous fission, or cosmic rays. The main contributing sources are the cryostat and the rocks around the laboratory. They represent the most dangerous background as they induce NR in the energy range of the region of interest. Neutrons captured on the TPC gadolinium-loaded PMMA emit a γ flash from which the event can be tagged.
- **Neutrinos:** Emitted by the Sun or produced in the atmosphere. They can interact through elastic collision with a nucleus or an electron, or they can interact via inelastic charged current interaction. Only

the coherent elastic neutrino-nucleus scattering ($\text{CE}\nu\text{NS}$) induces a NR signal. This NR background is a single scatter since neutrino cross-sections are very small. It cannot be suppressed nor rejected. However the expected rate is well known and the expected number of events is sufficiently low not to affect too much the sensitivity of DarkSide-20k.

In general, background can be reduced by shielding the detector and by using material with high radio-purity. Online rejection is possible by using vetoes surrounding the detector to tag background events, and offline rejection is done through dedicated analysis cuts. In any case the knowledge on the activities of the diverse background sources is a crucial point to understand the data at the analysis stage. Moreover, even if effectively rejected, a high background rate introduces dead-time in the detector which directly impact the final sensitivity. It also increases the risk of saturation in the data acquisition system.

2.2.4 Outlier backgrounds

Besides the NR and ER backgrounds, we call outlier background the classes of events which might artificially fall in the region of interest due to instrumental effects or random coincidence. The first class of outlier arises from ER events accompanied by Cherenkov radiation in the wall of the TPC. When crossing the acrylic, γ s travel faster than light in this medium, causing the emission of extra photons in a very short time window. This additional light increases the prompt fraction of the pulse as illustrated in figure 2.9. This creates a risk for the event to fall in the NR band.

Another class of outlier events arises from S1-only α events from the wall of the TPC in coincidence with unresolved top event reconstructed as S2-only (figure 2.10). Due to some inhomogeneities in the electric field, electrons close to the lateral walls can be absorbed by them during their drift. As a result, it is expected that these events do not give induce S2 signal but only a NR S1. The scintillation signal can also occur in the reflector itself or in the TPB coating which are $65\ \mu\text{m}$ and $3\ \mu\text{m}$ thick respectively. The scintillation properties of those materials are different from the one of argon, but it is still possible for an α event to produce an S1 that might fall in the WIMP search region [53]. For other particles, the mean free path is much higher than the thickness of the materials, so they are expected to produce also an ER signal in the TPC therefore they do not represent a dangerous background.

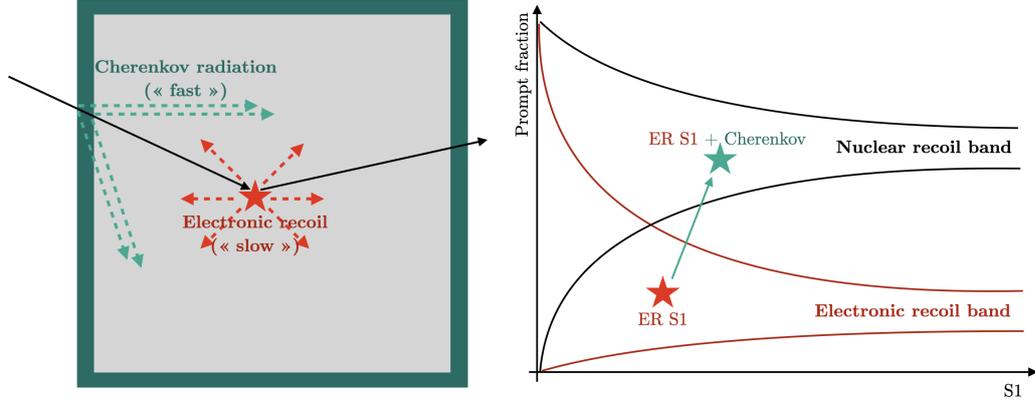


Figure 2.9: Schematic of the fast Cherenkov light produced in the walls for an ER event. The prompt fraction of the S1 pulse increases up to the NR band.

	Relative light yield	Characteristic time
1st	0.19	5.2 ns
2nd	0.81	2.3 ms

Table 2.2: Properties of the two TPB scintillation components, the total light yield is 0.41 photons/keV.

The scintillation in ESR and TPB was added in the simulation tools described in section 3.1.1 to produce α samples from the walls (table 2.2 and table 2.3). Given the activities in those material, the expected rate of WIMP-like S1 is computed. This is done by applying two cuts, an upper limit on the number of PE and a lower limit depending on the number of PE on f_{200} . That way, the rates are estimated at $31.37 \pm 0.14 \mu\text{Hz}$ and $1.58 \pm 0.06 \mu\text{Hz}$ for the ESR and TPB respectively.

On another side, interactions close to the top of the TPC may produce

	Relative light yield	Characteristic time
1st	0.53	17 ns
2nd	0.24	110 ns
3rd	0.14	2.1 μs
4th	0.09	48.0 μs

Table 2.3: Properties of the four ESR scintillation components, the total light yield is 0.35 photons/keV.

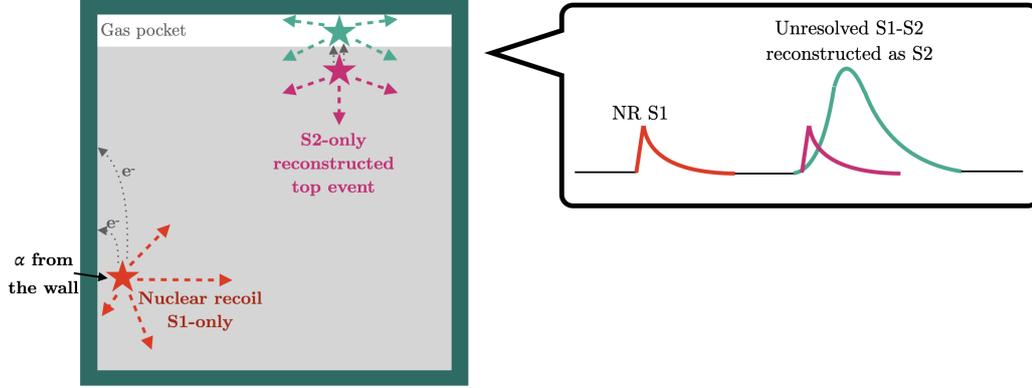


Figure 2.10: Accidental coincidence between S1-only event due to α events from the wall and S2-only reconstructed top event.

overlapping S1 and S2 which are reconstructed as single pulses. Since the amount of light is dominated by the ionization component, this pulse will be reconstructed as an S2. I estimated, by applying the reconstruction tools defined in section 5.2.2, the expected rate of such S2-only events to be 1.6 ± 0.4 Hz. The ^{39}Ar activity in the 2 mm below the gas pocket being completely negligible, those events are mostly induced by radioactive background from the top optical plane.

The rate of accidental coincidence that might induce a WIMP-like event can be computed by opening a 3.5 ms window after the S1 signal:

$$R = (R_{ESR} + R_{TPB}) \times R_{S2} \times 3.5ms = 0.18 \pm 0.05\mu\text{Hz} \quad (2.8)$$

Over a 10 year exposure, it corresponds to 57 ± 16 WIMP-like events. It is therefore a dangerous background that deserves precaution. This rate will be strongly reduced by cuts applied during the analysis such as energy cut, diffusion estimation to reject top events from S2, S1/S2 ratio, and position reconstruction on S1 to reject events from the walls. The role of the reconstruction, to which my contribution is later discussed, is therefore crucial to suppress this background.

2.2.5 Vetoes

DarkSide-20k has an onion-like structure as the core detector is surrounded by multiple layers of shielding and active vetoes, starting with the inner veto (figure 2.11). The inner veto is contained between the stainless steel vessel

and the walls of the TPC. It is filled with 32 tons of underground liquid argon. The inner veto will act as an active veto to tag emissions from the walls of the TPC and γ cascades from neutron capture on gadolinium. It also ensures that the TPC is not in contact with the stainless steel vessel, the latter having high γ activity.

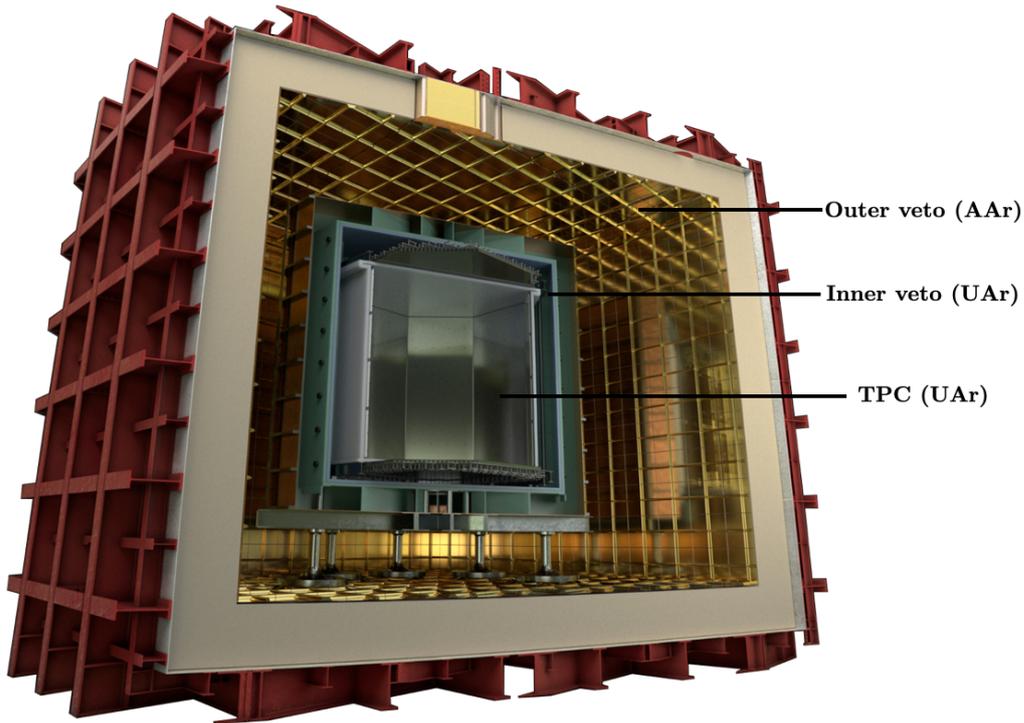


Figure 2.11: DarkSide-20k vetoes and cryostat.

In the inner veto, photo-detection modules will be located on the outside of the TPC walls, both on the walls and on the top and bottom parts as shown in figure 2.12. The modules in the vetoes will be the same as those of the TPC up to the amplifiers. The optical coverage however will be much lower (3.2% against the 97% of the TPC optical planes) as the tagging of γ events does not require high resolution for energy or position reconstruction. This will ensure a light yield of the order of 2 PE/keV.

The outer veto is defined as the volume between the stainless steel vessel and the cryostat, the last shell of the full detector. It consists in a very large volume of approximately 700 tons of atmospheric argon. Its main role is to tag cosmic muon which induce events with energy higher than 50 MeV. ^{39}Ar

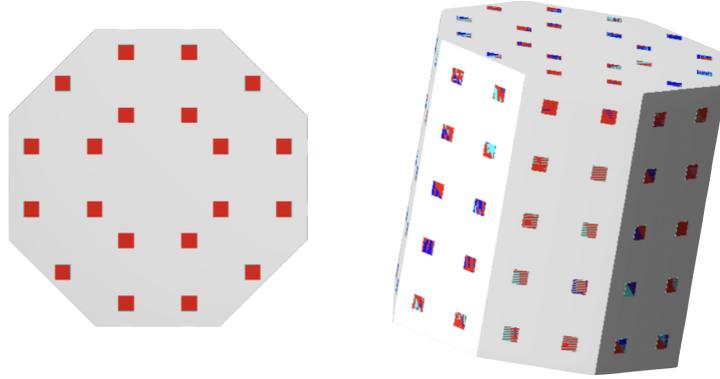


Figure 2.12: Disposal of the photo-detection modules for the inner veto, on the outside of the TPC.

decay are way below this threshold and there would therefore be no use of underground argon in this volume. Some optical modules will be detecting lights in the outer veto with very low light-yield. The light collection system is not yet definitive but the minimum requirement is 24 modules in order to reach a 1 PE/MeV light yield.

The external shell is the cryostat of DarkSide-20k, which is based on the one used in the ProtoDUNE project. ProtoDUNE is a prototype set at CERN by the DUNE collaboration. They filled a cryostat with liquefied argon to test the vertical drift of the future DUNE liquid argon neutrino experiment [54]. This is therefore a proven technology for cryogenics and structural constraints. It consists in a cubic structure of 10 m side made of a stainless steel membrane and insulating foam. This foam, crucial for thermal requirements, is however a strong neutron emitter and must receive high caution.

2.2.6 Photo-detection system

The SiPMs at the base of the DarkSide-20k photo-detection modules were developed thanks to a joint effort between the DarkSide collaboration and the Fondazione Bruno Kessler, a private research institute based in Italy, focusing on many fields including sensors and photonics. Their production is carried out by the industrial partner LFoundry. Recently, an ISO-6 dedicated clean room was installed at LNGS for the testing and assembly of the SiPMs (figure 2.13). It also has foreseen room for the assembly of bigger elements of the inner detector.

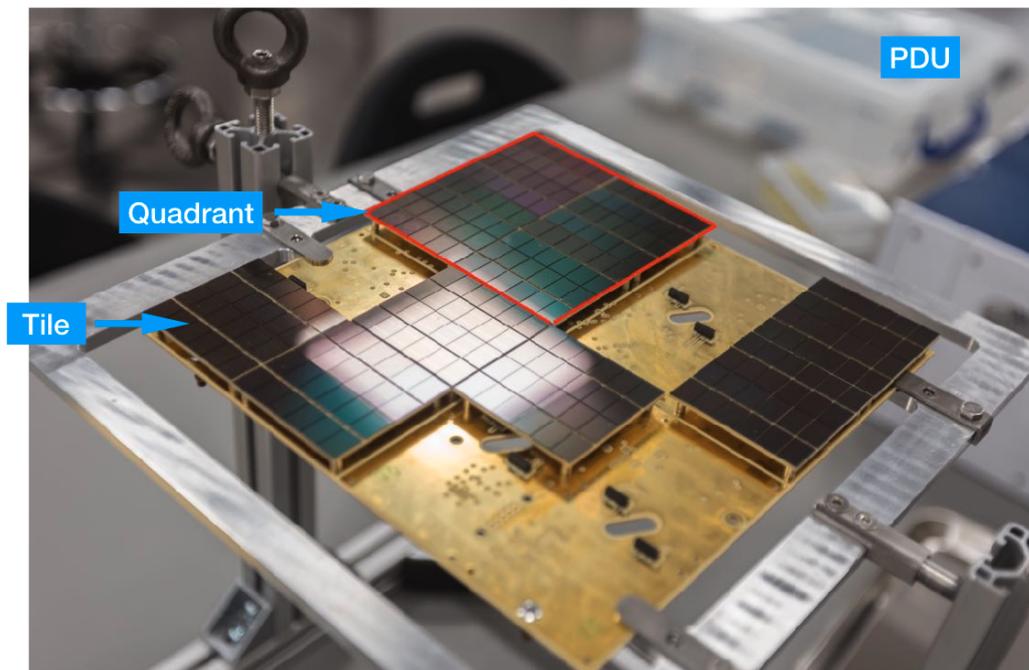


Figure 2.13: Mounting of a PDU in LNGS clean room.

SiPMs represent one of the most challenging aspect of the construction of the detector as DarkSide-20k will be the first large scale detector based on this technology. Previous liquid argon experiments, such as DarkSide-50 or DEAP-3600, a single phase detector located in SNOLAB in Canada and taking data since 2018, were based on PMTs [55]. The challenge with SiPMs mostly lies in building large surfaces of sensors while dealing with their high dark count rate. The advantage is that they can reach very low radioactivity which is impossible with the bulkiness of PMTs. The latter are still used in liquid xenon experiments which operates at higher temperature (165 K against 87 K for liquid argon). Since temperature strongly affects SiPM performances, they are less suitable in xenon detectors.

A SiPM is a combination of avalanche photodiodes (APD) and quenching resistances in parallel (figure 2.14). An APD is a semiconductor device that converts light into electric charge by photoelectric effect. The SiPM is placed under a biasing voltage. This voltage has to be several volts higher than the breakdown voltage of the APD, namely the voltage at which it becomes conductor. The difference between the biasing and the breakdown voltages is

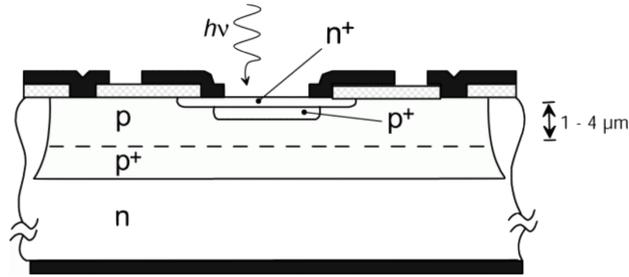


Figure 2.14: Working principle of the p-n junction of an APD. p and n regions are populated with free holes and free electrons respectively. The p^+ and n^+ area on the top indicates the active region in which the photo-electric effect occurs. The p region around is the one in which the avalanche occurs as the electric field is the strongest.

called the over-voltage. Because of the over-voltage, when a photon is absorbed by an APD, the charge carrier can trigger an avalanche, inducing a huge amount of extra carriers, making the current signal detectable.

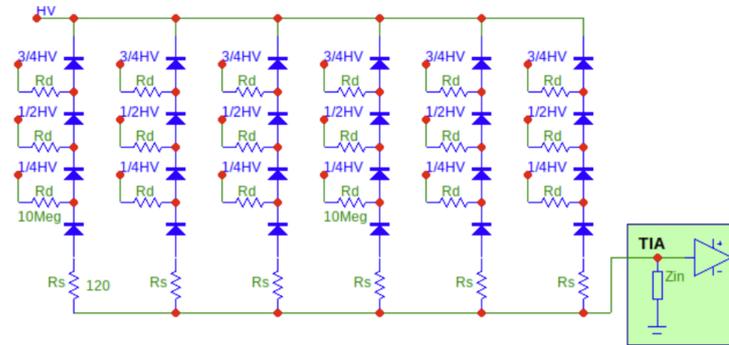


Figure 2.15: Equivalent circuit of the 24 SiPM tile.

The probability for an incoming photon to trigger an avalanche is the photo-detection efficiency and is around 50% in DarkSide-20k SiPMs. The number of charge carriers produced for one photon is the gain of the SiPM, it can be of the order of 10^5 and is linear with the over-voltage. After the avalanche, the voltage across the APD is null and increases exponentially back to the biasing voltage with a characteristic time being the product of the p-n junction capacitance and the quenching resistance.

SiPMs reach very high efficiency to single photo-electron detection. This is primordial for S1 detection as the mean number of scintillation photons

expected per channel is lower than 1 in the WIMP region of interest. As a counterpart SiPMs present a rather high dark count rate of the order of 40 Hz for $10 \times 10 \text{ cm}^2$. Due to thermal fluctuation, avalanches can be randomly triggered in the absence light. The dark count rate is the Poisson rate at which this occurs. It depends on the temperature and on the over-voltage. Another source of noise of SiPM is the afterpulse. During the avalanche, some electric charge can get stuck in an impurity of the material and be released some time after, inducing a second avalanche.

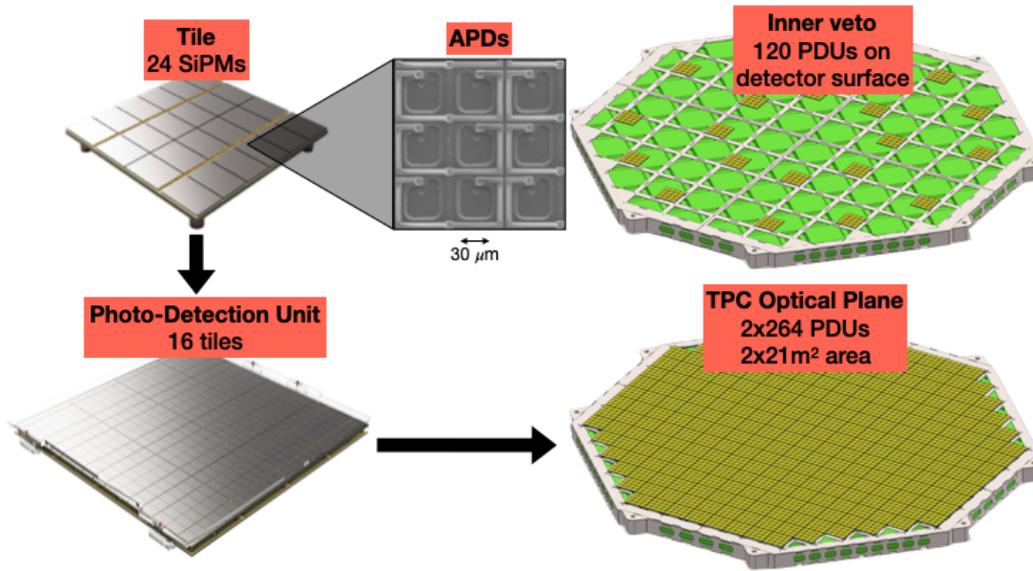


Figure 2.16: From SiPM APDs to the full DarkSide-20k optical planes.

24 SiPMs are regrouped in one tile of $5 \times 5 \text{ cm}^2$ [56]. The tile is coupled to the front-end board (FEB) which is a cryogenic transimpedance amplifier (TIA) (figure 2.15). The TIA is an operational amplifier used to convert the current of the SiPM cascade into voltage. The tile and the FEB form the photo-detection module (PDM). 16 tiles are grouped together forming one photo-detection unit (PDU) (figure 2.16). Within the PDU, groups of 4 channels are analogically summed into 1 readout channel, reducing by a factor 4 the total number of readout channels, then one PDU is 4 channels. The motivation and validation of the channel summing are detailed in section 4.2.1. In total, there will be 8338 and 1920 tiles, corresponding to 2112 and 480 readout channels in the TPC and inner veto respectively.

Chapter 3

Electronic recoil background simulation

The simulation tools of DarkSide-20k are crucial ingredients to validate the design and to develop the reconstruction and analysis strategies. This chapter describes the framework for simulation of particle interactions and light propagation developed by the collaboration. My contributions include works on a dedicated software to mix physical events in realistic time slices and on effective approaches for fast simulation of the light detection.

3.1 Tracking and optical simulation

The DarkSide-20k detector is modelled in the Geant4-based simulation package called G4DS. The software embeds also the description of DarkSide-10, DarkSide-50, and the test setups of the DarkSide collaboration [57]. It also contains custom argon response and optical models to simulate physics event up to the photo-detection. Its purpose is to test and validate the detector design, to evaluate the physics potential of the experiment, to optimise the strategy for data acquisition and reconstruction, and to define the data selection criteria. G4DS contains a custom optical model to describe the propagation of light in the detector. Additional tools are developed, with shared libraries and data format, to provide simulation and analysis tools at every level of the detector.

3.1.1 The G4DS package

In G4DS, several particle generators are available to start a tracking simulation. For instance, the primary particle can be shot at given a position and momentum, or an unstable isotope or a radioactive chain can be generated in a detector location. This location can be a generic shape, like a cube or a sphere, or a defined volume such as a detector element. Other specific generators are provided for cosmic muon or environmental neutrons.

All the properties of the detector materials and the specifically tuned physical processes are stored and used for the propagation of particles. In particular, two physics lists of interactions are used, the G4EmLivermore for electromagnetic physics and a custom made library of hadronic processes for neutron simulations. The tracking of the particle is simulated through steps during which the particle travels freely. A step ends when an interaction occurs or when the particle enters a new material. Upon interacting, the particle deposits some energy in the medium, the information of these energy deposits are stored. These information contain the nature of the interacting particle and the position, time, and energy of the deposit.

Taking the energy deposits as an input, the energy partition between ionization, scintillation, and quenching is evaluated as a function of the nature of the incident particle to distinguish ER from NR. The response of liquid argon to radiation is implemented in a custom-made Geant4 module which contains an effective description of the recombination probability $r(E)$. The latter was empirically parametrised as a function of the recoil kinetic energy E as:

$$r(E) = \operatorname{erf}\left(\frac{E}{p_1}\right) \times \left(p_2 \times e^{\frac{-E}{p_3}} + p_4\right) \quad (3.1)$$

erf is the error function and p_i are parameters fitted from DarkSide-50 data. This was measured for a 200 V/cm drift field and gave $p_1 = 3.77$, $p_2 = 0.277$, $p_3 = 113$, and $p_4 = 0.665$. The recombination probability as a function the energy is shown figure 3.1. The response model also includes a custom quenching function for NR.

Then, a custom optical model is used to propagate S1 photons from the interaction location. The propagation of the photons takes into account the absorbance, reflectivity, and refraction indices of the materials as a function of the photon wavelength. The absorption and isotropic re-emission from the wavelength shifter are also simulated. For S2, photons are simulated by

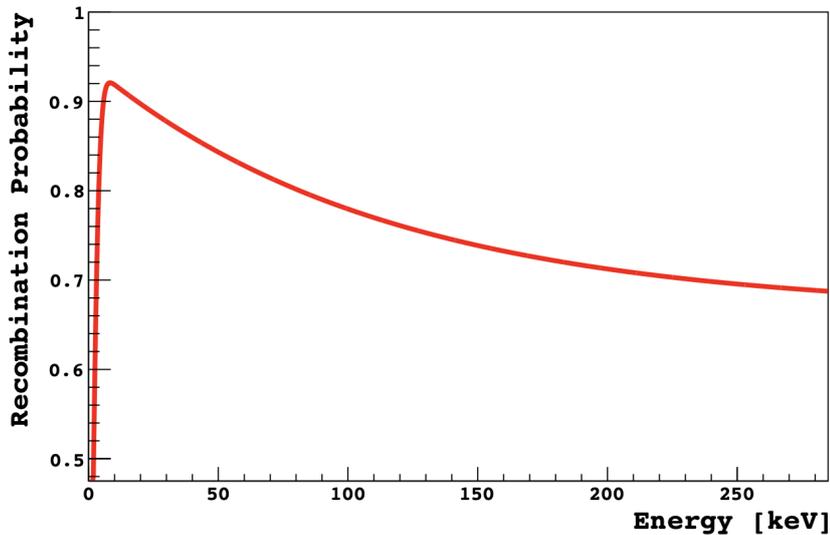


Figure 3.1: Effective recombination probability fitted on DarkSide-50 data [57].

computing the drift of the ionization electrons up to the gas pocket, applying the electro-luminescence model and diffusion smearing, and propagating the photons from the gas pocket.

Photo-detection units (PDUs) are integrated to the geometry for TPC and vetoes. When a photon reaches a PDU, a random number is generated to account for quantum efficiency. If the photo-electron (PE) survives the quantum efficiency, its time and channel are stored. Those PE information are the input for the photo-electronic and DAQ simulation and can thus be obtained from any physical process.

3.1.2 Time slice simulation

As DarkSide-20k will operate in trigger-less mode, the raw data will be stored in time slices. To reproduce the acquisition chain, simulated events from radioactive backgrounds are mixed in each time slice accordingly with their associated activity. For this scope, a dedicated package, called the event mixer, was developed.

The first step of the time slice production is the generation, in G4DS, of large statistic samples of individual background processes. Each event

of these files contains one decay of a specific isotope located in a uniformly distributed position of an element of the detector. The generated locations include:

- The **PDMs** for both the inner veto and TPC.
- The **Gd-loaded PMMA** scintillator on the lateral walls and in the optical planes.
- The **stainless steel vessel**.
- The **structural support for the grid** that will control the high voltage in the gas pocket.
- The **steel structures** on top and bottom of the TPC.
- The anode (on bottom) and cathode (on top) **PMMA windows**.
- The **reflective layer** placed on the lateral walls of the TPC 4 cm away from the walls. It regroups the 4 mm thick acrylic layer, the ESR and the TPB coating.

10^8 events are simulated for each location and for each isotope. Among these generated events, only a fraction will lead to at least one energy deposit in the TPC or in the inner veto, the other events will not induce detectable lights and are therefore not written on disk. Deposits in the outer veto are not stored as the following studies are only centred on the inner detector so far. The fraction of written f over generated events is stored as it is required by the event mixer for normalising the final rate.

The activities of all the detector materials are assayed by different groups of the collaboration (table 3.1). The assay works are performed in 4 underground laboratories and in 5 other facilities, combining several techniques. One is the use of high-purity germanium detector screening to measure with high resolution the γ spectrum of a sample [58]. For ^{210}Po , the activity is measured by radiochemical extraction [59]. Another method is the Inductively Coupled Plasma Mass Spectrometry (ICP-MS). The ICP-MS consists in ionising a sample of the material and detect the so-formed ions to identify the isotopic composition [60]. The complementarity of the techniques allows to have access to all isotope activities and to validate the results with independent measurements.

The measured activities of the cryostat foam are very high in comparison to the other materials (1.2×10^6 Bq in total). Since it is far from the centre

of the detector, the probability for a gamma emission to reach the TPC is very low. This is why there is no need for high radio-purity in the cryostat. Moreover, its gamma activities are low enough to exclude the cryostat gamma activity from the simulation.

	Activity [Bq]							
	Up ²³⁸ U	Mid ²³⁸ U	Low ²³⁸ U	²³² Th	²³⁵ U	¹³⁷ Cs	⁶⁰ Co	⁴⁰ K
PDM	27	12	360	8.3	1.2	1.1	1.2	220
Veto PDM	5.6	2.9	83	2.3	0.32	0.28	0.27	36
Plastics scintillator	12	11	12	4.0	7.6	2.4	1.3	100
Stainless steel vessel	21	6.6	13	35	0.96	4.9	10	46
Grid support	0.70	0.51	0.70	0.58	0.032	0.60	2.2	0.34
Steel structure	5.5	3.0	5.5	6.6	0.25	0.96	1.5	13
Anode and cathode	3.6	0.099	2.4	0.15	0.16	0.039	0.44	0.81
Lateral reflector and acrylic	0.53	0.41	2.0	0.10	0.028	0.076	0.22	1.5

Table 3.1: Background activity table of DarkSide-20k materials provided by the radioactive assays.

In addition to the material decays, ³⁹Ar is simulated in the TPC and in the inner veto. Its activity is extrapolated from DarkSide-50 and is expected to be 0.73 mBq/kg. It is therefore a major contribution to the detector rate (table 3.2). At this stage, only energy deposits are written, the light simulation is saved for the next step.

	TPC activity (Hz)	Inner Veto activity (Hz)
Material rate	44	146
³⁹ Ar	36	26
Total	80	172

Table 3.2: Expected rate in TPC and inner veto consider efficiencies of each process.

The activities and simulation of the ²³⁸U chain activity are separated in three (up, mid and low) because the secular equilibrium is broken twice for this decay chain. It is firstly broken by the ²²²Rn which is very volatile and can therefore enter or escape the material. As ²²²Rn is not exclusively produced by the decay of its parent, there cannot be any strict relation between their activities. It is broken a second time at the ²¹⁰Pb level which half-life is 22.2 years, longer than the experiment duration and any other isotopes in the chain. Thus the activities of its daughters are determined by its initial abundance and not by the activity of the rest of the chain. The three sub-

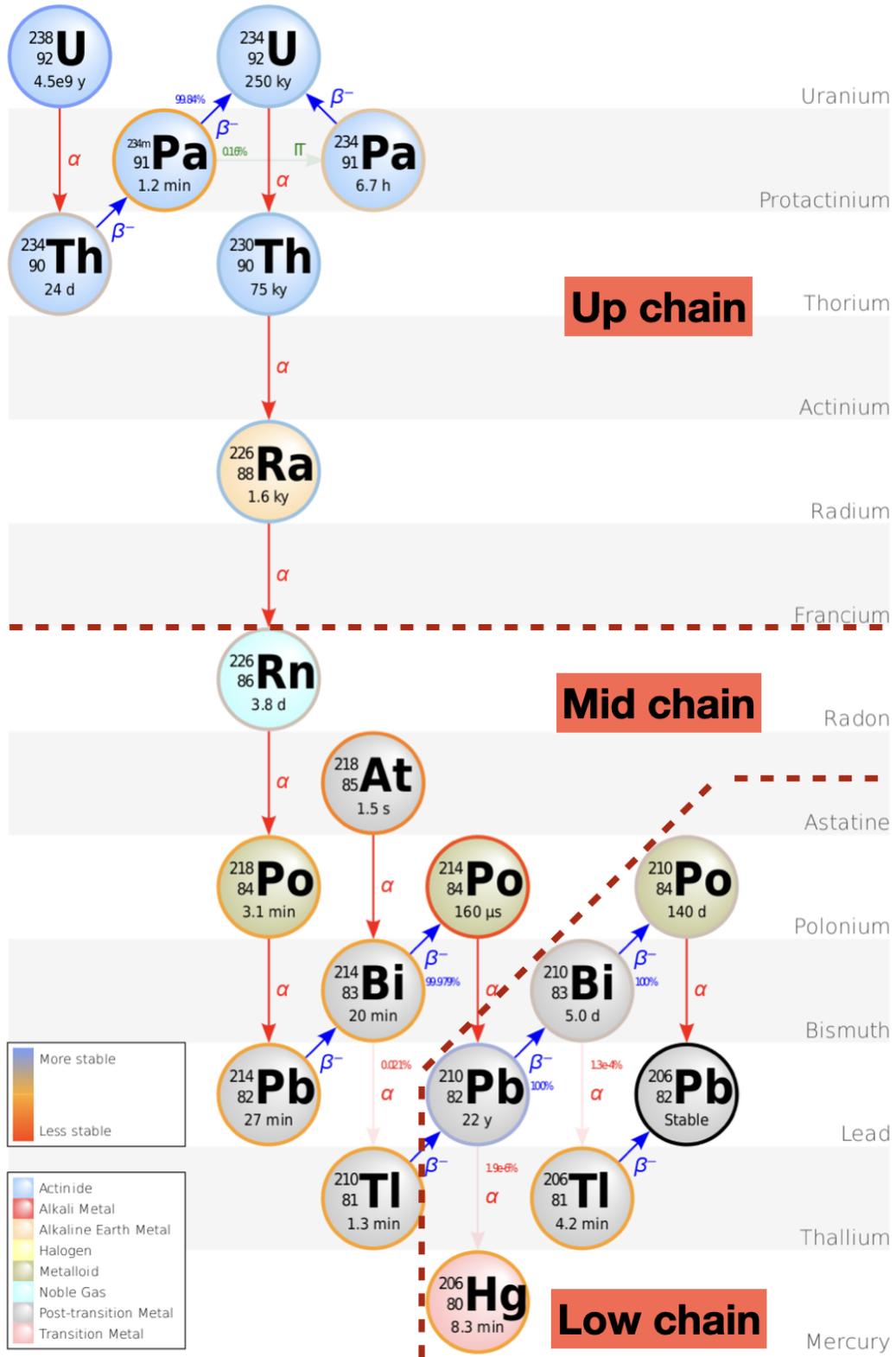


Figure 3.2: Full ^{238}U decay chain. Dashed lines represents the two breaks of the secular equilibrium before ^{222}Rn and ^{210}Pb .

chains of ^{238}U are indicated on figure 3.2.

The event mixer produces time slices with the same format that is used in G4DS. It is a custom binary format that can be converted to a ROOT file for convenience. This file is built by opening time windows and populating them with every background process following poisson distributions. The expected number of events for each location-source couple in a time gate of duration ΔT is computed as follow:

$$\mu = A \times f \times N_{iso} \times \Delta T \quad (3.2)$$

A , f , N_{iso} are respectively for each process the activity, the fraction of written events, and the number of isotopes in the decay chain, the latter is needed to correct the number of generated events accounting for the custom implementation of decay chains. Each event is then uniformly distributed in the time window. The output of the event mixer is a collection of these gates containing energy deposits from which the light can be simulated with G4DS or with the effective approaches described later. The background spectrum obtained for DarkSide-20k is shown in figure 3.3.

Before light simulations, an extra step has to be applied on the deposits. Since the modelling of liquid argon is based on macroscopic measurements while the tracking simulation with Geant4 is a microscopic description, the energy deposits cannot be directly used as an input of the scintillation and ionization model. They first have to be clustered. In other words, close-by deposits are merged before the producing any light. Deposits are merged if they are closer than 50 cm in the xy plane, 0.2 cm in the z direction, and less than 200 ns apart.

The data rates and pile-up probabilities are dominated by gamma sources, that is why at this stage, this model does not include NR backgrounds which face different simulation constraints. Since the neutron efficiency is very low, the study of this background requires the production of a very large number of events, which requires the implementation of biasing techniques in G4DS to be reached.

This work set an important progress toward the agreements between the simulation, physics, and data acquisition requirements. These realistic data samples are used in the following to study the constrain on online data treatment, data storage, observable reconstruction, and sensitivity studies. The whole framework is frequently updated with more precise geometry or with

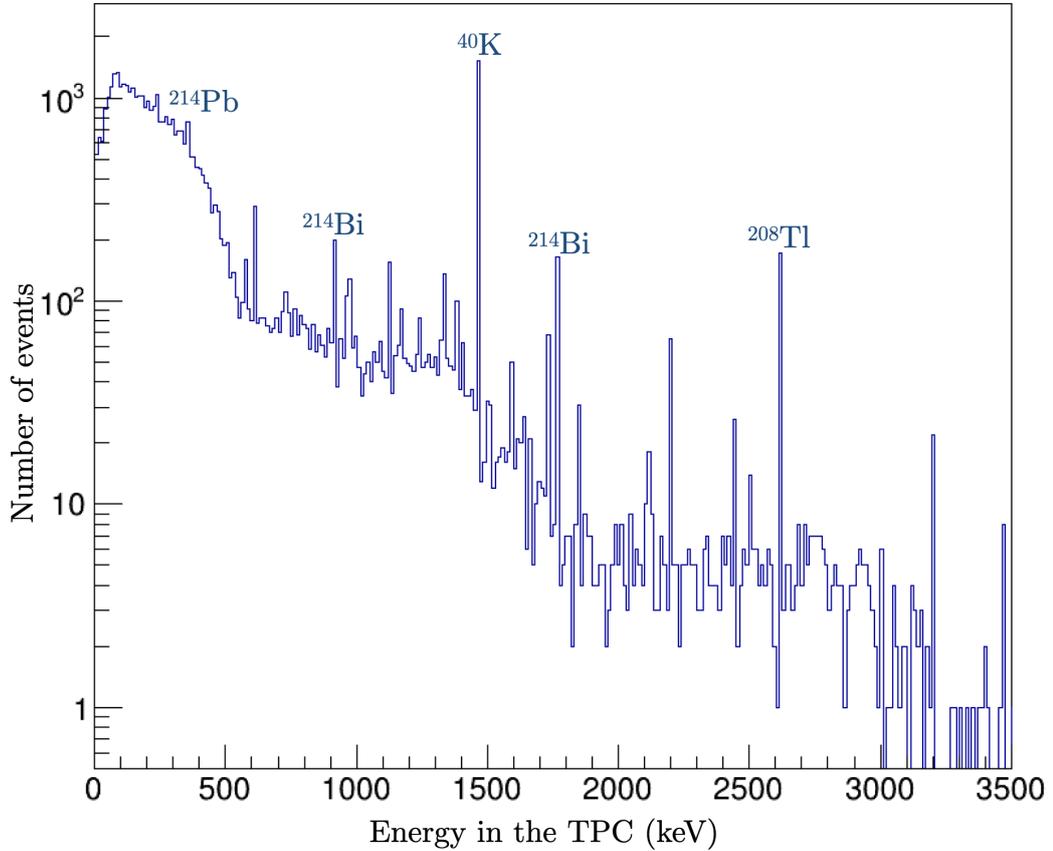


Figure 3.3: Electronic recoil spectrum of DarkSide-20k obtained with G4DS and the realistic sampling of background events. The spectrum corresponds to 500s of exposure.

improved modelling of the physics processes to provide more accurate and efficient tools for the collaboration.

3.2 Effective light simulation

Light pulse simulation is a very time-consuming process since it requires to compute the path of all single photons whose number can reach $\mathcal{O}(10^5 - 10^6)$ in an S2 pulse. For the production of high statistics samples, it is practically impossible to use the full optical model for all the pulses. Still, S1 pulses require the realistic simulation because the PSD could be strongly affected even by small imprecisions of an effective approach. However, S2 pulses and veto pulses can be effectively simulated using light maps.

3.2.1 Liquid argon response

To evaluate the number of photons and electrons that will be detected, the deposited energy is first shared among S1 and S2 as a function of the recombination probability, the latter depending on the electric field. In particular, for energy deposits above the extraction grid, the higher field induces less recombination and therefore more energy going in the ionization channel. An empirical function is used to compute the relative dependance of the recombination probability r as a function of the electric field \vec{E} :

$$r(\vec{E}) \propto \frac{p_1}{p_2} \left(\frac{\|\vec{E}\|}{1 + p_1 \|\vec{E}\|} \right) + p_3 \quad (3.3)$$

p_1 , p_2 , and p_3 are three parameters fitted from external measurements whose values are respectively 1.890, 1.2455, and 0.2051 [61]. This relation is used as a scaling with respect to the energy dependence at 200 kV/cm extracted in DarkSide-50. Then, the number of electrons is a gaussian smearing of the S2 energy divided by the quanta energy:

$$N_{e^-} \sim \mathcal{N} \left(\frac{E}{w}, F \times \sqrt{\frac{E}{w}} \right) \quad (3.4)$$

E and w are the deposited visible energy (after quenching) and the work function of 19.5 eV respectively [62]. F is a Fano factor that was measured to be 0.116 in DarkSide-50 [63]. In general, the Fano factor is defined as the ratio between the standard deviation of a random variable and the standard deviation in the case where the process follows a Poisson law. The number of photons is then computed as the total number of quanta minus the number of electrons in order to ensure energy conservation.

Once the number of electrons is computed, the number of photons emitted in the gas pocket for each of them follows a Poisson distribution with g_2 as mean value. In DarkSide-50, g_2 was directly measured by selecting electrons extracted from the cathode by photoelectric effect induced by UV photons either from a scintillation or an ionization pulse. This kind of events, called echoes, are efficiently selected by imposing the time coincidence with the progenitor S1 or S2 pulse equals to the maximum drift time. In DarkSide-50, this provided a value of $g_2 = 23.1 \pm 0.3pe/e^-$ [51]. The fit of g_2 is shown on figure 3.4.

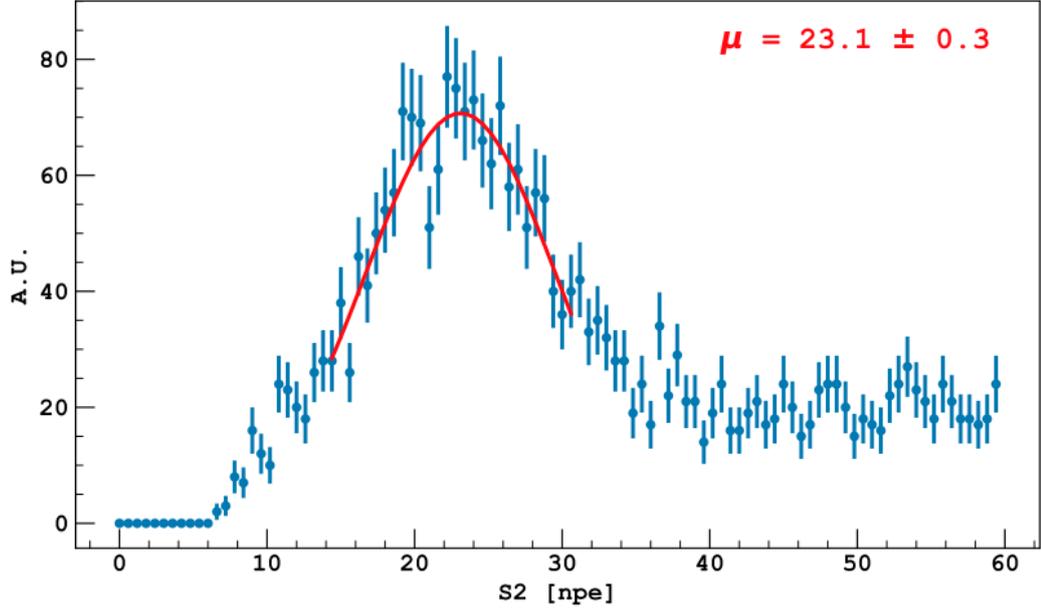


Figure 3.4: Gaussian fit of single electrons extracted from echoes in DarkSide-50.

The quantitative understanding of liquid argon relies on a measurement from 1974 of the average energy $\langle w_i \rangle$ required to form an ion pair. The definition of $\langle w_i \rangle$ arises from the balance equation for noble gases:

$$T_0 = N_{ex}E_{ex} + N_i(E_i + \epsilon_i) \quad (3.5)$$

Here, the energy absorbed in the medium T_0 has three contributions: the energy expended by ionization E_i , the energy expended by excitation E_{ex} , and the kinetic energy of the delivered electrons ϵ_i [64]. This allows to define $\langle w_i \rangle = T_0/N_i$ which is different from the previously defined work function w . Since w is the average required energy to produce either an ion or an exciton while $\langle w_i \rangle$ is the average energy to produce an ion, the two are related to the excitation-to-ionization ratio N_{ex}/N_i as follow:

$$w = \frac{\langle w_i \rangle}{1 + N_{ex}/N_i} \quad (3.6)$$

The 1974 measurement was performed under the strong hypothesis that the band structures in solid argon also exist in liquid states. Moreover it relies on a measurement of the excitation-to-ionization ratio $N_{ex}/N_i = 0.26$ performed by the DESY and only reported in an unpublished internal note. Later, the value of this parameter appeared to be inconsistent with some

other publications using $N_{ex}/N_i = 0.21$ [65].

In conclusion, it seems that the framework on which is built the understanding of liquid argon response remains questionable. The main caveat is the absence of a comprehensive description valid at both low and high energy. The possibility to develop a semi-empirical model avoiding those discrepancies was preliminarily studied and could in the future provide a description of the response and partition between scintillation and ionization that does not rely on the recombination probability.

3.2.2 Light maps for S2 pulses

The S2 light map associates, for a given 2D position in the gas pocket, a probability distribution across the photo-detection channels. The light map is built by simulating, with the full optical model in G4DS, photons isotropically and homogeneously distributed in 10 cm side voxels then measuring the channel distribution probability. This dimension is a compromise between the fidelity of the result and the time required for the generation. A second map was produced with 5 cm side voxels for slightly finer results. Once produced, the light map allows to sample PE channel given the x and y coordinates and the number of PE of the interaction as shown in figure 3.5.

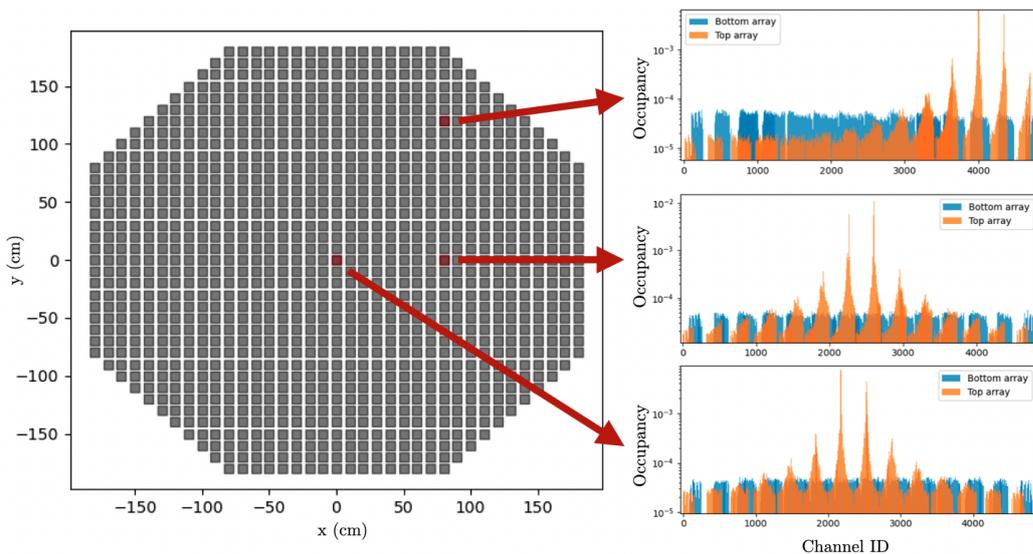


Figure 3.5: Channel probability distribution at three different positions.

The drift time is then computed as a function of the depth of the deposit. It has to be taken into account that the drift velocity depends on the electric field, which is higher above the extraction grid than in the rest of the TPC. In the TPC, the field is 200 V/cm implying a drift velocity of 0.93 mm/ μ s while it increases to 3.5 kV/cm above the grid, for a velocity around 3 mm/ μ s. Some smearing is added to the drift time of each electron in order to emulate the diffusion effect. As the latter is analog to a random walk, the smearing follows a gaussian distribution whose standard deviation scales as the square root of the interaction depth (figure 3.6). The drift time in the gas is also considered and modelled as a uniform distribution in the whole gas pocket.

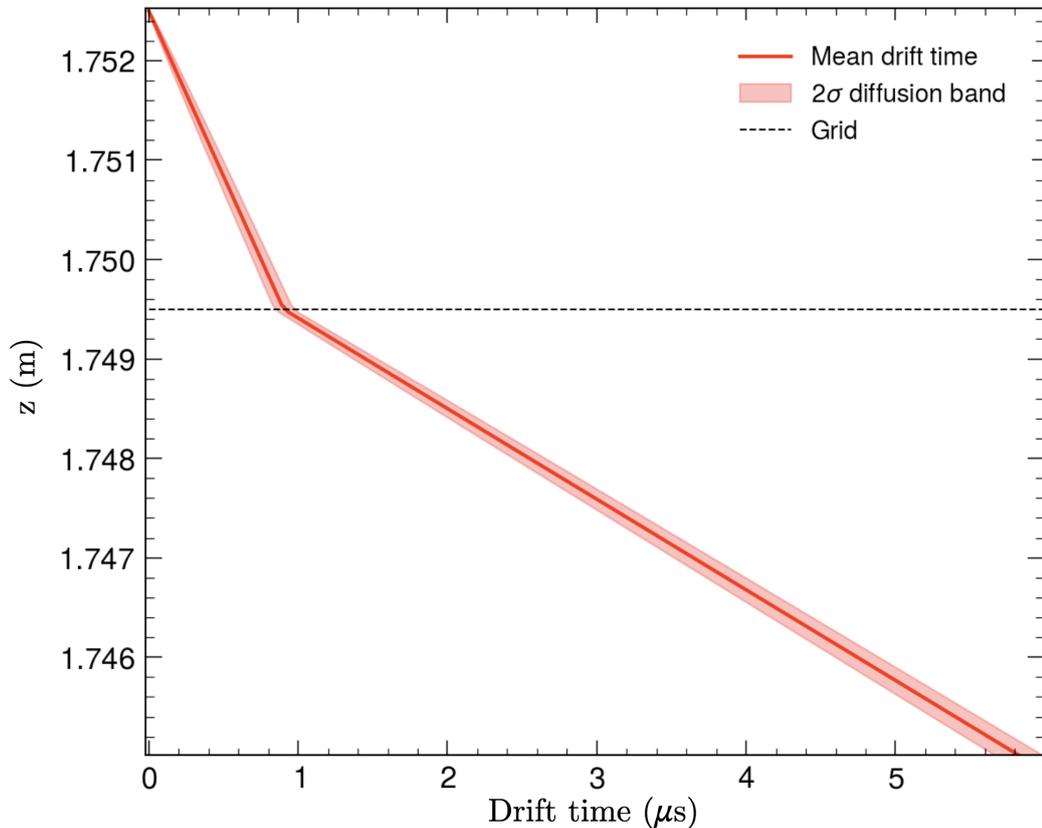


Figure 3.6: Electron drift time as a function of the z -coordinate of the deposit, which is 0 at the centre of the TPC. The dashed line indicates the grid above which the electric field is stronger. The red area indicates the fluctuation of the drift time due to the diffusion effect.

The photon time-of-flights are generated following only two distributions,

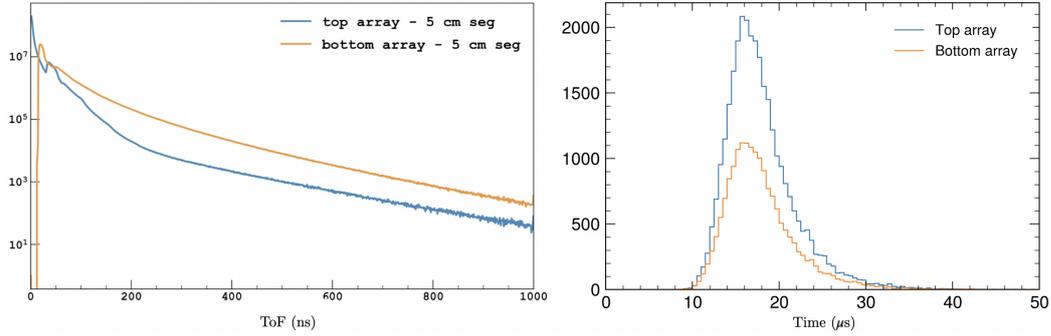


Figure 3.7: Time-of-flight distribution for top and bottom optical plane (left) and example of an S2 pulse generated with light maps (right).

one for photon reaching a top PDM, and one for photon reaching a bottom PDM (figure 3.7). A full position dependent map like the channel distribution is not produced as the effect is marginal with respect to the fluctuation of the drift time. As shown in figure 3.8, this effective approach provides results very compatible with the full optics simulation and is between 50 and 100 times faster.

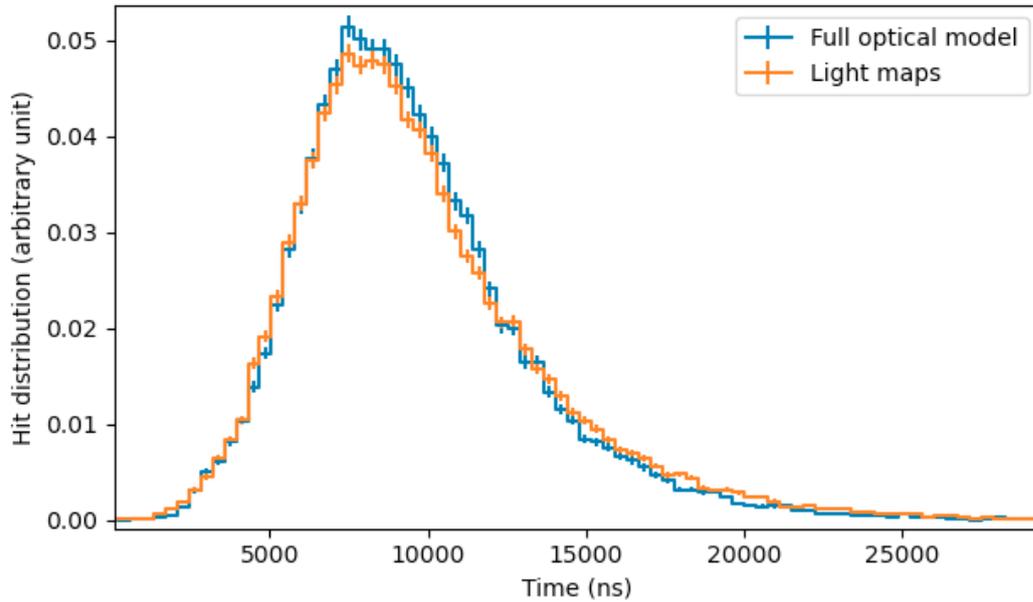


Figure 3.8: Comparison of an S2 pulse simulated with the full optical model in G4DS and with the light maps approach.

3.2.3 Light maps for inner veto pulses

In the case of the inner veto, light maps were also developed to process high statistics samples in a reasonable time. Precise energy and position reconstructions are not foreseen in the inner veto, reducing the need for highly accurate simulation.

To build the light maps, the volume is split into 20 or 30 cm³ voxels. Channel distribution functions are produced similarly to what was done in the TPC for S2. However the time-of-flight is in average higher in the inner veto than in the TPC. Because of the low optical coverage, photons are very likely to do many reflections before getting absorbed on a photo-detection module. The time-of-flight distribution is in principle different for each couple of event position and channel index, which would require for each of them one PDF. However this extra dimension would severely affect the speed and practicality of the method. An effective approach with some level of approximation is therefore needed to simulate time-of-flight following the light maps.

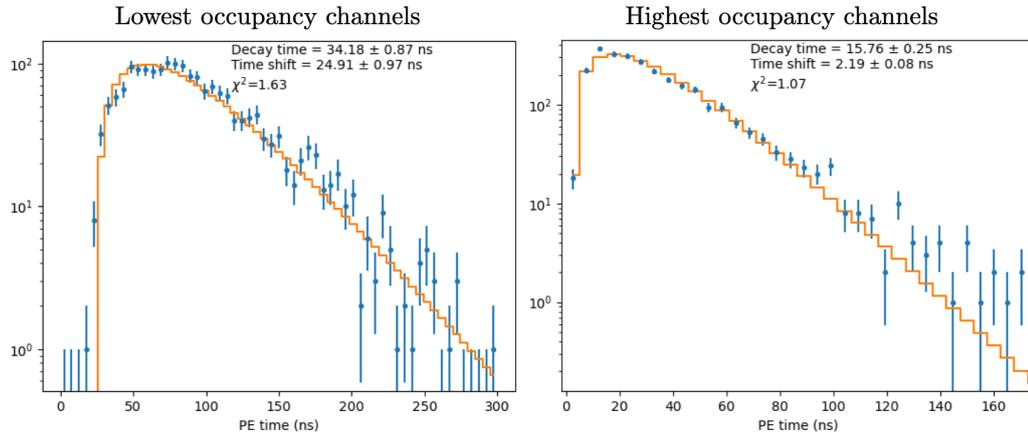


Figure 3.9: Time-of-flight photon distribution in low and high occupancy channels fitted with equation 3.7.

As a starting point, it is assumed that the occupancy, namely the probability for the photon to reach a given channel, decreases with the flight distance for the photon to reach it. Channels close to the event position will be more populated than those on the other side of the detector. Therefore, we expect the average time-of-flight to decrease with the occupancy. This hypothesis is verified with G4DS by isotropically shooting multiple photons from different locations of the inner veto. Then photo-detection channels are

regrouped by occupancy. The time-of-flight distribution is studied for each occupancy group and is fitted with the following function:

$$f(t) = \begin{cases} A(1 - e^{-\frac{(t-\delta t)}{\tau_{env}}})e^{-\frac{(t-\delta t)}{\tau_{decay}}}, & \text{if } t \geq \delta t \\ 0 & \text{otherwise} \end{cases} \quad (3.7)$$

Here, the decay time τ_{decay} and the time shift δt are the two free parameters to be fitted. τ_{env} is a rise time strongly correlated to the other two parameters, therefore it is fixed at $1 \mu\text{s}$. A is just a normalisation factor that can be expressed as:

$$A = \frac{1 + \tau_{env}/\tau_{decay}}{\tau_{decay}} \quad (3.8)$$

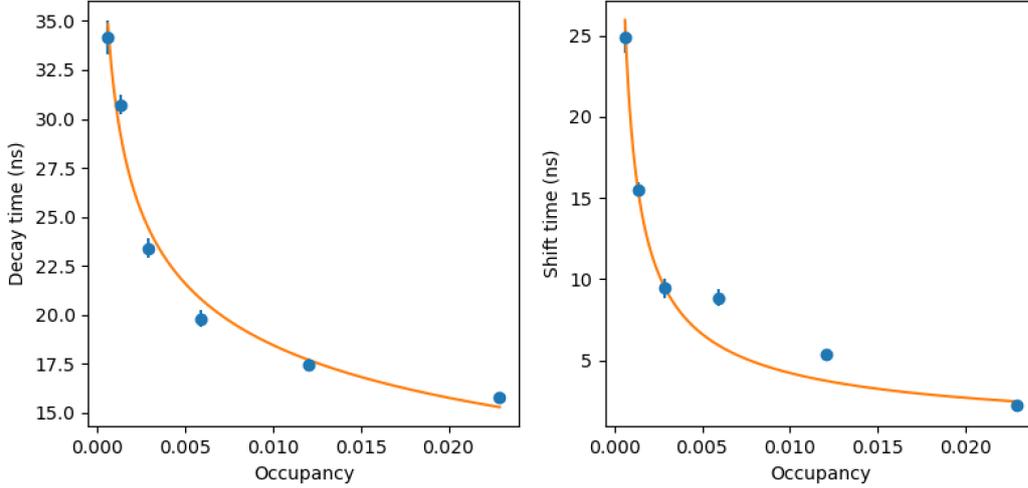


Figure 3.10: Power law fit of τ_{decay} and δt .

As expected, the characteristic times τ_{decay} and δt decrease with occupancy and the curves obtained are very similar from one position to another (figure 3.9). The aim of this study is to propose a PDF for the time-of-flight that only depends on the occupancy μ of the detection channel. τ_{decay} and δt must then be parametrised in order to be expressed as a function of μ . For different mean occupancy those two parameters were measured and fitted with power laws. Their dependency to the occupancy is shown in figure 3.10. This gave the following dependencies:

$$\begin{cases} \tau_{decay} = (2.7 \times 10^{-4} \mu)^{-0.227} \text{ns} \\ \delta t = (10.98 \mu)^{-0.651} \text{ns} \end{cases} \quad (3.9)$$

The time-of-flight can then be sampled by computing the PDF given the occupancy of the channel (figure 3.11). This approach gives results very compatible with the full optical model. The agreement between G4DS and the effective approach was assessed at different locations of the inner veto (figure 3.12). The observed remaining discrepancies do not represent an issue since the photo-electron time will in any case be dominated by the scintillation time (figure 3.13). As it can be seen, the time-of-flight implementation adds finer details to the pulse shape providing a more realistic simulation.

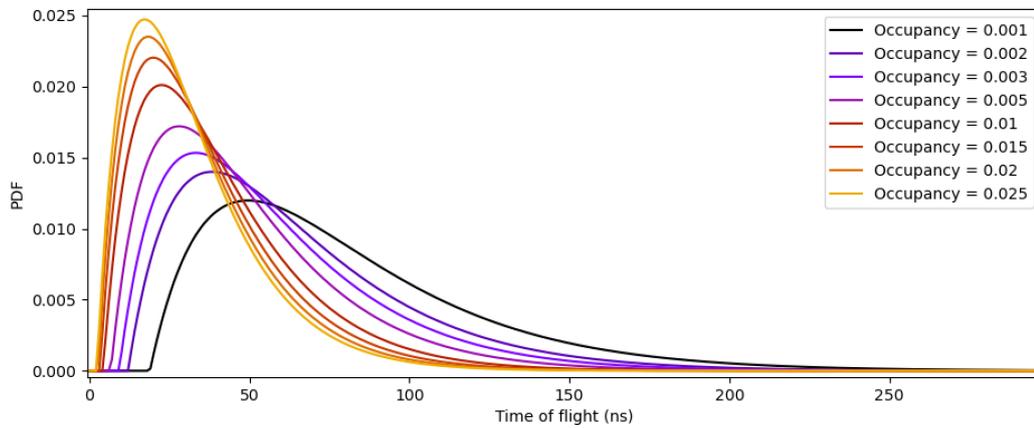


Figure 3.11: Time-of-flight PDF computed for different channel occupancies.

As an estimation of the total simulation speed, 10 ms of full background exposure were produced interactively on cluster. Over this time window, the track simulation is estimated to contribute for 0.2 ms. This first number is very approximate as the tracks are previously generated process by process and location by location, regardless of the time window or activities. Then the random combination of the background sources takes 20 ms, the clustering takes 40 ms. The S1 light generation with the full optical model takes 480 ms. The light maps generation takes 120 ms for S2 and 10 ms for the inner veto pulses.

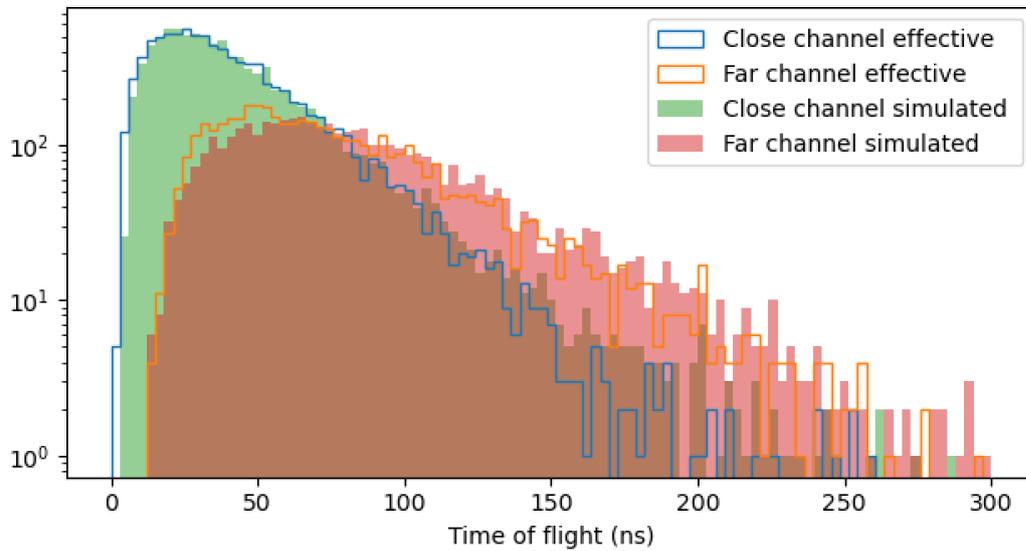


Figure 3.12: Time-of-flight distributions for high occupancy channel and low occupancy channel from full optics simulation and from effective generation. Produced in a different location as the one from which the parameters were extracted.

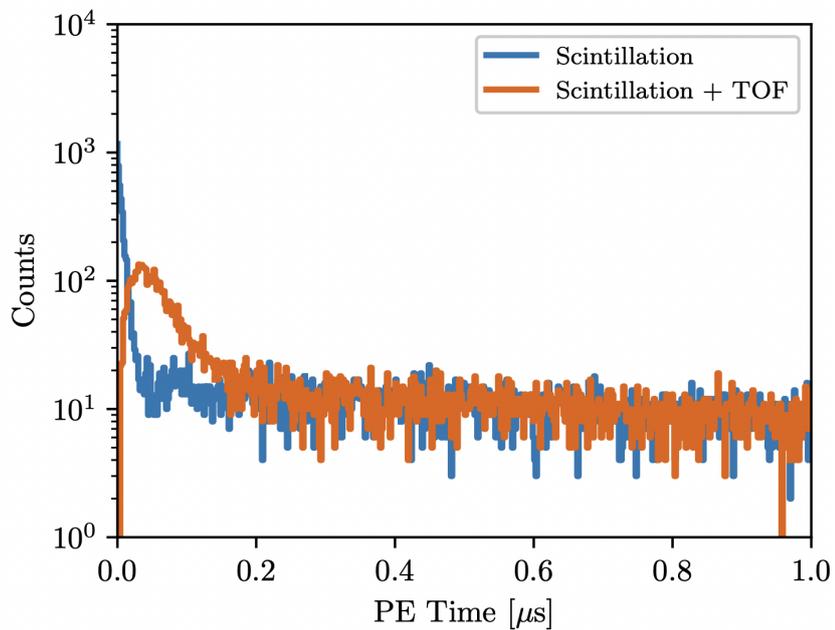


Figure 3.13: Photo-electron time distribution in the inner veto with and without time-of-flight.

Chapter 4

DAQ emulation and Event Data Model

The tracking and light simulation presented in the previous chapter is the input of the electronic and data-acquisition (DAQ) emulation software. These tools were used to develop and optimise the hit finder algorithm which is at the core of the online data-taking. Then, we defined and validated the event data model of DarkSide-20k by taking into consideration the physics requirements of the experiment and the constraints from DAQ and computing.

4.1 DAQ emulation

DarkSide-20k will operate in a trigger-less mode, meaning that all channels will continuously acquire data regardless of each other. The aim of such uninterrupted data stream is to avoid bias from some trigger decision, at the cost of extra challenges in data management. The strategy for the online treatment of the PDU waveforms must be defined in view of the full DAQ chain. This motivates the building of a software to emulate the whole data-taking, from the PE detection in time slices to the writing of raw-data on disk.

4.1.1 DarkSide-20k data acquisition

In total, DarkSide-20k counts 2720 channels, including 2112 for the TPC, 480 for the inner veto, and 128 for the outer veto. The signal from the PDU is digitised by CAEN VX2740s operating in Digital Pulse Processing mode allowing to treat independent trigger request for each channel. One of those digitisers can be connected to 64 channels and apply custom trigger condi-

tion. The sampling rate of DarkSide-20k is 125 MHz, or 1 sample every 8 ns. The constraints on data flow prevent to write on disk the waveforms from 2720 channels during 10 years, therefore another processor chain is required to process them online and store only relevant information. This is done at the level of the Front-End Processors (FEPs) which are connected to 2 digitisers each, equivalent to 128 channels (figure 4.1).

The output of the FEPs are transmitted to a network data switch which handles the data of the whole detector and split them into time slices. The duration of one slice is still under discussion, with a current baseline of 100 ms. Therefore, one slice contains the data of all channels for 100 ms.

The slices are processed online with a computing cluster, hence called Time Slice Processors (TSPs). The data switch transmits one slice to a TSP which first writes data on disk, and then apply a first online reconstruction. Beside these two tasks, the role of the TSPs will be extended to online monitoring and supernova trigger system within the SNEWS (SuperNova Early Warning System) network [66]. TSPs operate in an asynchronous way so that a TSP does not receive any new slice before it has processed the previous one. That way, if 50 TSPs are available, each of them has 50×100 ms to process one time slice. The number of TSPs that make up the farm is not yet defined and will depend on the processing to apply online.

4.1.2 Zero-length encoding

The trigger condition at the digitiser level is a time-over-threshold of 0.5 PE over 18 samples (or 144 ns). The SiPM response has an exponential tail with a characteristic time of 300 ns. Therefore, this condition ensures very high single PE efficiency. The threshold of 0.5 PE strongly relies on the online PE calibration that has to be done online and channel by channel. Once a segment is defined by the time-over-threshold, it is extended by 150 samples before and by 300 samples after.

The time-over-threshold is applied at the digitiser level through a zero-length encoding (ZLE) on the digitised waveform. The ZLE is a zero-loss compression methods which consists in replacing series of N consecutive zeros with one zero preceded by N (figure 4.2). The waveform region which does not pass the time-over-threshold is set to 0 and compressed like so, leaving only above threshold segments that we call ZLE segments. The compressed waveform segments are then sent to the FEP.

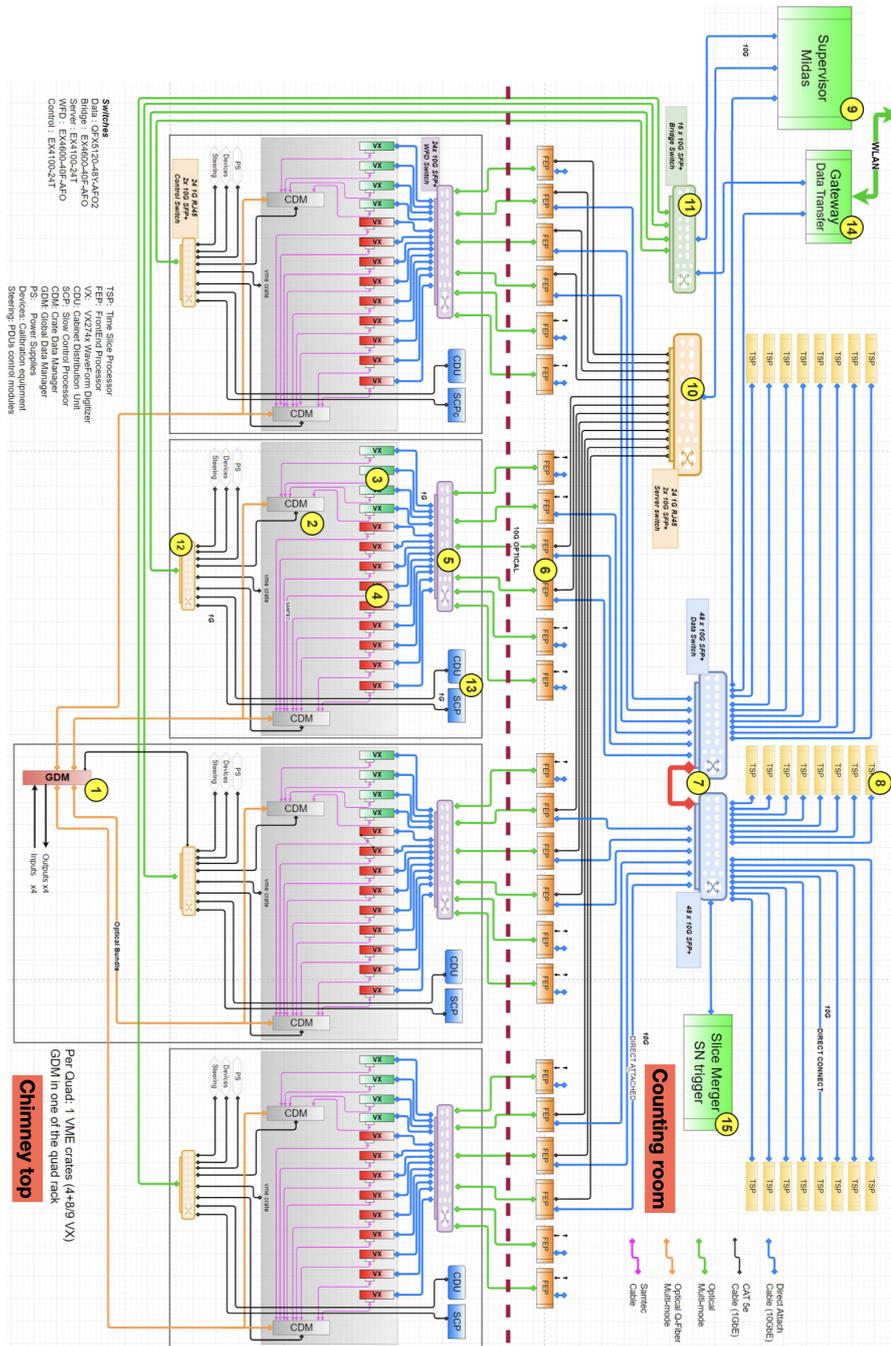


Figure 4.1: Overall DarkSide20k DAQ. The GDM (1) is connected to the CDMs (2) to distribute clock and commands to the veto (3) and TPC (4) digitisers. The digitisers are connected through a network switch (5) to the FEPs (6). The FEPs are connected to a network data switch (7) all located in the counting room away from the detector. Then, the overall data is distributed to the TSP (8) cluster. The Midas DAQ server (9) manages the data traffic from digitisers to TSPs via separated network switches (10, 11, 12). External network is provided via the gateway server (14). Room for additional processors is anticipated (15).

Raw waveform													
3	8	5	4	0	0	0	0	0	4	3	0	2	1

ZLE compressed waveform											
3	8	5	4	5	0	4	3	1	0	2	1

Figure 4.2: Example of zero-length encoding

4.1.3 DAQ emulation and fast waveform generation

The emulation of the DarkSide-20k DAQ starts with an input file produced by the event mixer containing a collection of photo-electron (PE) structures, each of them containing time and channel information. Then waveforms are generated as the superimposition of the PE SiPM response and noise around the baseline. The waveform generation is however a long process, especially with realistic noise from power spectra. To speed up the simulation, waveform segments are only generated around the physical Monte Carlo PEs of the input file. This reduces by two orders of magnitude the total waveform length to generate. Outside those chunks, channels are populated with dark noise. As the waveforms are not simulated in these regions, instead, an effective approach of the ZLE algorithm and hit finder is applied.

This was implemented in a highly optimised Python software using the Numba library. Numba is an open source compiler which automatically optimise python function to reach speed comparable with C or Fortran [67]. The full DAQ emulation process takes around 150s to simulate 1s of data interactively on a single Apple M1 core. This enables the complete simulation of the data- from the Monte Carlo energy deposits to the raw data of DarkSide-20k for significant durations.

4.2 Hit finder

The FEPs receive the ZLE segments from the digitizer and apply an important data reduction before the writing of the raw data. For this purpose, a hit finder algorithm was developed to identify SiPM avalanches from the

waveforms. It aims at measuring the times and heights of hits over the baseline fluctuations. The hit finder algorithm (HFA) was optimised to have high performance at low occupancy regime, as expected in S1 pulses, to maximise the PSD of the NR and ER.

4.2.1 Grouping and response templates

The sum of PDM signals in the photo-detection system of DarkSide-20k was motivated by the simplification of the channel read-out and the reduction of radioactivity from the cables and electronics. It implies the sum of the baseline fluctuations and therefore a reduction of the signal-to-noise ratio (SNR). Grouping the channels also reduces the capability of position reconstruction. The decision of the channel summation is then a compromise which relies on the performance of the HFA [50]. The following work demonstrated the viability of a 4 tile summing at the hit finder level.

Three grouping configurations were considered. The tile configuration or 1x1, namely without any grouping, is the one maximising the SNR. The 2x1 configuration and the 2x2 configuration consists respectively in summing 2 and 4 tiles together. As electronic noise RMS scales with the square root of the number of summed tiles, the SNR decreases with the inverse behaviour.

In order to improve the SNR, a new electronic configuration was proposed in which a compensation is added to the TIA of the tile. A capacitor is connected in parallel with a resistor, basically behaving as a low-pass filter. This feature reduces the noise at high frequency, therefore increasing the SNR. The cost is a slowing down of the SiPM response. Two configurations, one with a 68pF capacitance, the other with 33pF, were selected to be compared to the old one and were implemented in the electronic simulation (figure 4.3). Were also compared, for each configuration, two values of over-voltage, 7 VoV and 8 VoV. A higher over-voltage ensures higher SNR but also increases the dark noise.

4.2.2 Electronic noise simulation

The baseline fluctuation due to electronic noise originally implemented in the simulation software was gaussian. However this approximation is no more valid with the reduced high frequency noise induced by the compensation. The waveform generation was then upgraded to read data from a noise power

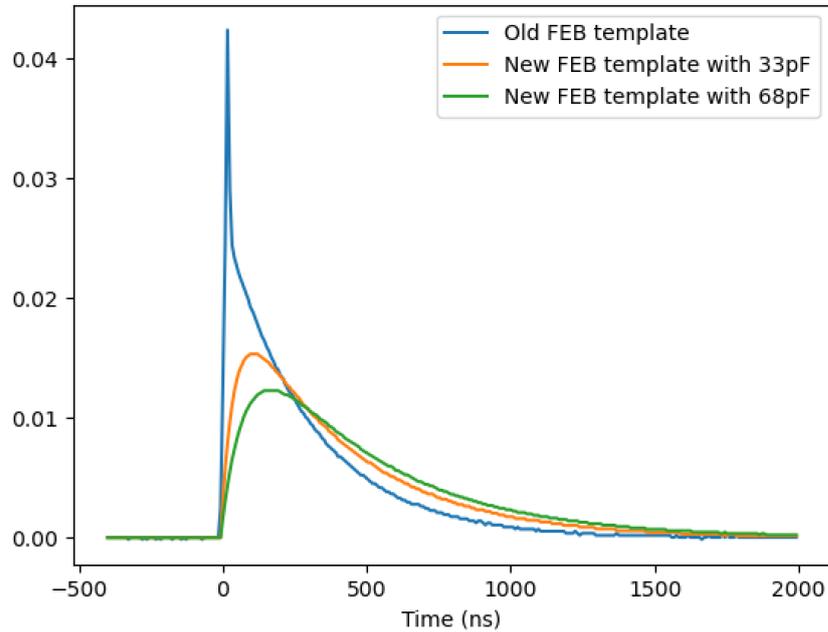


Figure 4.3: Templates for old and new FEB, with the two capacitor values, used in the DAQ emulation software.

spectrum input file (figure 4.4). This spectrum is obtained by computing the Fourier transformation on a real waveform. It gives the intensity as a function of the frequency and was measured for each FEB configuration and for each grouping scheme.

To simulate a waveform, an array of complex phases is firstly randomly generated, one for each frequency. Then the inverse Fourier transform of the power spectrum is performed and multiplied by the phases. The real part of this function gives a realistic noisy baseline. An example is shown figure 4.5. Within this method, the correlations between the frequency phases are not included. However, the impact of this approximation is considered negligible.

The simulation of realistic noise is much more time consuming than the gaussian generation and became the bottleneck of the electronic simulation chain. Therefore it required to be highly optimised in the software. This was done by producing a long baseline buffer from which segments of the size of the waveforms are selected. It appeared that realistic noise badly affects the hit finding efficiency both in terms of true hit identification and fake hit purity.

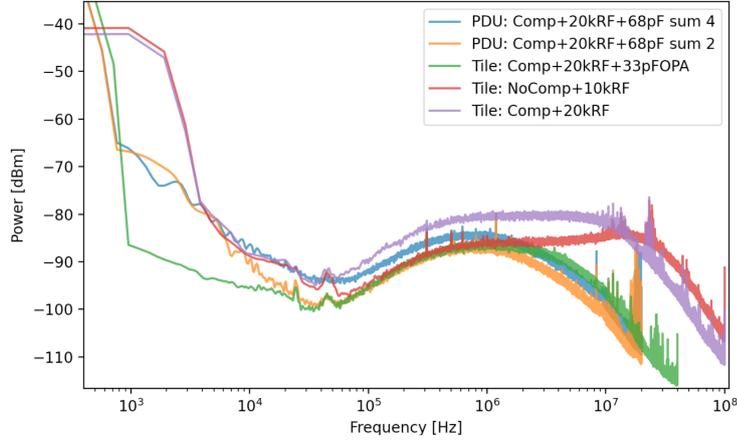


Figure 4.4: Noise power spectra for different FEB configurations

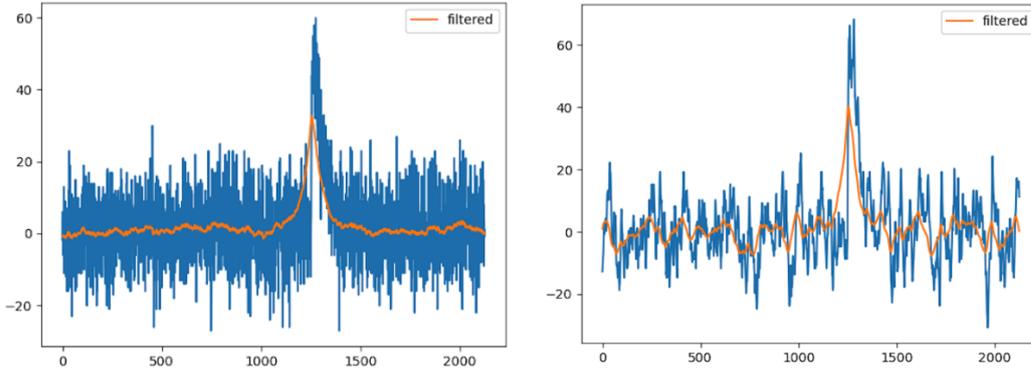


Figure 4.5: Gaussian noise waveform (left) and realistic noise waveform from power spectrum (right), both with a single hit shot.

4.2.3 Matched filtering algorithm

The hit finder is based on a matched filtering approach. To maximise the SNR, the waveform is convoluted with a time-reversed template of the single photo-electron response. The output of a hit through the filtering is a cusp signal with an optimal SNR. For instance, with the 2×2 grouping configuration, 68 pF compensation, and 7 VoV, the raw waveform SNR is 6.6 and it rises to 9.6 after the filtering.

The filtering template used as SiPM response is the sum of a gaussian and an exponential tail for the original FEB template with no compensation. The gaussian is computed through a moving average (MA) and the exponential tail is computed as an auto-recursive function (AR). Therefore we denote

AR-MA the filtering performed with this pattern.

With the compensation, it appeared that using only the exponential component (AR filter) gives compatible efficiency and reduces the computation time by 70%, which is crucial for online processing. Besides, the difference between the two filtered waveforms seems sensitively null, as shown in figure 4.6. The SNR of the filtered waveform, defined as the ratio between maximum and baseline RMS both computed after filtering, is compared between the two filtering methods. The filtered SNR is slightly increasing with the AR filter (between 3% and 5% more).

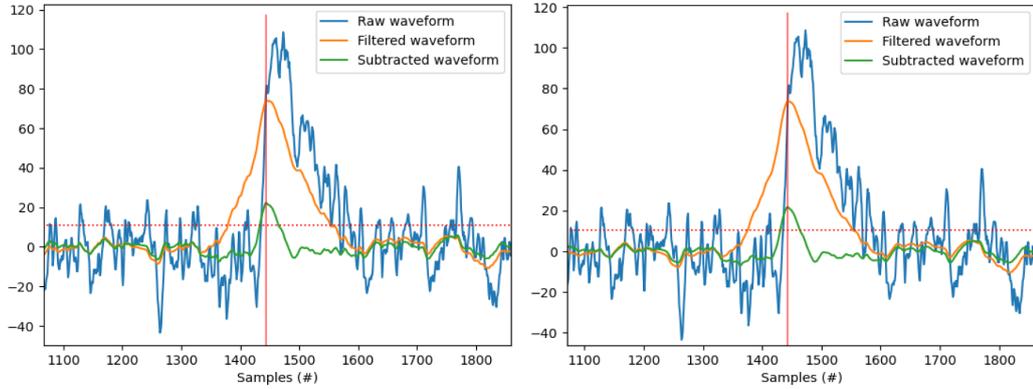


Figure 4.6: AR (left) and AR-MA (right) filtered waveforms. The horizontal dashed line is the threshold applied on the subtracted filtered waveform and the vertical line is the abscissa of the so-found hit.

The filtered waveform is then processed by the HFA. It starts by subtracting the running mean of the filtered waveform to isolate hits. Hits are then identified as segments where the subtracted filtered waveform is above a discrimination threshold. This threshold is proportional to the RMS of the filtered waveform baseline with a coefficient which is one of the parameters to be optimised.

For each hit identified, are first recorded the channel and time. The time of the hit is defined as the time of the maximum of the filtered waveform. The prominence of the hit is computed and stored. It is define as the difference between the filtered waveform and the running mean. The prominence is a measure of the height of the hit which allows to quantify efficiently the number of photo-electrons cascades occurring in the channel. In particular, if two PEs in pile-up are too close to be separately identified, the prominence

of the single found hit can be high enough to recover the total number of PEs. In addition, the charge of the hit is defined as the integral of the filtered waveform within the segment above threshold.

4.2.4 Configuration optimisation

The HFA has to be optimised to assess the viability of all different grouping configuration (1x1, 2x1, and 2x2 tiles grouping), FEB configuration (with 33 pF or 68 pF capacitance on the motherboard), and over-voltage (7 VoV or 8 VoV). The first step is to define a cut in prominence to reject fake hits which are induced by large fluctuation of the baseline. Since their prominence is in general less than the one of a single PE hit, the cut can be define as a fraction of the average prominence value of one PE hit. The quality of the cut in terms of purity and efficiency can vary with some parameters of the HFA but the optimal threshold appeared always at the same value. It is therefore fixed at 0.6 PE (figure 4.7). In the following, everything is done with the AR filtering, whose performance was shown to be compatible with the AR-MA one.

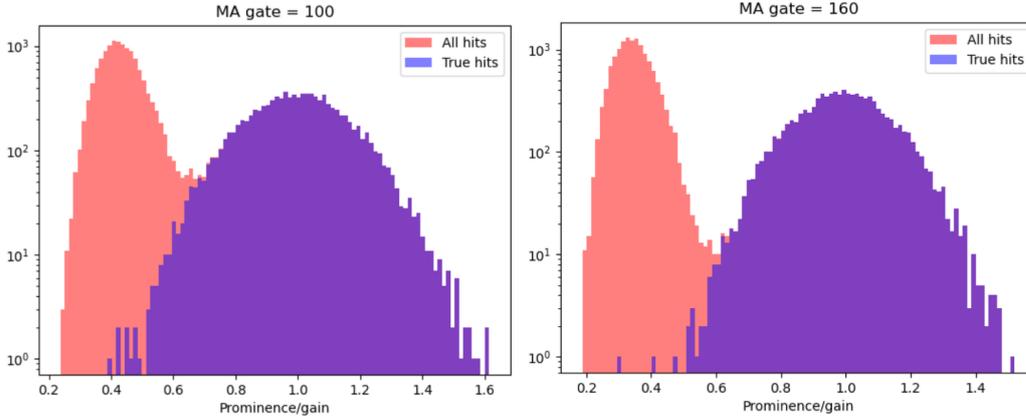


Figure 4.7: Prominence distribution of simulated hits for two different values of MA gate. The true hits are identified with time matching the true PE time.

The parameters that are the most crucial and configuration dependent are the time gate for the moving average subtraction (MA gate) and the number of RMS that defines the threshold. For each electronic configuration we perform a scan of those two parameters along which, three figures of merit are evaluated. The first one is the single hit efficiency measured with one PE

shot. It is defined as the probability to detect the PE at the expected time. The efficiency is computed over 1000 waveforms without any correlated or dark noise. Under no grouping or 2x1 grouping, the high SNR of the new FEB with capacitance ensures a 100% efficiency. It starts to decrease with the 2x2 grouping. In that case, the MA gate and RMS threshold have to be properly optimised.

To understand the rate of fake hits, waveforms are generated without any PE. They are passed to the hit finding algorithm, any hit reconstructed is therefore a fake one. After processing, it was observed that most of them were rejected due to low prominence. In particular, the fake hit rate was shown to be much smaller than the dark count rate.

The probability of a fake hit identified next to a true PE was also tested. When the filtered waveform intercepts the threshold, some fluctuations might lead to the identification of a second hit. This was measured by generating a single hit and counting the number of fake hits close to the trigger time. Those hits have very low prominence and none of them survive the prominence cut.

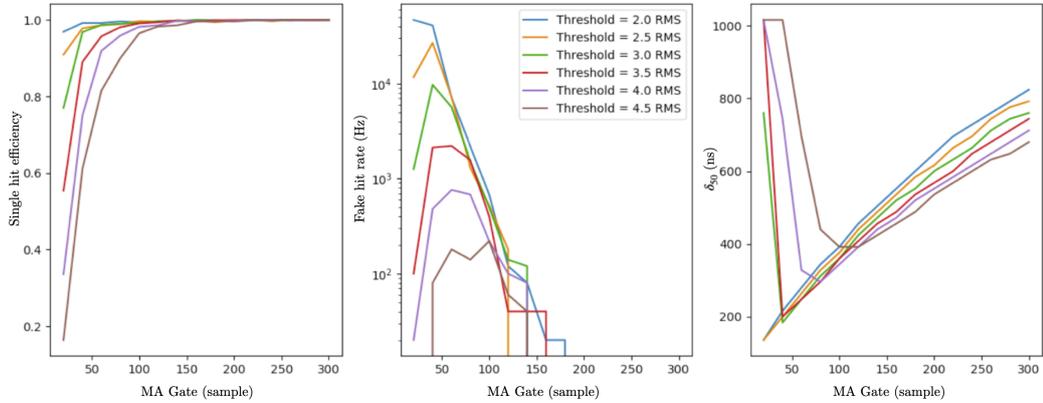


Figure 4.8: Single hit efficiency, fake hit rate, and δ_{50} evaluated along MA gate and threshold with 68 pF capacitor under 2x2 grouping configuration operating at 7 VoV.

We tested the efficiency of the algorithm on waveforms with two true PEs separated by some delay (figure 4.9). In the first place, we look at the capability of the HFA to identify both hits separately at the right times and with prominences compatible with 1 PE. At too low delay, this efficiency is 0 since the responses overlap and only one hit is reconstructed. If the hits

	SNR	MA gate	Threshold (RMS)	Single Hit Efficiency	Fake Hit Rate (Hz)	δ_{50} (ns)
8 VoV, 68 pF	8	160	4.5	0.9989	<4.6 (90% CL)	472
7 VoV, 68 pF	6.6	160	3	0.9969	<4.6 (90% CL)	520
8 VoV, 33 pF	6.5	160	3	0.9962	6	472
7 VoV, 33 pF	5.5	280	4	0.9957	2	584

Table 4.1: Optimisation results for 2x2 grouping configurations.

are sufficiently distant in time, the peak responses do not overlap and it is equivalent to detecting 2 isolated PEs. We denote δ_{50} the delay at which the two-hit efficiency is 50%. In other words, δ_{50} is the minimal delay at which the probability to properly identify two hits is more than 50%. The three observables used for the optimisation are shown on figure 4.8 for one configuration as a function of the MA gate.

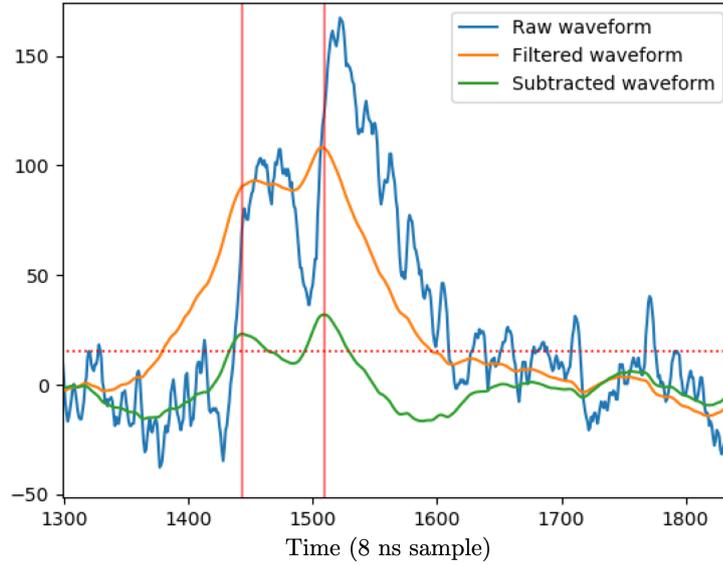


Figure 4.9: Hit finder over a waveform with two PEs separated by a 512 ns delay.

Out of all the obtained results, we select only the cases where no fake hits were observed over 100 ms and 100% of the single hits were identified in 1000 waveforms of 20 μ s. Among the remaining, the optimal set of parameters is defined as the one with the lowest δ_{50} . Many configurations are tested, so the choice criteria might be satisfied because of statistical fluctuation. Therefore, the reported results come from a second measurement done with higher statistics (all with a 2x2 grouping) (table 4.1).

Single hit efficiency is very close to 1. The fake hit rate is, for comparison, lower than the dark count rate (of the order of 10 Hz per tile at 7 VoV). These results ensure that the PDUs can safely be operated under a 2x2 grouping at 7 VoV. This is verified by checking the 2-hits efficiency curve with the 68 pF and 33 pF capacitances (figure 4.10). With the 33 pF configuration, the longer MA gate (280 samples) affects the recovery of the 2-hits efficiency at high delay. Therefore we disfavor this configuration.

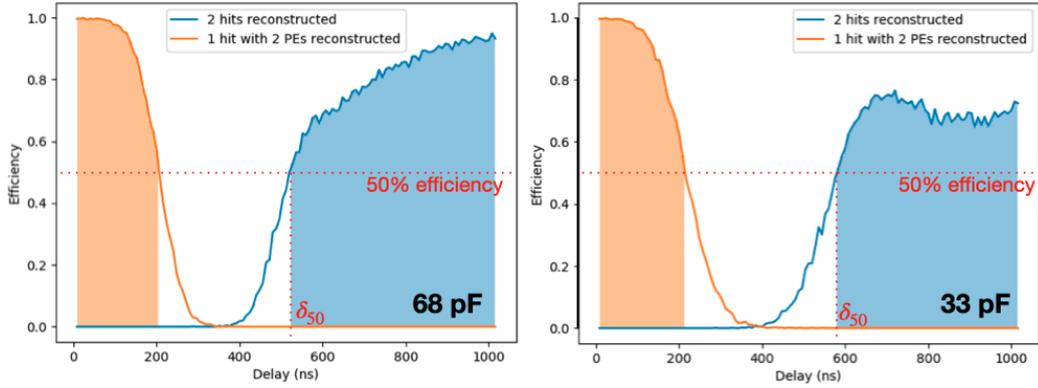


Figure 4.10: Two hits efficiency curves for 68 pF (on the left) and 33 pF (on the right) both under 2x2 grouping operating at 7 VoV.

With the obtained optimal configuration, the single photo-electron (SPE) resolution can be measured for the two amplitude observables of a hit, namely, the prominence and the charge integral. The SPE is defined as the standard deviation of the single PE peak divided by the mean value. They are measured by performing a gaussian fit over the prominence and charge distributions. With the optimal HFA parameters of the selected configuration, the SPE resolution is 14% for the prominence and 19% for the charge integral (figure 4.11).

4.2.5 Laser calibration

In order to compare the prominence and the number of detected PEs, each PDU has to be calibrated individually. The first characteristic to measure is the gain, namely the conversion factor between PEs and prominence. This gain can be derived with laser calibration which consists in exposing the SiPMs with pulsed laser in single-PE regime. The laser is brought through the detector via an optical fiber at the end of which light is diffused towards

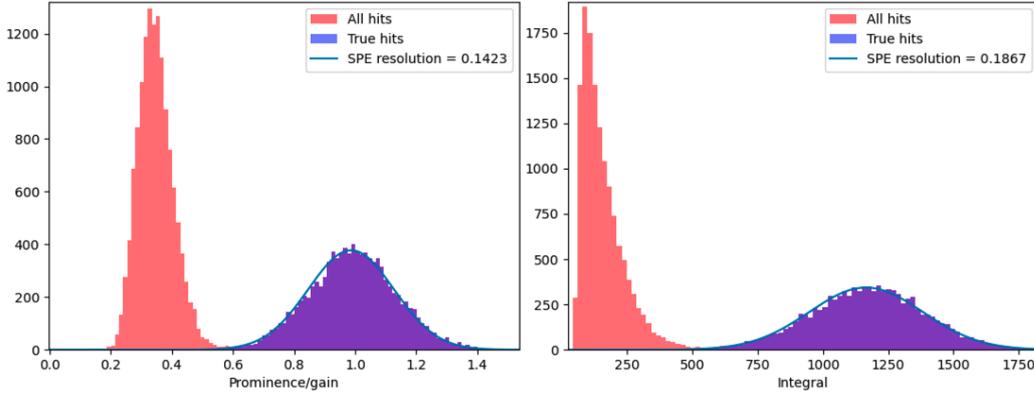


Figure 4.11: Single photo-electron resolutions of the HFA for 68 pF, 2x2 grouping, and 7 VoV for prominence (left) and charge (right).

all the PDUs. The acquisition can be triggered by the laser pulsing. In DarkSide-20k, such a laser calibration system is foreseen with multiple fibres to allow the whole optical planes to be relatively homogeneously exposed.

The prominences and the charge of the reconstructed hits will populate equally separated peaks forming the so-called finger-plots. Each peak corresponding to a number of photo-electrons detected by the SiPM (figure 4.12). Finger-plots were also produced from the charge integral around the laser trigger region. Hit based plots give better resolution, but these ones include the pedestal peak corresponding to 0 PE. A linear fit of the charge of the peaks gives a measurement of the gain (figure 4.13).

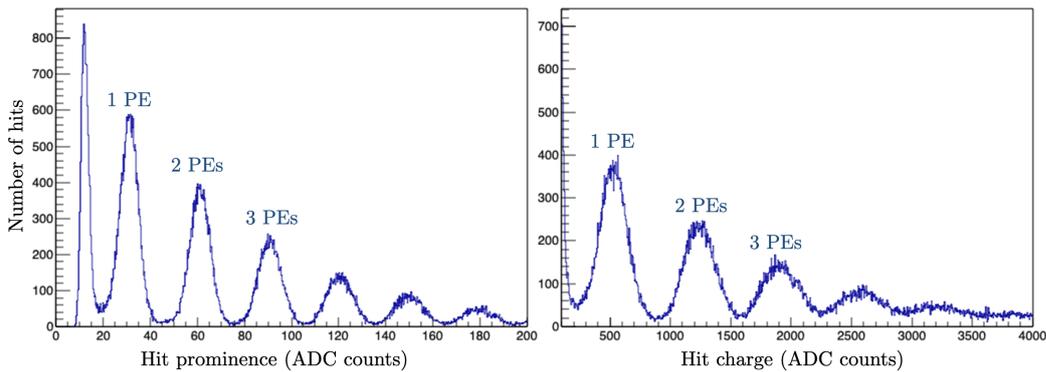


Figure 4.12: Hit prominence (left) and charge (right) finger plots from one channel of Setup-1 laser calibration. The prominence cut is not applied at this stage, hence the fake hit peaks at low prominence and charge.

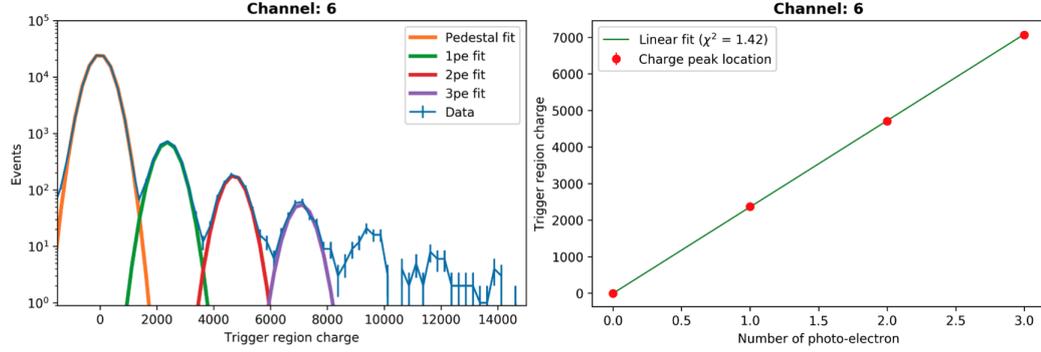


Figure 4.13: Finger-plot from the charge computed around the trigger region (left) with gaussian fit of the first 4 peaks and linear fit (right) of the mean value of the identified peaks.

The data shown come from a cryogenic test setup in Naples called Setup-1 aiming at characterising the PDUs properties at low temperature. The data are acquired in a dewar filled with liquid nitrogen where a PDU is shot with a pulsed diode laser. The motherboard contains 25 Photo Detection Modules (PDMs) which are digitised in 25 channels with a sampling frequency of 250 MHz. Each channel is composed of 24 SiPMs. The external laser trigger rate is 100 Hz and a 5000 samples (20 μ s) long window is acquired.

4.2.6 Noise characterisation

The data from the laser calibration gives information on the noise characteristics of the PDU. We distinguish two types of noise: correlated and uncorrelated noise. Among correlated noise, internal cross-talk occurs when a photon escape the SiPM during the avalanche to trigger a neighbour APD of the same channel. This process is instantaneous, given the proximity of the SiPMs. Let's keep in mind that a cross-talk avalanche can induce another cross-talk, and so on. To quantify its probability, we look at the distribution of the finger-plots. If they were no cross-talk, the distribution would be Poisson since all PEs are independent events. In reality, the distribution follows the Vinogradov model [68]. It is a deviation of the Poisson distribution with the internal cross-talk probability as extra parameter.

To establish the statistical model for cross-talk over a Poisson distribution, let's consider a true PE occupancy μ and let x be the probability for one avalanche to trigger a cross-talk. Each PE leads to a sequence of cross-talk with probability x and must end with no cross-talk with probability $1 - x$.

Then, the probability of having m cross-talks given n true PEs has to be proportional to the following quantity:

$$P(m_{CT}/n_{PE}) \propto x^m(1-x)^n \quad (4.1)$$

There are as many sequences as true PEs and they contain in total m cross-talk, each of them occurring with a probability x . However, one does not know how many cross-talk were associated to each sequence. Then, the proportionality coefficient is the number of ways to distribute m cross-talks to n PEs that is equal to $\frac{(n+m-1)!}{(n-1)!m!}$. To get the probability to have m cross-talks and n PEs, the latter probability is multiplied by the Poisson probability to have n PEs:

$$P(m_{CT}, n_{PE}) = \frac{e^{-\mu} \mu^n (n+m-1)!}{n! (n-1)!m!} x^m (1-x)^n \quad (4.2)$$

Let now $N = n + m$ be the total number of avalanches which is the final observable. It must be considered that the probability to see 10 avalanches is the sum of the probabilities to have 1 PE and 9 cross-talks, or 2 PEs and 8 cross-talks, or 3 PEs and 7 cross-talks, etc. The total probability can be written in general summing equation 4.2 for n going from 1 to N with $m = N - n$, such that:

$$P(N) = \sum_{n=1}^N \frac{\mu^n e^{-\mu}}{n!} \frac{(N-1)!}{(n-1)!(N-n)!} x^{N-n} (1-x)^n \quad (4.3)$$

This expression is the Vinogradov models and gives the probability of detecting N PEs given an occupancy μ and an internal cross-talk probability x . This model was validated with a toy Monte Carlo which samples a number of true PE from a Poisson distribution and iteratively generate cross-talks for every avalanche (figure 4.14). In this calculation, the possibility for an avalanche to induce two or more cross-talks is not considered. It is assumed that this approximation is valid as long as x is low enough.

In order to measure x without fitting the finger-plots, we can first extract μ from the population of the pedestal bin which is not affected by cross-talk since it correspond to the situation where no PE is produced. Its relative population with respect to the total is equal to $e^{-\mu}$. Then we can measure the difference between the population of the 1 PE bin and the expected population in the case of a purely Poisson statistics. The fraction of events missing is the fraction of them where cross-talk happens and are moving to higher statistic bins:

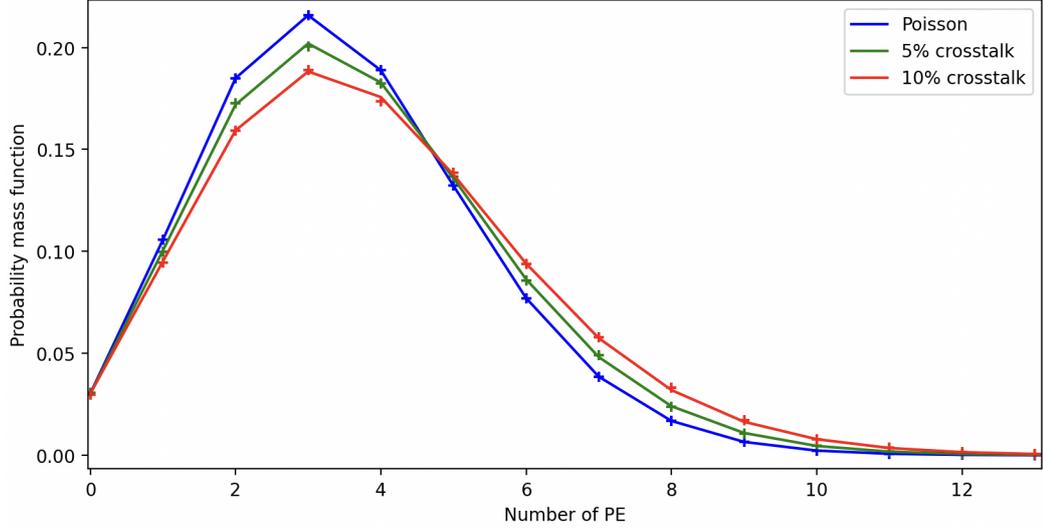


Figure 4.14: PE distribution for a cross-talk probability of 0%, 5%, and 10%. The true PE occupancy is set at 3.5. The lines are computed with the Vinogradov model and the crosses are obtained from a toy Monte Carlo simulation.

$$\frac{P_{pois}(1) - P(1)}{P_{pois}(1)} = x \quad (4.4)$$

Here, $P_{pois}(1) = \mu e^{-\mu}$ is the Poisson probability of having one photoelectron. In the Setup-1 data, the value of x for all channel fluctuates around 0.3.

In addition to internal cross-talk, SiPMs can produce external cross-talks which are based on a similar principle. The photon emitted by the avalanche escapes the SiPM and reaches another channel. The time difference between the two hits depends on the time-of-flights and the probability is not an intrinsic property but depends on the geometry of the detector. The external cross-talk rate can be estimated by looking for time correlation of dark noise. This is however hard to estimate in the simulation and is therefore not included at this stage.

The last correlated noise is the after-pulse. During the avalanche, a charge carrier can be captured by some impurity and be release only some time later, inducing a secondary avalanche. The release time follows an exponential law. The SiPM capacity is recharging at the same time and can be incomplete when the avalanche occurs. Therefore, the amplitude of the after-pulse will

be smaller the sooner it happens after the primary hit. If the release time is small enough, the after-pulse will not be identified by the hit finding algorithm (figure 4.15). The after-pulse probability was measured to be 0.05 on average with a characteristic time around $1 \mu\text{s}$.

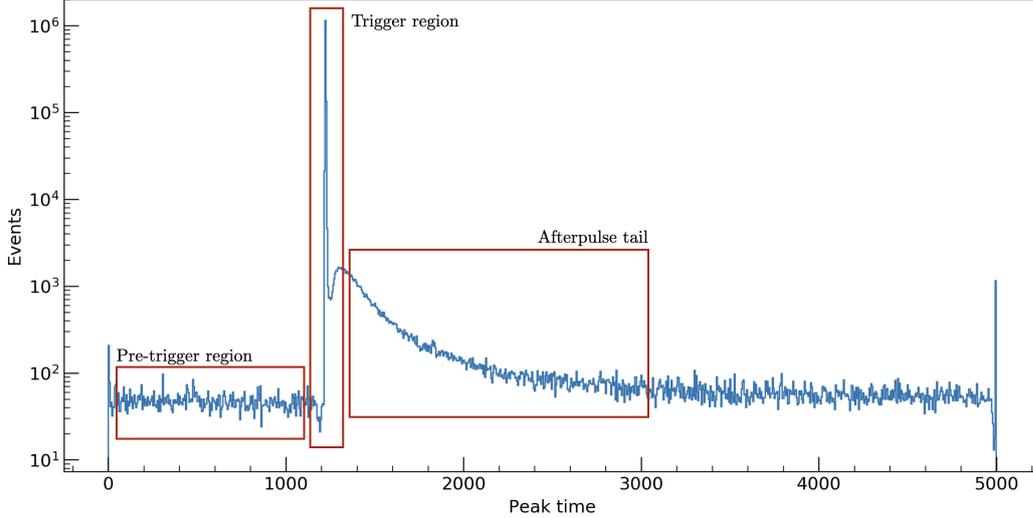


Figure 4.15: Hit time histogram in the 5000 sample acquisition window. The external trigger is at 1220 samples. The well between the trigger peak and the after-pulse tail correspond to after-pulses that are too close to the primary hit and with too low amplitude to be detected by the HFA. The peaks in the beginning and the end of the window are populated by artefacts of the filtering at the edge of the waveform, so they are excluded from the analysis.

The uncorrelated noise is the dark count rate. As those events are independent, they follow Poisson statistics with a constant rate. In order not to account for physical PEs and correlated noise when measuring this rate, we count those hits in the pre-trigger region, namely the region in the acquisition window before the laser time. The measured dark count rate is 40 Hz for $10 \times 10 \text{ cm}^2$ which corresponds to one DarkSide-20k read-out channel.

A deep understanding of the responses and efficiencies of the SiPMs is crucial to validate the photo-detection design of the experiment, to highlight possible caveats, and to optimise the online processing. Several prototypes were specifically setup to study different SiPM behaviours, providing the required data to build a reliable simulation of the DarkSide-20k photo-electronic. The results obtained for the noise and responses of the SiPMs are

given as input to the DAQ emulation software.

4.3 Event data model

The event data model is the definition of the raw data format of the detector. In the case of the trigger-less DAQ of DarkSide-20k, the physical events will only appear at the end of the reconstruction. The raw data contains the time slice information in a compressed format. They are based on the output of the FEPs that are the hits and ZLE segments information. The definition of the stored information must satisfy many requirements starting with the physics goals of the experiment. The data storage limitation and the available processing capability must also be taken into account. Therefore, this work required a very wide view of the experiment from the DAQ to the physics analysis.

4.3.1 Physics requirements

The physics potential of DarkSide-20k in the search for WIMPs with mass above $20 \text{ GeV}/c^2$ primarily relies on the efficient suppression of the ER background. This search region is referred as the high-mass in opposition with the S2-only analysis searching for signals of lower mass candidates. In the high-mass search, once multiple-scattering events are rejected, the residual background is mostly suppressed by the PSD.

The low-energy threshold in the high-mass channel is set at 35 keV_{nr} . This NR energy corresponds to the light that produces a 8 keV_{er} ER, the difference coming from the higher quenching of NR in argon. This threshold is designed to guarantee the $>50\%$ acceptance band for WIMPs and less than 0.1 events leaking from the ER band, integrated over the full energy range and exposure. The high-energy threshold is indicatively set at 200 keV_{nr} (or 60 keV_{er}) in order to fully contain signals up to $10 \text{ TeV}/c^2$ WIMP mass. Given the light yield of the TPC, the region of interest in S1 photoelectrons is comprised between 80 and 1200 PE. This energy region being the primary goal of the experiment, it is from now on referred to as the golden region.

The most critical area of the golden region is the one near the low-energy threshold in the $(S1, f_p)$ parameter space. This is because the NR and ER bands significantly overlap below this energy. The variables derived from the acquired waveforms must then guarantee the resolution to disentangle NRs

and ERs down to 80 PEs, while keeping linearity in the energy response for the full S1 range.

For what concerns S2 in the golden region, the main requirement is the accurate reconstruction of the electroluminescence pulse shape in order to discriminate multiple-scattering events and pile-ups of S1 and S2 pulses. It must also give a good estimator for the z -position from the electron diffusion effect along the drift. Furthermore, the linearity of the reconstructed S2 would allow to exploit the ER/NR separation and alpha event rejection of the S2/S1 ratio.

The range below 8 keV_{er} is referred to as the silver region. It is of interest for the search of light dark matter particles with the S2-only analysis. It also has excellent potential in the detection of core-collapse supernova neutrinos via CE ν NS. In this range, the requirements are the high detection efficiency for S2 signals induced by single electrons and excellent reconstruction of the S2 pulse shape. The latter is essential to reject low-energy multiple-scatter events and pulses pile-up. The estimation of the z position from the electron diffusion effect is also crucial as it is, in an S2-only analysis, the only way to perform fiducialisation along the vertical axis.

Finally the region above the 60 keV_{er} is mainly devoted to understanding and modelling of the background by looking at the characteristic α and γ peaks from radioactive decays. The identification of full absorption peaks from gammas and alphas requires good linearity in the S1 response, albeit with less stringent relative resolution than the one required in the golden range.

4.3.2 Summed waveforms

For the purpose of pulse shape reconstruction and data quality monitoring, we considered the possibility of storing summed waveforms in addition to the hits and ZLEs information. Two waveforms are produced by summing all the ZLE segments of each optical plane. The segments identified by the ZLE algorithm have to be previously normalised by the channel gain. To control the size of the stored data, the resulting WFs are downsampled by a factor 10. The time sampling is then 80 ns instead of 8 ns which is still enough for any pulse analysis as the time scale for an S2 is $\mathcal{O}(10 \mu\text{s})$.

At this stage, the waveforms would be mostly populated by isolated hits

from dark count rate. This is solved by applying a time-over-threshold before downsampling of 2 PEs for at least 80 ns. There is no risk of truncating physical pulses as this is done after the summation, so even for low energy S1, the height of the waveform is way above 2 PEs. Finally the waveforms are written with zero-length encoding compression to further reduce the data size. An example of summed waveform after downsampling and time-over-threshold is shown in figure 4.16.

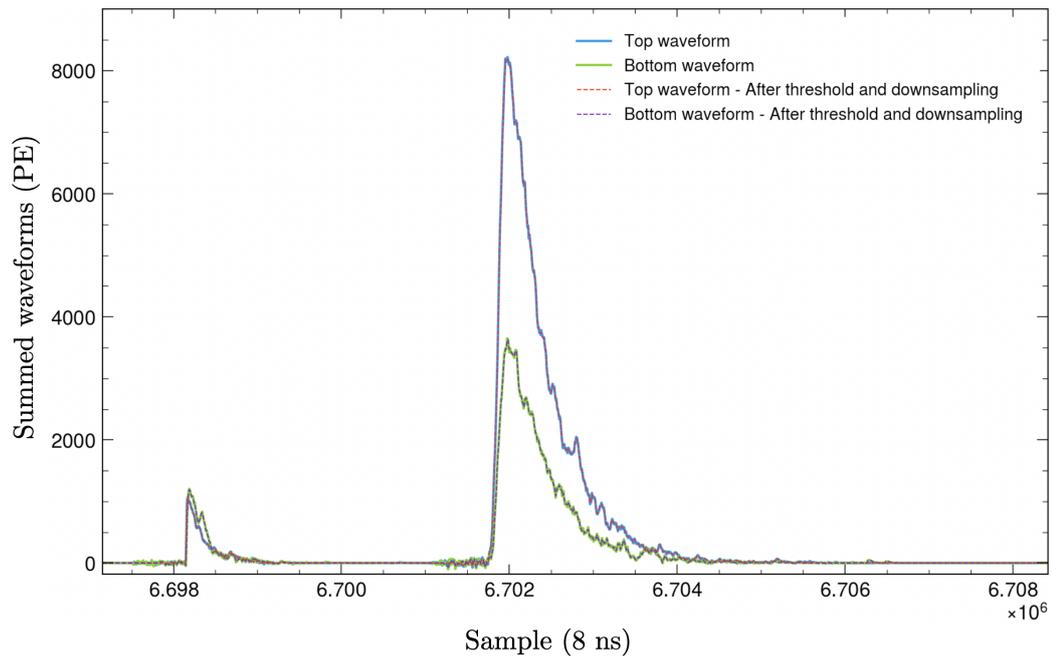


Figure 4.16: Summed waveforms before and after time-over-threshold cut and downsampling for an S1 and an S2. The waveforms are here normalised by the amplitude of one PE assuming ideal gain calibration.

The summed waveforms would definitely be a powerful tool for pulse shape analysis especially for S2 pulses as the HFA is not optimised for them. The redundancy of the information between the hits distribution and waveforms is a precious feature to understand the data during the analysis. However, the online implementation of the summing is not trivial. In principle, at the FEP levels, the waveforms are lost and the TSP only receives processed hits and ZLEs information. A solution could be to partially sum the waveforms in the FEP for all the channels it is connected to, namely 128 channels. Moreover, the summed waveforms strongly relies on online channel-based gain calibration. This requires the development of a periodic

fast analysis of finger plots. The strategy, adopted to be as conservative as possible, is to consider the summed waveforms in the data rate computation but to not develop any of the reconstruction algorithms relying on them.

4.3.3 DAQ output data structure

The output of the DAQ will be a file containing a collection of consecutive time slices (table 4.3.3). The baseline for the raw data format of DarkSide-20k is centred around the ZLE segments and the hits found by the HFA (figure 4.17). The time slice data format developed for the event data model consists in a header and the following five structures:

- **ZLE**: with information associated to ZLE segments. It includes the channel, the integral of the waveform in the segment, the time difference with respect to the slice timestamp, the duration of the segment in samples, and the number of hits contained. The latter is used when reading the slice to associate hits and ZLEs.
- **ZHIT**: containing the prominence and the sample position of the hit with respect to the beginning of the ZLE segment. The hit charge are not stored as they are redundant with the ZLE charge and do not provide SPE resolution as good as the prominence. The channel number does not have to be written as it can be recovered from the associated ZLE.
- **IVHIT**: associated to the inner veto hits. The inner and outer veto structures are for the moment assumed to be identical to the TPC hits. This can be adapt later on to match the specific requirements of the vetoes.
- **OVHIT**: associated to the outer veto hits.
- **WF**: containing the downsampled and compressed summed top or bottom waveforms.

The header contains general information of the slice, an ID and a timestamp, the size of each structure, and monitoring information. The structure sizes allow to eventually skip slices or structures, making the reading of the file faster. It also serves as a check to probe any issue in the writing. The monitoring variables are the mean and the RMS of the baseline for all channels. The writing and reading of this structure as a custom binary file is

implemented in the DAQ emulation tool to ensure efficient manipulation and flexibility. The size of the time slice depends on the event rates in the three detectors following:

$$\begin{aligned} \text{tslice size [bytes]} = & 22032 + 18 \times n_{\text{zles}} + 6 \times n_{\text{zhits}} \\ & + 8 \times n_{\text{ivhits}} + 8 \times n_{\text{ovhits}} \\ & + 4 \times n_{\text{wf_top}} + 4 \times n_{\text{wf_bot}}. \end{aligned}$$

Level	Variable	Type	Size (bytes)
HEADER	id	int32	4
	n_zles	int32	4
	n_zhits	int32	4
	n_ivhits	int32	4
	n_ovhits	int32	4
	n_wf_top	int32	4
	n_wf_bot	int32	4
	time	float32	4
	mean_baseline	float32	4×2750
rms_baseline	float32	4×2750	
ZLE	channel	int16	2×n_zles
	timediff	float32	4×n_zles
	integral	float32	4×n_zles
	nsamples	int32	4×n_zles
	n_zhits_in_zle	int16	2×n_zles
ZHIT	sample	int16	2×n_zhits
	prominence	float32	4×n_zhits
IVHIT	timediff	float32	4×n_ivhits
	charge	float32	4×n_ivhits
OVHIT	timediff	float32	4×n_ovhits
	charge	float32	4×n_ovhits
WF	wf_top	float32	4×n_wf_top
	wf_bot	float32	4×n_wf_bot

Table 4.2: The time slice data structure.

The reported hit structures for inner and outer vetoes is only preliminary. The DAQ emulation is not yet developed for the vetoes and is needed to determine the data model requirement. Still, the inner veto data rate is expected to be smaller than in TPC because the optical coverage is smaller

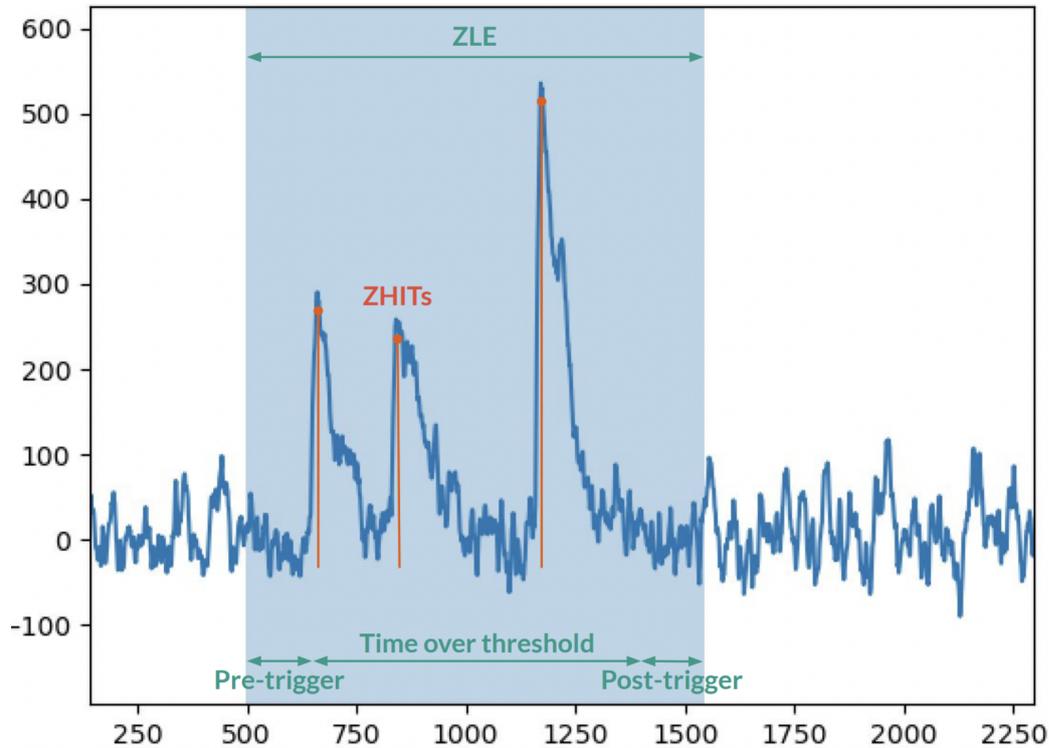


Figure 4.17: A simulated waveform in which a ZLE containing 3 hits is identified.

and there are no S2 pulses.

4.3.4 Digitizer transfer rate and data storage

Given the data structure previously developed and assuming nominal background activity and photo-electronic noise, the data rate can be estimated with the event mixer and the DAQ emulation software. The obtained rate of writing data to disk is around 24 MB/s from TPC only (figure 4.18). This was obtained by applying the full simulation chain to produce in total 10 s of data. The data rate is almost equally shared between ZLEs and hits data structures because one ZLE is on average associated to 3.5 hits and its structure is 2.7 times larger in size (figure 4.19). The contribution of the summed waveforms to the data storage is less than 4%. The transfer rate for 1 FEP is in average 40 MB/s (figure 4.20).

To stress the system, the simulations were repeated by increasing the

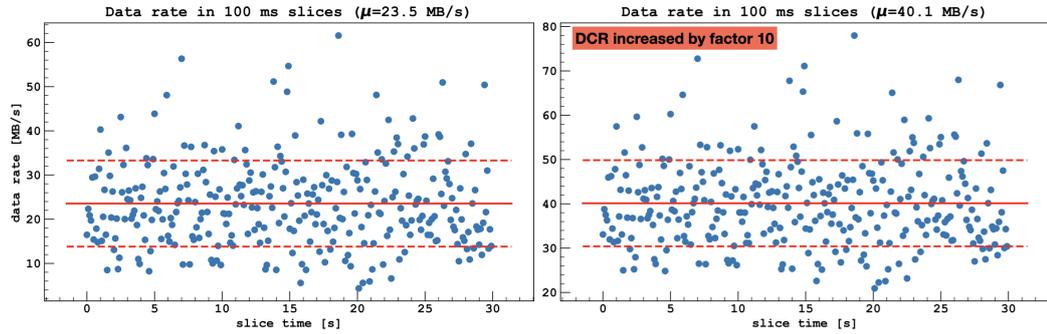


Figure 4.18: Expected data rates of 100 ms slices with nominal dark count rate (left) and 10 times increased dark count rate (right).

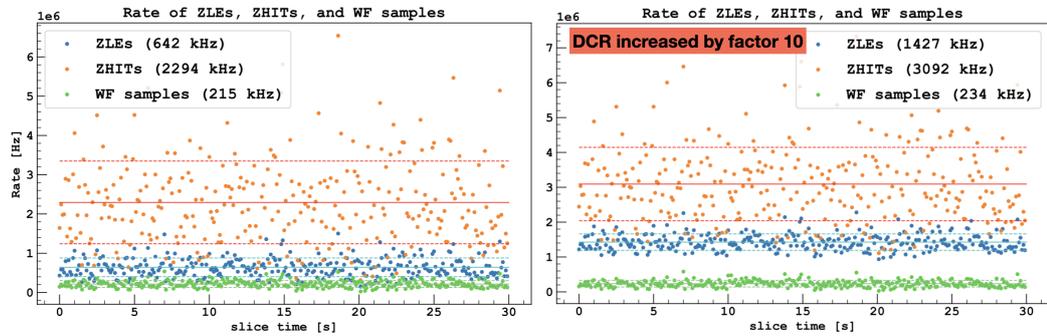


Figure 4.19: Distribution of ZLEs, hits and waveform rates with nominal dark count rate (left) and 10 times increased dark count rate (right).

dark count rate by a factor 10. The rate of ZLEs increased by 2, the hit rate by 50%, and the rate of the summed waveform samples by 10%, which reinforces the confidence that the time-over-threshold cut rejects efficiently dark noise. The digitizer-to-FEP mean transfer rate increased from 40 to 64 MB/s and the rate of data to disk from 24 to 40 MB/s.

In conclusion, we estimated that with this event data model design, the data acquisition is able to tolerate the nominal background rate even with an unexpected increase of dark noise. It must as well foresee calibration runs during which the TPC will be exposed to radioactive sources, inducing an higher event rate. The activity of the different sources are tuned accordingly, in order to avoid DAQ saturation.

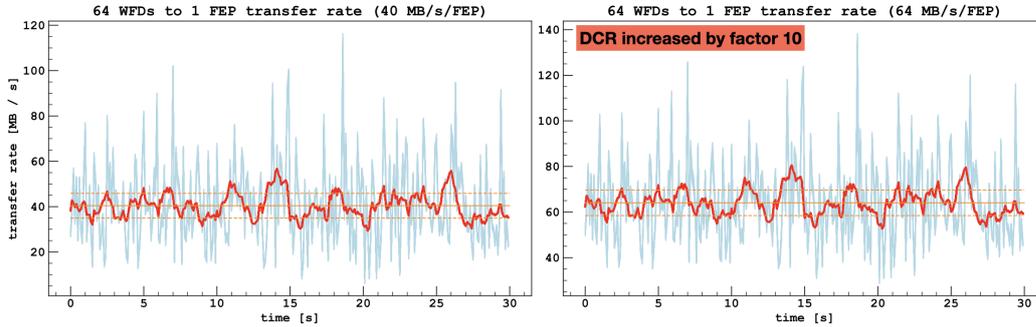


Figure 4.20: Expected transfer rate per digitiser for 100 ms slices (blue) and 1 s slices (red) with nominal dark count rate (left) and 10 times increased dark count rate (right).

4.4 Physics validation

The definition of the event data model must be guided by the physics goals. It was developed and validated by applying it on simulated data and checking that the main physics observables can be reconstructed. ER and NR samples were simulated uniformly in the TPC in the energy range of the golden region, namely $[20, 500]$ keV_{nr} and $[8, 200]$ keV_{er}. To avoid biases or uncertainties due to pile-up or inefficiency of the clustering algorithm, S1 and S2 are independently simulated and reconstructed.

4.4.1 PE detection efficiency

The first feature to validate is the PE detection efficiency. Photo-electron finger-plots can be produced by histogramming ZLE charges and hit prominences from S1 pulses as these, differently from S2, are under single-PE regime at the channel level. The single PE resolution is estimated at 26.0% for the charge variable and 17.4% for the prominence. The resolution is altered with respect to the HFA study because the PE pile-ups in S1 pulse smear the prominence reconstruction. Once again, all channels are simulated with the same gain at this stage. The finger plots obtained through the whole DAQ emulation and from the Event Data Model are shown in figure 4.21.

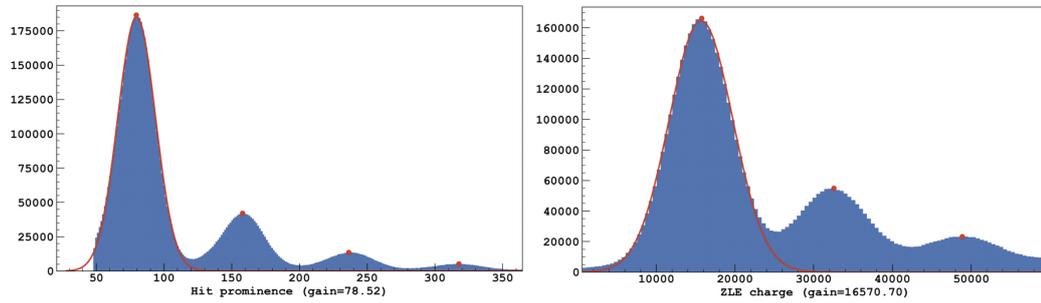


Figure 4.21: Finger plots from hit prominence (left) and ZLE integral (right) from S1 pulses. The truncation of the prominence is due to the 0.6 PE cut of the HFA.

4.4.2 Energy reconstruction

The primary energy estimator is the total light of the S1 pulse. To validate the capability of reconstructing energy, the event data model includes the 3 following observables:

- **Number of hits** per pulse regardless of their prominence values: this variable is not affected by intrinsic charge or prominence resolutions. It is linear in low channel occupancy regimes (<0.1 PE/channel). Moreover, the number of hits is blind to internal cross-talk, as these occurs simultaneously with primary avalanches. Therefore, it is expected to provide the best performance in PSD, because it minimises fluctuations from instrumental noise. However, this variable suffers from saturation effects, as photoelectrons occurring on the same channel and in close time (<500 ns) can only be partially resolved.
- **Sum of the prominence** of all identified hits, normalised by the gain: such a variable partially recovers the non-linearity of the number of hits as it scales with the number of hits occurring simultaneously (figure 4.22). However, the prominence variable cannot recover the whole inefficiency of the hit finder, and therefore the associated non-linearity. This specifically happens for second avalanches following a first one after 100 to 300 ns. It is also affected with some fluctuation due to internal cross-talk.
- **Sum of the ZLE integral** per pulse, normalised by charge gain: as the ZLE finder is highly efficient, the charge variable has good linearity with the number of photoelectrons (figure 4.23). However, it integrates also instrumental noise such as after-pulses and cross-talks, which, in

addition to the intrinsic baseline fluctuations, worsens the resolution. In addition, electronic and DAQ saturation can affect the charge for S2 pulses and for very high energy S1.

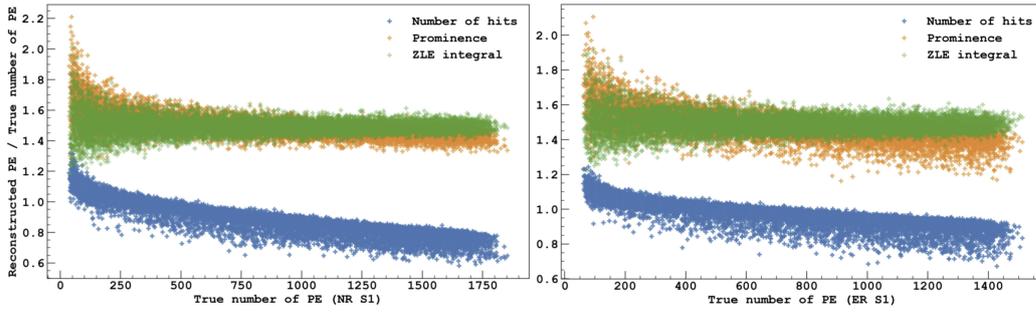


Figure 4.22: Ratio between energy estimator and true number of PEs as a function of energy for S1 NR (left) and ER (right).

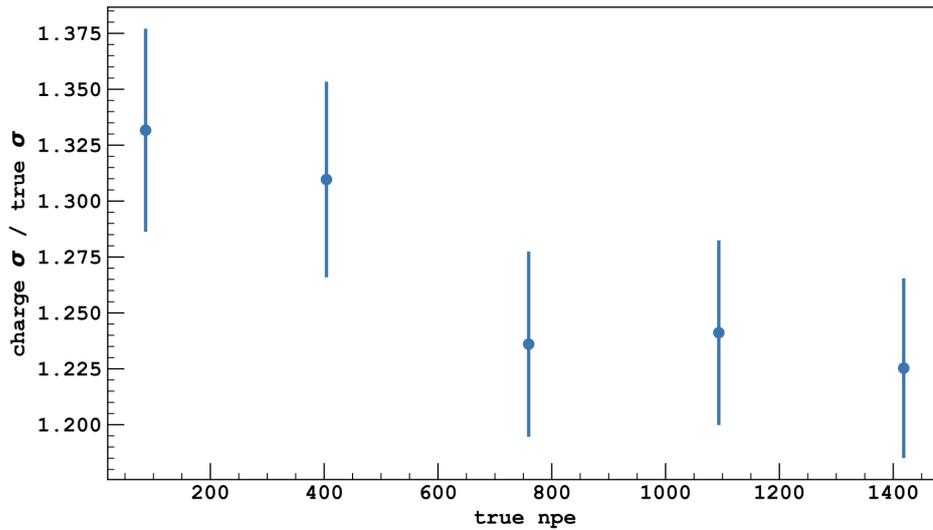


Figure 4.23: Ratio between standard deviations of the ZLE reconstructed number of PEs and the true number of PEs as a function of the energy.

4.4.3 Pulse shape discrimination

PSD of S1 pulses in the golden region is evaluated by computing f_{200} , the fraction of light contained in the first 200 ns of the pulse. It is computed on the two hit-based observables, namely with and without prominence weights.

The ZLE charge variable is not expected to provide good separation as its resolution is the worst of the three observables at low energy. Indeed, the lower bound of the golden region is the most crucial as it is the one near which the ER and NR bands are the closest.

As mentioned previously, no pulse finder algorithm is applied at this stage and, therefore, the pulse start time has no uncertainty. It appears, as expected, that the number of hits ensures better separation than the prominence, since the former is less sensitive to correlated instrumental noise. Figure 4.24 shows the separation between ER and NR in the golden region. The low statistics of the analysed NR and ER samples limit the precision of the two bands but are enough to compare the two observables and to affirm that the PSD requirements can be fulfilled.

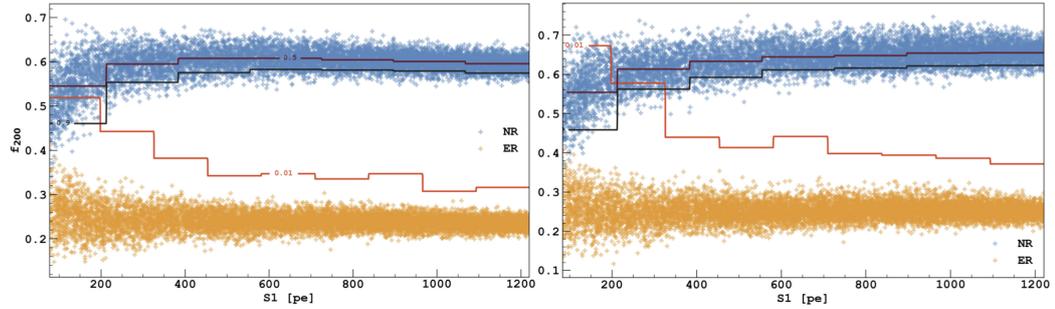


Figure 4.24: f_{200} computed from hits without (left) and with prominence weights (right) as a function of S1 computed from the ZLE charge variable. The red line is the maximum leakage curve for ER of 0.1 event in 100 ton year. The black and dark red lines are the NR acceptance curve at 90% and 50% respectively.

4.4.4 Position reconstruction

The primary goal concerning position reconstruction is high resolution in xy plane from S2 pulses. The ZLE integral gives better resolution than with the two other variables (figure 4.25). To quantify the resolution that can be reached, we developed a basic algorithm consisting in identifying the 12 ZLEs of the pulse with the largest charges. The reconstructed position is the barycentre of the ZLEs channel position weighted by their charges. The resolution along the x or the y axis for ERs was measured to be 3 cm at 10 keV_{er} and 2 cm at 100 keV_{er}. This precision is enough for fiducialisation

and can also ensure good multiple scatter identification and pulse association.

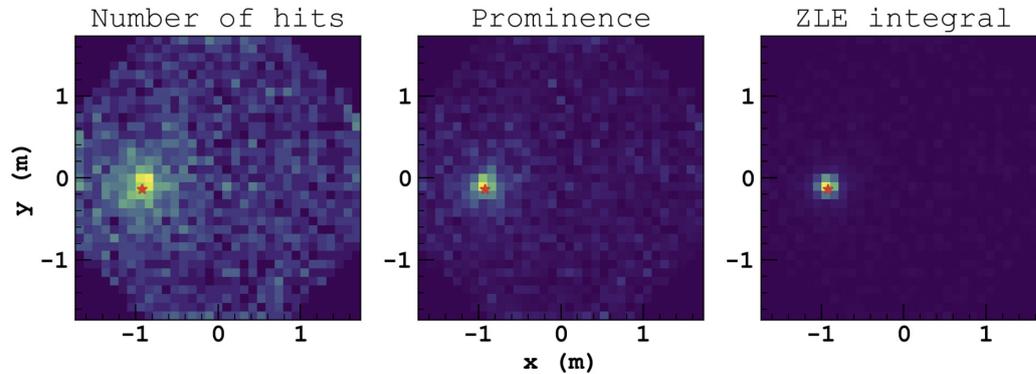


Figure 4.25: Channel distribution of an S2 pulse for the 3 observables. The red stars indicates the true position of the event.

More advanced algorithms will provide better reconstruction from S2 in the xy plane and will also allow reconstruction of the depth of the interaction from the diffusion. The latter is necessary for low-mass search without S1 in which it is the only way to perform fiducialisation in the z -axis, and also for strong S1-S2 pulse association to reduce dead time or accidental coincidences due to pile-up. This study is yet to be done, but hints of the diffusion can already be seen from the proposed data structure. As it can be seen on figure 4.26, S2 pulses appear much wider from the hit distribution and from the summed waveforms.

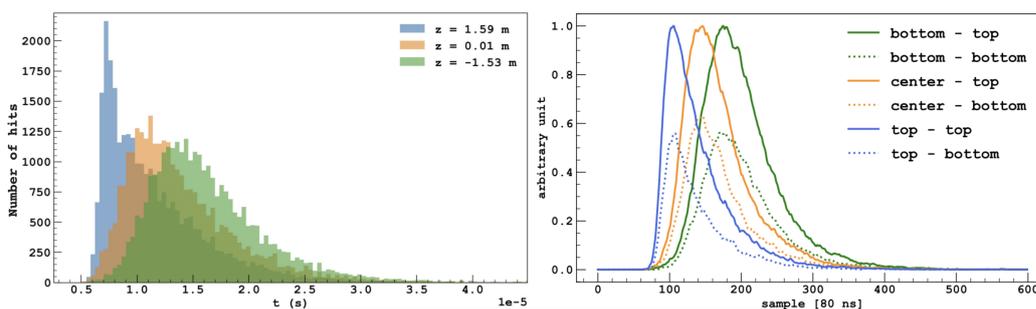


Figure 4.26: Hit time distribution (left) and summed waveforms (right) of 3 different S2 pulses, one from the top, one in the center and one on the bottom of the TPC.

Chapter 5

Data reconstruction

The scope of the reconstruction is to extract the information on the physical events, starting from the raw data defined in the Event Data Model. A physical event consists in the sequence of pulses induced by the same interaction. The reconstructed event must contain information on the energy, position, pulse shape, and topology of the interaction.

The reconstruction framework is divided in four layers depicted in figure 5.1. The pulses have first to be identified within the hit distribution (layer 0). They undergo a first basic reconstruction and are classified between S1, S2, or unresolved pulses (layer 1). Then, more advanced algorithms are applied to reconstruct position and specific observables according to their natures (layer 2). The last step is the event building, which consist in the definition of the physical events by associating S1 and S2 pulses on the basis of the previously reconstructed variables (layer 3).

I developed the first three layers of the reconstruction software, from a clustering based on a custom Bayesian Blocks algorithm to preliminary high-level pulse reconstruction. The event builder is not addressed yet. This reconstruction scheme was then implemented in a modular framework that ensures efficient offline reprocessing as later discussed.

5.1 Layer 0 - Clustering

The clustering is the first step of the reconstruction, and the core of the pulse finder. Its aim is to distinguish between physical hits and dark noise. The algorithm searches for clusters of hits induced by physical pulses of light. As DarkSide-20k operates in trigger-less mode, the clustering has to be applied

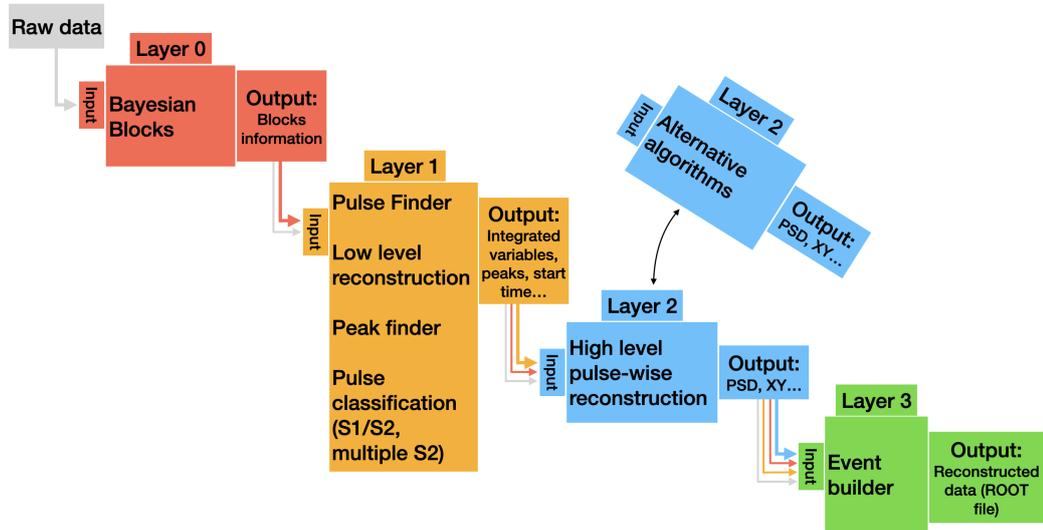


Figure 5.1: Modular reconstruction framework in 4 layers. It ensures the possibility to replace a set of algorithms by another while keeping compatibility with previous and next layers. Raw data refers to the time slices as defined in the Event Data Model.

over the whole acquisition time and must therefore be carefully optimised to keep the computing time under control. The input of the clustering algorithm is the hit time distribution and it is designed to identify significant excesses in the hit rate which are then interpreted as pulses.

5.1.1 Bayesian Blocks

The proposed approach is based on the Bayesian Blocks, commonly used in astronomical time series analysis [69]. It aims at detecting localised structures, pulse shapes, and intensity variations in counting data. The output is a segmentation of the time series into intervals called Bayesian Blocks inside which the data rate is constant within Poisson-compatible fluctuations.

The block lengths are not fixed which makes the algorithm sensitive to any structures regardless of their durations. The principle of the method is to iteratively add "change points", defined as the edges of the Bayesian Blocks [70]. The positions of the change points are the ones that maximise a so-called fitness of blocks, which is a function that describes how well the signal rate is constant within a block. The fitness of blocks is additive, so that the fitness of multiple blocks is the sum of the fitness of each block.

In general, the hit rate can be described, as a function of the time t and a vector of parameters Θ , by a model $M(t, \Theta)$ [71]. By considering sufficiently small bins, each containing either 1 or 0 hit, the likelihood can be written as:

$$\mathcal{L}(\Theta) = \prod_{i \in P} \mathcal{P}(1, M(t_i, \Theta)) \times \prod_{i \in Q} \mathcal{P}(0, M(t_i, \Theta)) \quad (5.1)$$

P and Q are the set of bins with respectively 1 and 0 hit. $\mathcal{P}(n, \mu)$ is the Poisson probability to have n hits for a given occupancy μ . In our case, $M(t, \Theta)$ is a Poisson distribution with constant rate R , such that within a bin of duration δt , the occupancy is $\mu_{bin} = R\delta t$. The likelihood can then be written as:

$$\begin{aligned} \mathcal{L}(R) &= \prod_{i \in P} \mu_{bin} e^{-\mu_{bin}} \times \prod_{i \in Q} e^{-\mu_{bin}} \\ &= (\mu_{bin} e^{-\mu_{bin}})^N (e^{-\mu_{bin}})^{M-N} \end{aligned} \quad (5.2)$$

Here, N is the total number of hits and M is the total number of bins in the block, so $M - N$ is the number of bins without hit. For the fitness of blocks to be additive, it is defined as the logarithm of this likelihood, giving:

$$\begin{aligned} \ln \mathcal{L}(R) &= N (\ln \mu_{bin} - \mu_{bin}) - (M - N)\mu_{bin} \\ &= N \ln \mu_{bin} - M\mu_{bin} \end{aligned} \quad (5.3)$$

The log-likelihood must not depend on the arbitrary binning. We replace μ_{bin} with its value $R\delta t$ and notice that $\delta t \times M = T$, where T is the total duration of the block. Then, the log-likelihood can be rewritten as:

$$\ln \mathcal{L}(R) - N \ln \delta t = N \ln R - TR \quad (5.4)$$

The $N \ln \delta t$ term is put on the left hand side to highlight that it does not depend on the model. Therefore, the optimisation of the fitness of blocks is equivalent to the maximisation of the right hand side of this equation. This log-likelihood is indeed maximum for $R = N/T$.

At this stage, the optimal segmentation is obtained by identifying as many blocks as hits, which would result in a wrong overfitted model. To avoid this, a penalty factor is included to the log-likelihood by adding the number of blocks as a parameter of the segmentation. To that end, the total likelihood is multiplied by a prior probability $P(N_{blocks})$ on the number of blocks:

$$P(N_{blocks}) = \frac{1 - \gamma}{1 - \gamma^{N+1}} \gamma^{N_{blocks}} \quad (5.5)$$

This prior is a geometric function of a single parameter γ , normalised for N_{blocks} between 0 and the total number of hits N . With $0 < \gamma < 1$, the prior is decreasing with N_{blocks} , penalising segmentations with high number of blocks. This parametrisation can be interpreted as a tuning of the false positive probability.

5.1.2 Fast Bayesian Blocks algorithm

For a given hit time series, the Bayesian Blocks algorithm returns the optimal segmentation (figure 5.2). The series is first divided in time cells containing only one hit. Two loops are performed over the cells. The log-likelihoods are computed for blocks starting at cell j and ending at cell i for i going from 1 to the number of cells and for j going from 1 to i (figure 5.3). The method ensures to get the best partition possible without testing all of them, exploiting the following theorem: a sub-partition \mathcal{P}' of an optimal partition $\mathcal{P}(I)$ is an optimal partition of the subset I' that it covers.

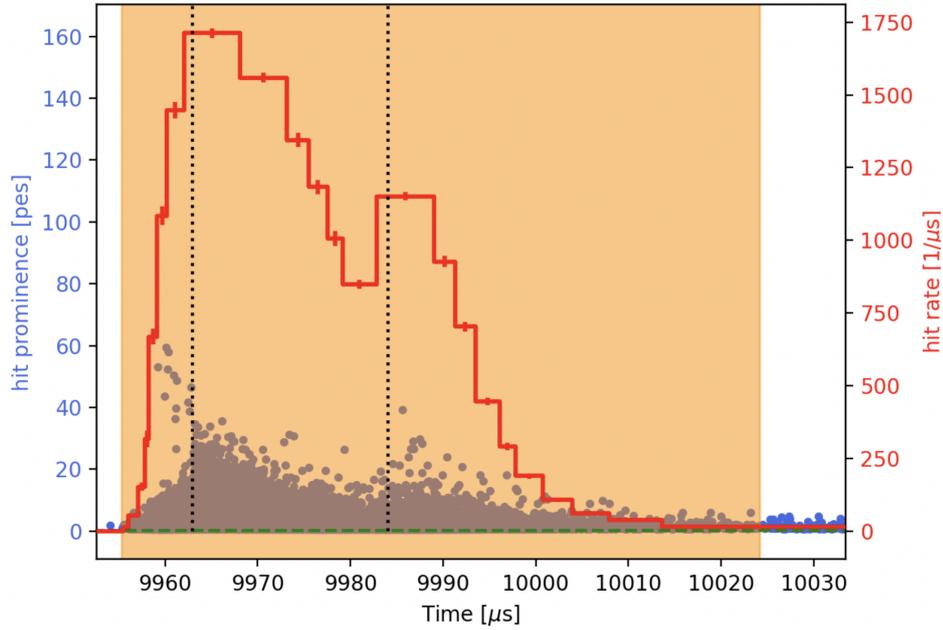


Figure 5.2: Example of a preliminary cluster reconstructed from the Bayesian Blocks for two S2 pulses. The blue points are hits with a height proportional to their prominence. The red segments are the Bayesian Blocks. The dotted vertical lines are the true MC times and the horizontal green dashed line is the rate threshold.

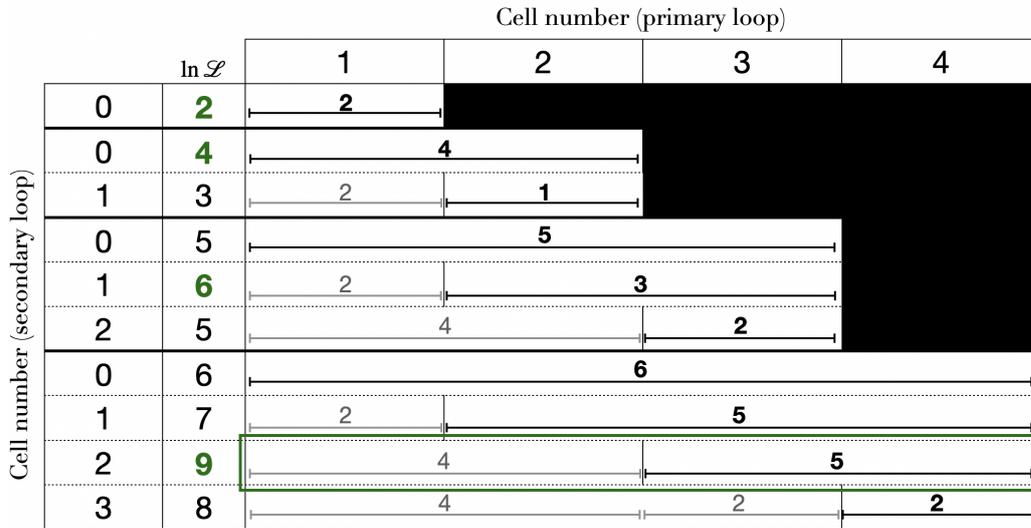


Figure 5.3: Loops in the bayesian blocks algorithm in an example with 4 cells. In the primary loop, the log-likelihood is computed for a block starting from every previous cell to the last one considered. The so-computed log-likelihood from cell i to N is added to the one of the best segmentation previously found between 1 and i . The green square shows the optimal segmentation of the 4 cells with a total log-likelihood of 9.

As two loops are performed over the cells, the computing time scales with the square of the number of hits. This is unsuitable to process the data of DarkSide-20k. We developed a solution which consists in applying the algorithm on chunks of time slices, each containing the same number of hits, fixed at 1000 hits per chunk. This approach however affects the performance on the edges of the chunks. To solve this, we make each chunk overlap on a fraction of the next one. The change points found in the overlap region are the ones of the previous chunk if they are in the first half and the ones of the next chunk otherwise.

Two overlap fractions were tested, either 50% or 100%. 100% overlap means that the full chunk is overlapping with the previous or the next one, in other words the end of the chunk i is the beginning of the chunk $i + 2$. If the overlap is 50%, the first (last) quarter of each chunk is overlapping with the previous (next) chunk (figure 5.4). In the 50% overlap case, it is also required to merge consecutive blocks whose rates are compatible within 3σ . The threshold to define the merging condition was optimised so that the 50% and 100% overlap give the same efficiency.

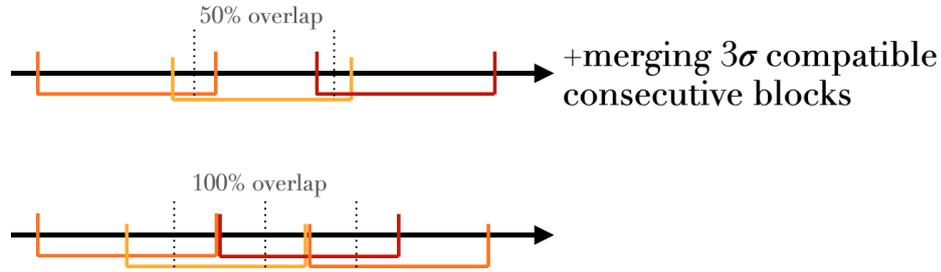


Figure 5.4: 50% and 100% chunk overlap on which bayesian blocks are computed. The dotted lines represent the limits before which change points are taken from the previous chunk and after which they are taken from the next chunk.

As each chunk contains the same number of hits, the required time to apply the algorithm is the same for every chunk and the total computing time scales linearly with the number of hits. In figure 5.5, the computing time is shown before and after the implementation of the chunk approach. It appeared that the 50% overlap is significantly faster and provides equivalent performances than the 100% overlap, therefore it is used from now on. In addition to this important upgrade, other improvements of the code allowed to significantly lower the computing time. They include Numba implementation and technical simplification of the log-likelihood computation. The full process takes around 3 s per second of data on an Apple M1 core. This speed is satisfactory, although it will very likely remain the longest step of the reconstruction.

To provide the possibility of further increase of the computing speed, we implemented a downsampling option. It consists in applying the Bayesian Blocks algorithm on a subset of the hit series. Given a downsampling factor n , one hit out of n is selected and the algorithm is applied the same way as previously described.

5.2 Layer 1 - Pulse finder and classification

A first set of algorithms for the reconstruction was developed in the layer 1. It starts with the pulse finder, based on the Bayesian Blocks found in layer 0. Within each pulse the three energy observables are integrated and the PSD estimator f_{200} , defined as the fraction of hits in a 200 ns prompt window of

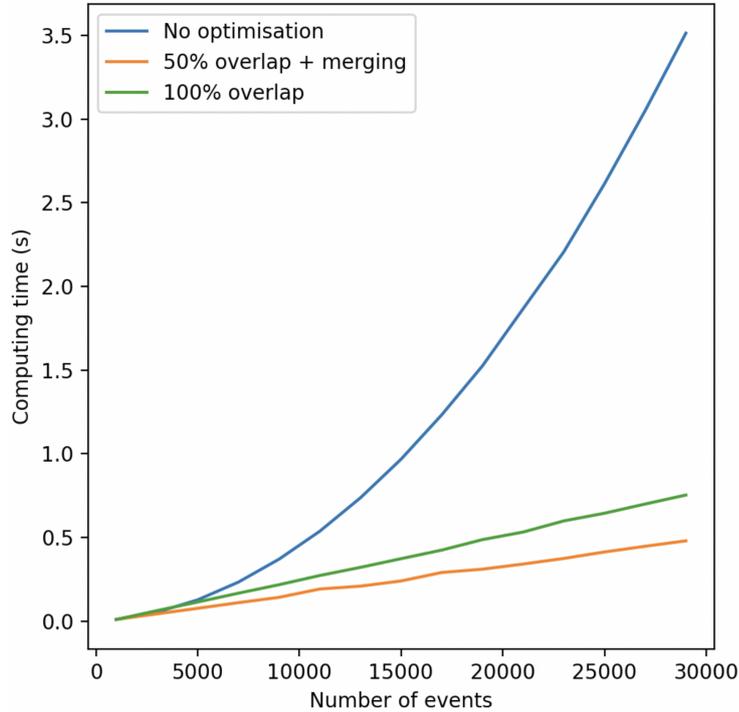


Figure 5.5: Computing time needed to process bayesian blocks algorithm without chunk optimisation and with 50% and 100% chunk overlap. This test is done interactively on a single core, before other optimisations.

the pulse, is computed. A peak finder was also developed to identify pulse structures that are not resolved by the Bayesian Blocks. Still in layer 1, these variables are then used to perform the pulse classification.

5.2.1 Pulse finder

The hit rate within each Bayesian Block is computed and used to define preliminary clusters. The rate threshold above which consecutive blocks form a cluster is set at $0.4 \text{ hits}/\mu\text{s}$, which is the equivalent rate of an S2 induced by 0.5 electron. Since the single electron resolution is less than 25 %, this threshold ensures close to 100% efficiency in identifying S2 pulses down to 1 electron. S1 detection showed the same performance in the region of interest for the standard WIMP search ($98.8\% \pm 0.2\%$ at 10 keV_{er}).

Once the cluster is found, we apply an algorithm to check for unresolved pulses. These adjacent pulses are produced by S2s from multiple scatter

events, by pile-up, or by events from the top for which S1 and S2 are close in time. Since a pulse with sufficiently high energy is segmented in multiple blocks, information on the pulse shape is partially contained in the rates of the blocks. In particular, a rate increase in the middle of a preliminary cluster is interpreted as a secondary pulse. This is implemented by splitting the cluster into pulses at local minima of block hit rate. Figure 5.6 shows an example of a preliminary cluster identified from the Bayesian Blocks and split into pulses.

In case of close pulses, the tail of the earliest one overlaps with the next, therefore, considering them as two independent structures would bias the reconstruction. For this reason, the split of preliminary clusters is done only if the rate of the local minima is below a threshold set at 1 hit/sample. This approach motivates the development of a peak finder algorithm to apply within the pulses in a next step of the reconstruction.

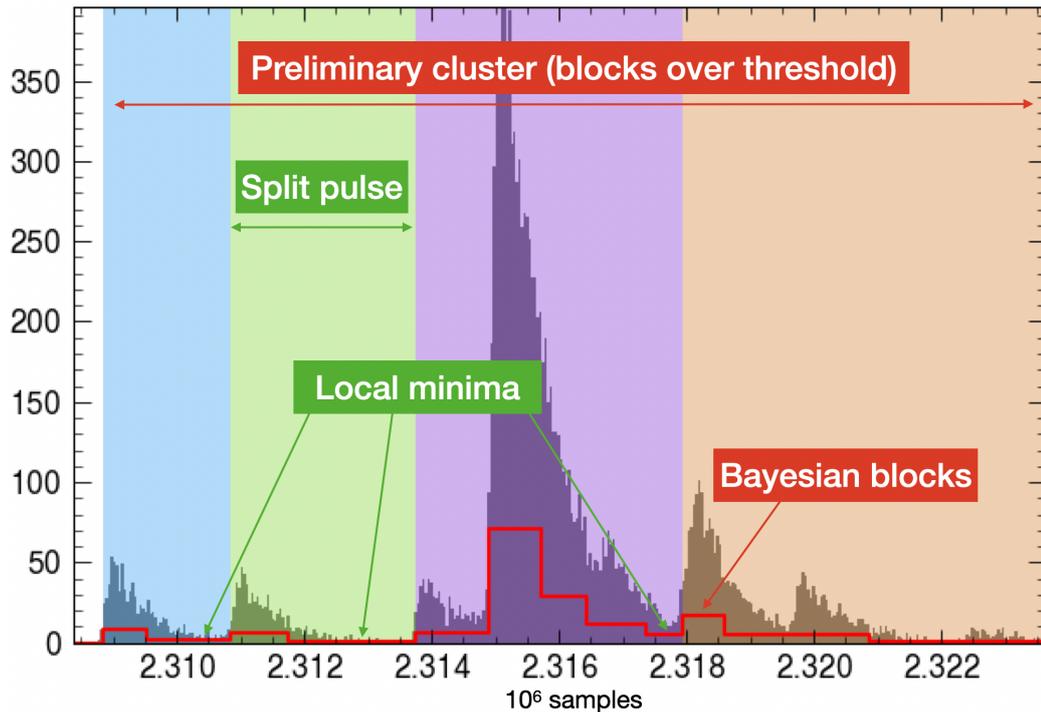


Figure 5.6: Bayesian block based pulse finder on a hit distribution from multiple scatter gamma. The height of the blocks is proportional to their hit rates.

The efficiency of the pulse finder is evaluated on the full background sim-

ulation. We defined two observables: the inefficiency as the ratio between unidentified pulses and the total number of pulses from the Monte Carlo truth; then, the false positive probability defined as the ratio between the number of identified pulses that do not contain any true Monte Carlo pulse and the total number of identified pulses. Those metrics are evaluated with and without hit downsampling and used to optimise the γ parameter of the Bayesian Blocks. The performances and the impact of the downsampling factor are reported in table 5.1.

These metrics allow to fix a maximum downsampling factor of 10 above which the inefficiency is no longer compatible with 0 within the error. With this downsampling the computing time is reduced by around 50%. Since it also decreases the fluctuations, the γ parameter must be optimised differently for each downsampling value. The downsampling remains an option that is not used in the following studies.

	Inefficiency	False positive
Definition	$\frac{\text{Unidentified MC pulses}}{\text{Number of MC pulses}}$	$\frac{\text{Empty identified pulses}}{\text{Number of identified pulses}}$
Without hit downsampling	0.02 % \pm 0.02%	0.27 % \pm 0.09%
With $\times 10$ downsampling	0.02 % \pm 0.02%	0.36 % \pm 0.09%

Table 5.1: Inefficiency and false positive probability of bayesian blocks pulse finder with and without downsampling.

5.2.2 Peak finder

The pulse finder based on Bayesian Blocks has very high efficiency and is rather good in resolving pulses with separation down to tens of μs with the local minima split method. Still, another step is required to properly identify substructures in clusters with closer pulses. A peak finder algorithm, applied on each pulse, is developed to tag more subtle shapes in the hit time distribution. It is important to keep in mind that the concerned events, mostly multiple scatter events, are not of interest for dark matter search, the precision of the reconstruction is therefore less crucial than their identification.

The challenge of the algorithm is to efficiently identify peaks from S1 and S2 pulses with energies spread across all the dynamic range of the detector. Moreover S2 pulse shapes strongly depend on the depth of the event, due to the diffusion effect. At the same time, it must be fast enough for the reconstruction scheme. We propose then an algorithm taking as input the hits of

the pulse and producing first an histogram of the time distribution. The bin size is the most crucial parameter of the whole algorithm. A too small binning will induce fake peaks identified on statistical fluctuations, while a too large binning would prevent separation of very close pulses. In the following, the size of one bin is fixed at 20 samples (160 ns).

Similarly to the hit finder algorithm at the SiPM level, we propose the use of the AR filter presented in section 4.2.3 to increase signal-to-noise ratio of the rate. The effect of the filtering can be seen on figure 5.7. The hit histogram is convoluted with a time-reversed template of an exponential decay. The aim is to exploit the decay shape common between all type of pulses, although their characteristic times are very different between S1 and S2. After optimisation it appeared that a decay time of 6 bins, so 960 ns, gives the best performance for this binning.

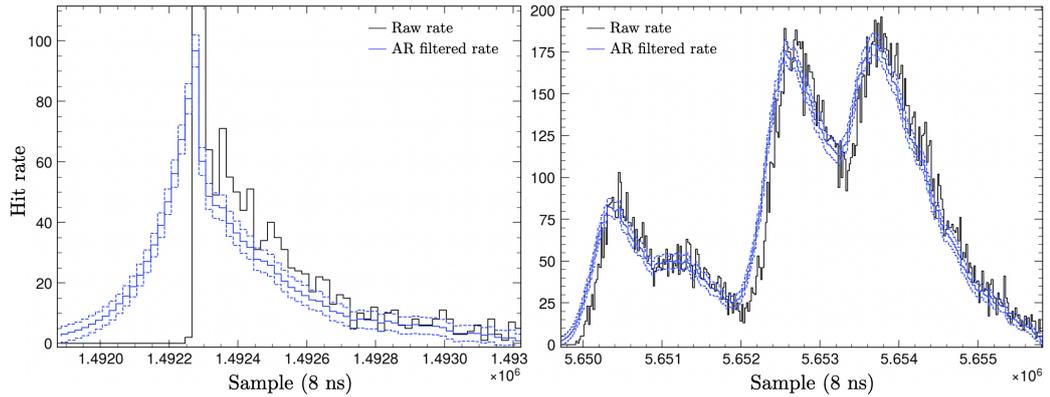


Figure 5.7: Hit rate and filtered rate of an isolated S1 pulse (left) and consecutive S2 from multiple scatter (right). The dashed lines are the error bands of the filtered rate.

Once the filtered rate is obtained, error bands are computed to define a threshold above which an increase will be interpreted as a new peak. As the content of each bin follows Poisson statistics, the error is written as:

$$\sigma(R) = A \times \sqrt{R} + B \quad (5.6)$$

Here, R is the rate in the bin and A and B are two additional parameters to optimise. While A is simply a factor to control the confidence level of the peak determination, B is an extra term for protecting against fake peaks due to fluctuations at low hit rate. Optimisation gave $A = 1/3$ and $B = 2$.

The peak finding based on the filtered rate is done through a single loop over the bins. A flag is used to determine regions in the histogram during which the rate is, within 1σ , either decreasing or increasing. Within each increasing regions, the bin with the highest rate is stored as a peak. The output of the peak finder is shown on a complex pulse in figure 5.8.

Within an increasing (decreasing) region, for sequence of bins with compatible rate within 1σ but whose central values has decreasing (increasing) trend, the error is divided by the square root of the number of consecutive decreasing (increasing) bins. This feature ensures some extra sensitivity to small peaks overlapping with the rise of a higher one for which the rate does not clearly decrease between the two.

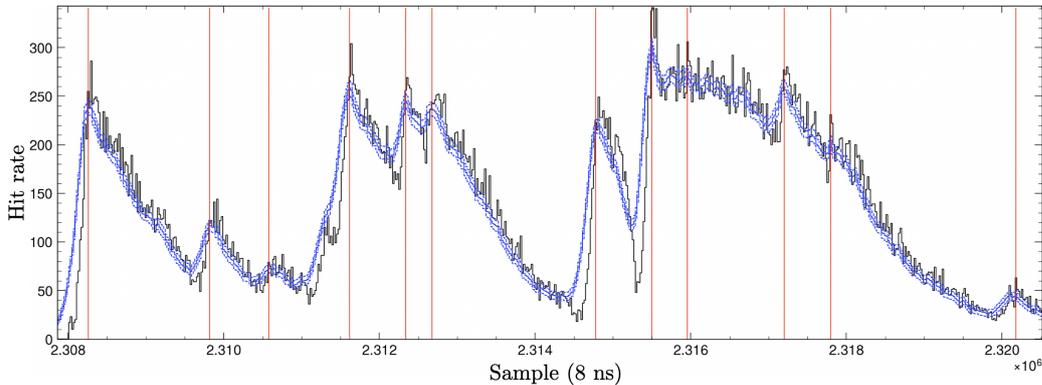


Figure 5.8: Peak finder algorithm performance on a complex pulse.

The purpose of the peak finder is to identify multiple peaks that were not resolved by the Bayesian Blocks. A 10 s sample of the full background simulation was generated to assess the performance of the algorithm. The metrics are only based on proper identification of single or multiple pulses. The timing or the precise number of peaks in complex pulses are not of interest at this stage. Over the full background simulation, the probability of identifying fake peaks in a single pulse are $0.0\% \pm 0.1\%$ and $0.2\% \pm 0.1\%$ for S1 and S2 respectively.

The probability of properly identifying multiple structures in pulses with consecutive S2s depends on the time difference between the peaks. This probability was estimated on pulses containing multiple true Monte Carlo S2s as a function of the standard deviation of their times. It does not take into consideration the energies of the deposits. If a pulse with relatively

small amplitude overlaps a much higher one, it will be harder to resolve. It is then expected that the efficiency in separating deposits with very different energies is low. The separation showed in figure 5.9 is then a conservative estimation of the capability of the peak finder. It was mainly produced for the optimisation of the algorithm parameters.

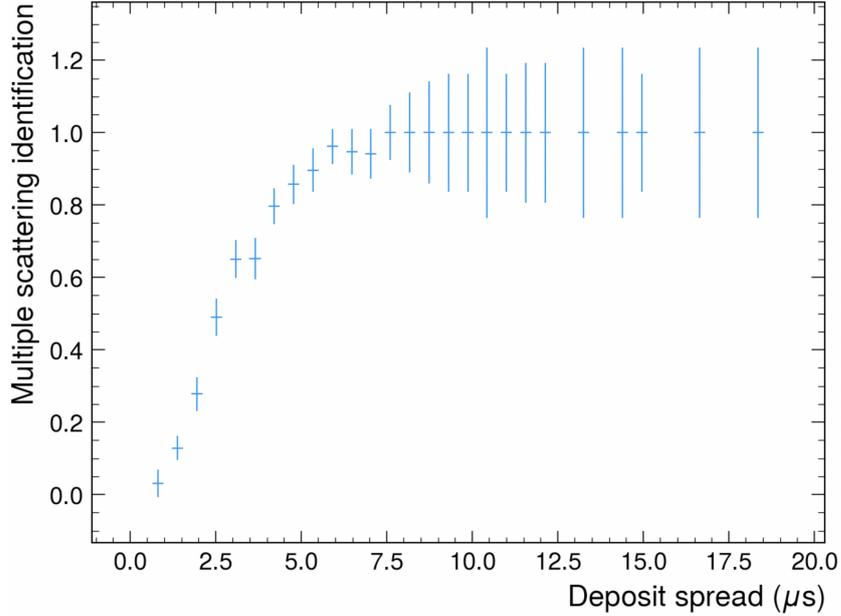


Figure 5.9: Probability of reconstructing more than one peak in multiple S2 pulses as a function of the standard deviation of the times of the true pulses.

The peak finder ability to separate close pulses is of interest for the estimation of the S1-S2 unresolved rate. Figure 5.10 shows an example of an S1 and an S2 pulses that are resolved by the peak finder but not by the pulse finder. The separation capability as a function of the depth of the interaction was measured on ER events generated at the top of the detector, several centimetres below the gas pocket. We evaluated first the contribution from the pulse finder, defined as the probability for the algorithm to identify two separated pulses from the Bayesian Blocks. Then, we compute the improvement provided by the peak finder to reconstruct two peaks when only one pulse is found. The two separation powers are shown in figure 5.11. Since the delay between S1 and S2 is here of the same order as the peak finder binning ($\sim 1 \mu\text{s}$), it appears that its contribution to the separation is marginal.

Finally, regarding computing consumption, it takes on average 0.2 s to

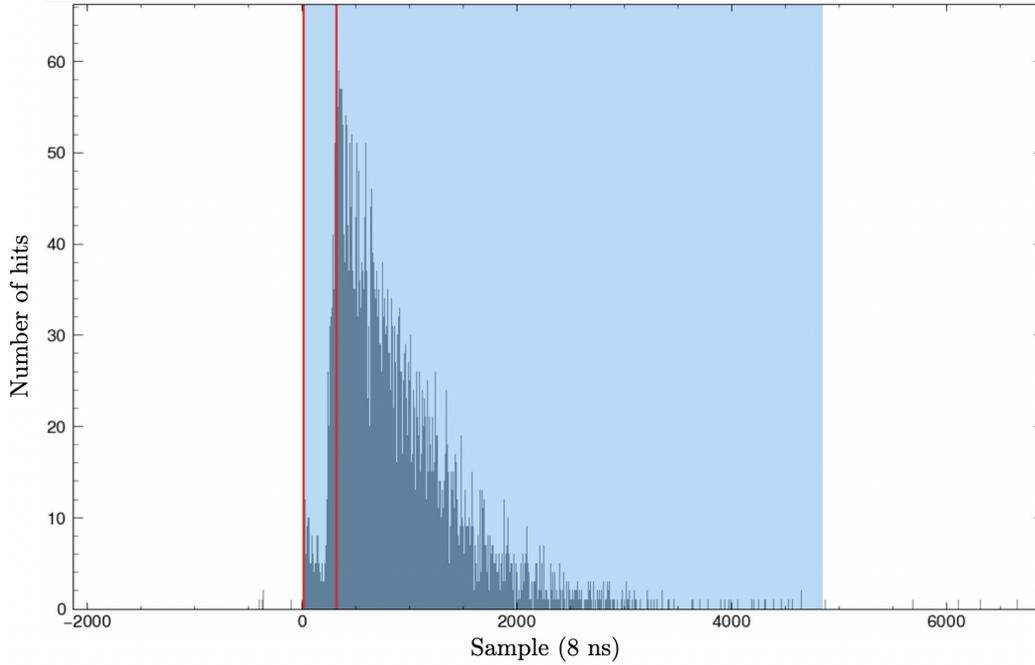


Figure 5.10: An S1-S2 event unresolved by the bayesian blocks but in which two peaks are reconstructed and indicated by the red vertical lines. The blue area is the pulse.

process 1 second of data with this algorithm, which for comparison is more than 10 times less than the Bayesian Blocks.

5.2.3 Low-level pulse reconstruction and classification

The low-level pulse reconstruction determines the integrated variables from the three energy observables, namely the number of hits, the sum of the prominences, and the ZLE charge integral. The time window in which they are computed is fixed at $8 \mu\text{s}$ to contain the full triplet component of S1 pulses. Over this $8 \mu\text{s}$ is also computed f_{200} on the hit distribution. The software provides the possibility of storing these observables integrated over other time windows in order to reconstruct relevant energy information for different pulse natures.

The start time of the pulse provided by the Bayesian Blocks is not precise enough and has to be redefined after the pulse finding. Since no classification is done at this stage, it has to be a generic estimation for both S1 and

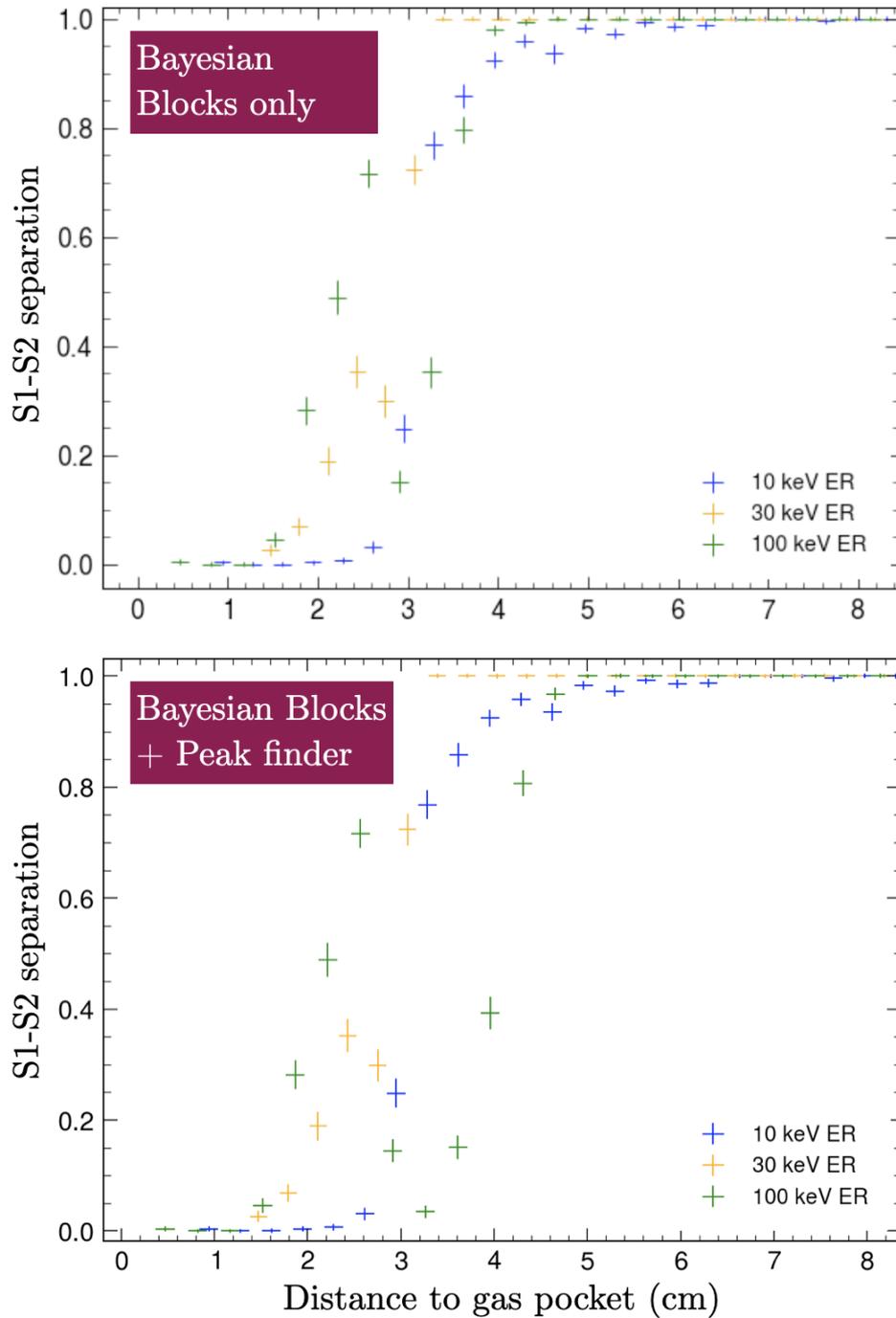


Figure 5.11: Probability to distinguish S1 and S2 as a function of the depth of the interaction, for different energies, with the Bayesian Blocks clustering and with the peak finder. Here, the higher field above the grid is not applied.

S2 pulses. The pulse is first extended by 800 ns in the past and the second percentile of the hit distribution in this window defines the start of the pulse. The algorithm was tested on the full background samples and the reconstructed start time was compared to the true Monte Carlo time of the pulse. The distribution of the difference between reconstruction and Monte Carlo truth are shown in figure 5.12. The obtained timing resolutions are 7 ns and $1.4 \mu\text{s}$ for S1 and S2 respectively. This generic start time is only used for the classification which relies on the PSD estimator f_{200} for which those resolutions are sufficient. We neglect the offset between the true time and the reconstructed since it is due to the scintillation rise time and can if necessary be corrected a posteriori. However this method works poorly for multiple pulses, hence the need for a higher level start time algorithm after classification.

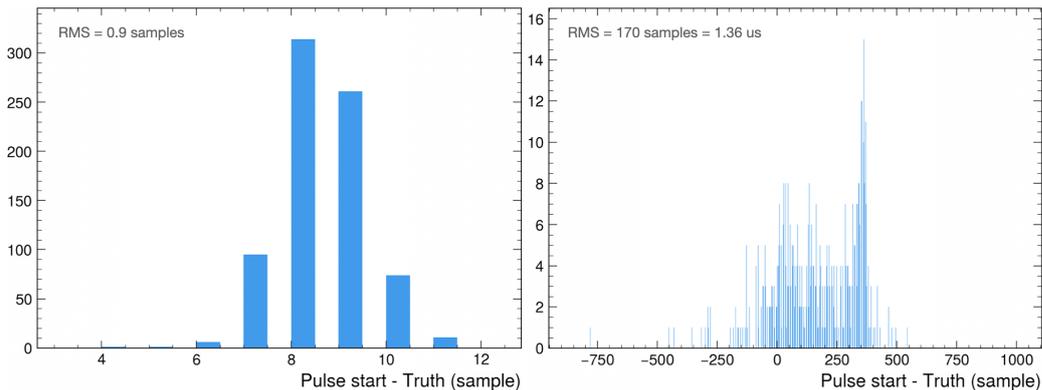


Figure 5.12: Difference between reconstructed start time and true MC time for S1 (left) and S2 (right) pulses.

The classification is done at the last step of the low-level reconstruction. All pulses with more than one peak are tagged as multiple pulses. If the time until the next pulse is shorter than the $8 \mu\text{s}$ window, the integration is not done and the pulse is tagged as truncated. The S1 versus S2 classification relies on the combination of the number of hits N_{hits} and f_{200} . Similarly as for ER and NR, S1 and S2 can be distinguished since the latter has a much lower prompt fraction. The S1 and S2 regions in the (N_{hits}, f_{200}) are indicated in figure 5.13. If N_{hits} is too low, or if the pulse lies between the S1 and S2 bands, it is not classified at this stage, requiring more advanced reconstruction.

5.3 Layer 2 - High-level pulse reconstruction

The layer 2 contains the reconstruction of more advanced variables such as the position of the pulses. The final position of the event is determined: on the z -axis by the drift time based on pulse association; in the xy plane from the ZLEs distribution of S2. However, pulse-wise position reconstruction before the event builder, with a resolution of the order of 50 cm, are needed for S1-S2 association in case of pile-up. For example, in case of two single-scatter events producing a sequence of pulses like S1-S1-S2-S2, the position from S1 and from S2 have to be estimated to disentangle the two events at the event builder level. This will also allow the rejection of random coincidences between uncorrelated S1-only and S2-only events as explained in section 2.2.4.

For S1, the xy position can be evaluated from the hit channel distribution and the z position can be estimated from the top-bottom asymmetry (TBA) of light between the two optical planes. For S2 the xy position is reconstructed from the algorithm described in section 4.4.4 and z is evaluated from the diffusion effect with an estimator based on the hit time distribution. Preliminary algorithms dedicated either to S1 or to S2 pulses were implemented for these purposes in the layer 2.

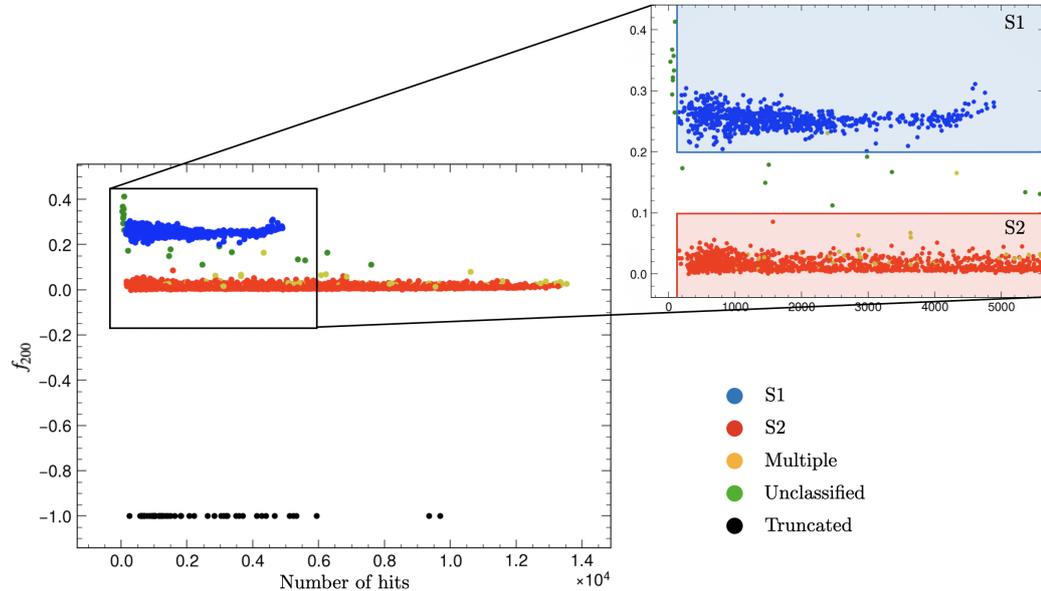


Figure 5.13: Pulse classification over 10 s of full ER background. For tagging, truncated pulses are given a -1 value for f_{200} .

5.3.1 S1 position reconstruction

High spatial resolution is hard to achieve on S1 because of multiple factors: the low number of photons in an S1 pulse, the reflectors which diffuse the light, and the TPB isotropic re-emission. For xy reconstruction, the channels detecting more than 5 hits are selected [72]. This excludes channels only populated by dark noise. The barycentre of the selected channels is computed, taking the number of hits as weight. The use of the hit variable is motivated by the low occupancy in S1 pulses.

The z position of the interaction is also reconstructed from the S1 pulse by exploiting the TBA of the light distribution. The TBA is defined as the relative fraction of light detected in the bottom plane with respect to the total light. It was evaluated with the three usual observables. It appears on figure 5.14 that the ZLE charge provides the best resolution and monotony. The asymmetry of the distribution from top to bottom is due to the reflection of the light on the gas pocket interface. To project z from the measured TBA, the distribution is fitted with this function:

$$z = 1.75 \times \operatorname{erf}(A \times \text{TBA} + B) \quad (5.7)$$

A and B are free parameters with negative values fitted on the ZLE charge distribution. These algorithms achieve around 60 cm resolutions both in xy and z .

5.3.2 z from S2 pulse shape

A first estimator of the diffusion was built by selecting a subset of the hit distribution in S2 pulses. The hits between the third and tenth percentiles of the time distribution are selected. This discarded hits before the start of the pulse which can be affected by some inefficiency of the pulse finder. The tail of the distribution is also cut out as it does not contain any information on the diffusion. The standard deviation of this subset is computed both with and without prominence weights. The impact of weighting the hits with their prominence is small, as it can be seen on figure 5.15, but it increases the resolution of the estimator especially at high z . The distribution is then fitted with the following function:

$$z = A \times \sigma^2 + B \quad (5.8)$$

σ is the standard deviation of the hit subset and A and B are the two free parameters. The result of the function is clipped between -1.75 m and

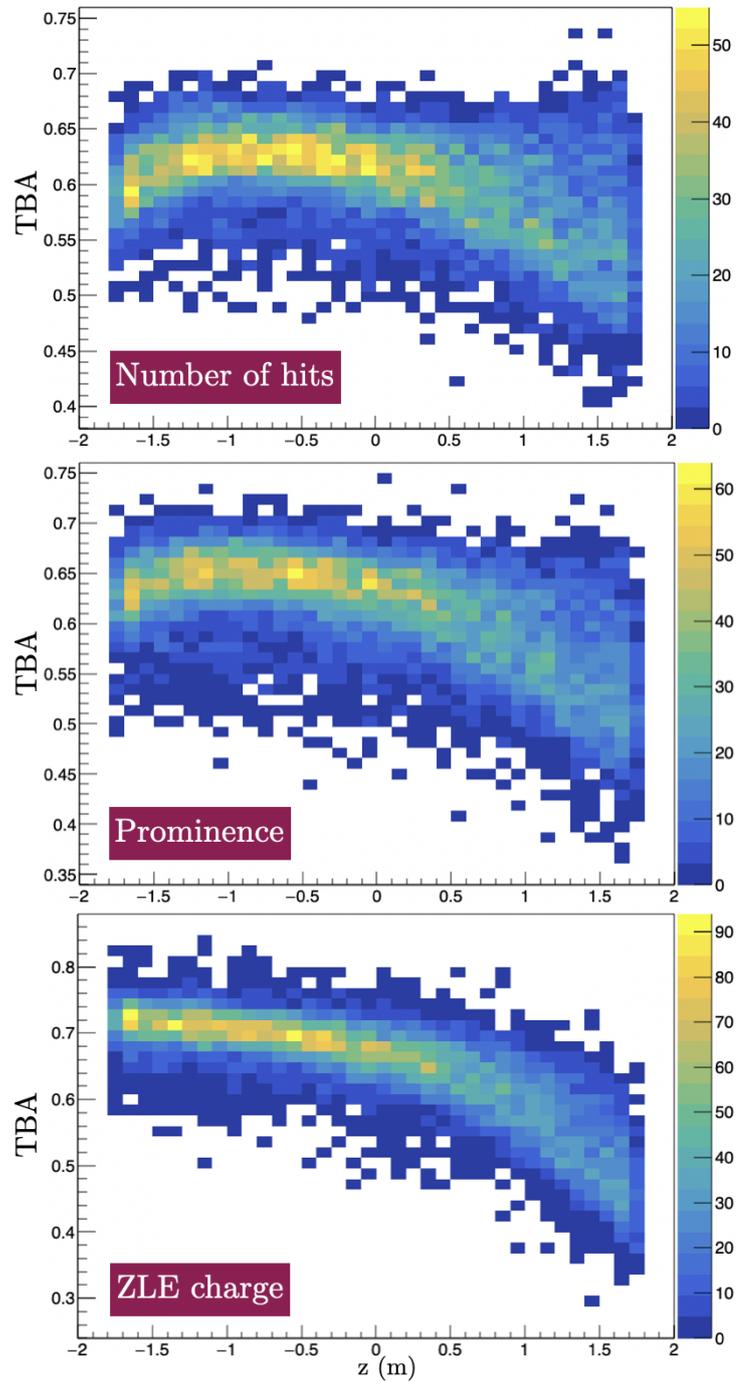


Figure 5.14: TBA as a function of z computed on S1 pulses for the three energy observables.

1.75 m to force the computed position to match the detector geometry. The resolution obtained for the z position from S2 is once again around 60 cm.

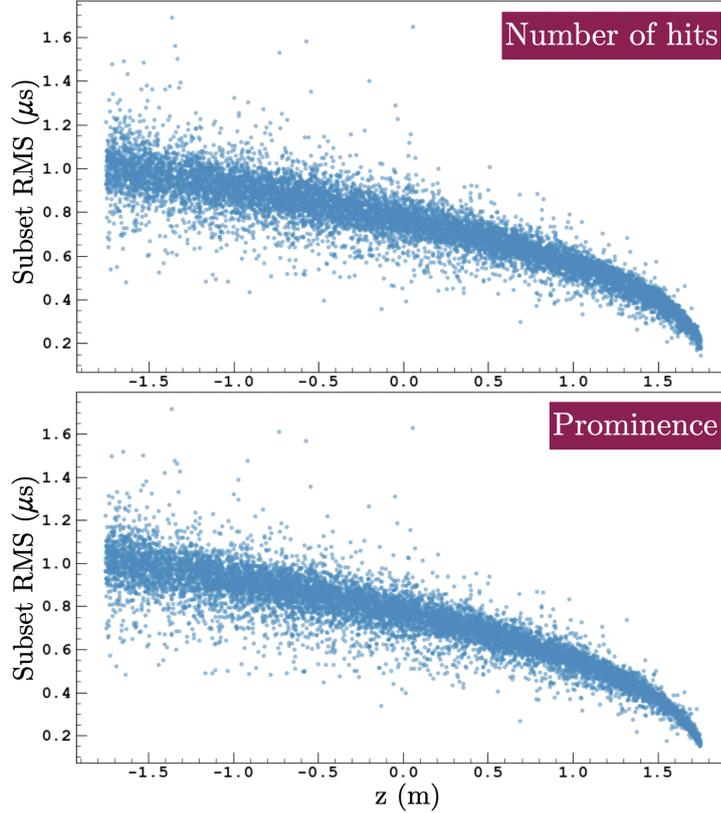


Figure 5.15: Standard deviation of the S2 hit subset, between third and tenth percentiles, as a function of the z position of the interaction with and without prominence weights. This was obtained for ER events uniformly distributed in the TPC, with energy in the ROI.

The resolution that are reached in this scheme are far from what could be done with some more advanced techniques involving for example machine learning. For S1, the use of convolutional neural networks showed potential to reach resolutions in the order of 40 cm [72]. For S2, fitting the hit distribution with parametrised pulse shape could also improve the z resolution from diffusion effect. This framework provides templates for the collaboration to develop these advanced algorithms. It ensures the possibility to work in parallel on the different layers and provides the inputs to start the development of the event builder.

5.4 Modular framework

The evaluation of the computing requirements is crucial to estimate the needs in CPU usage during and after the data-taking. Most of the reconstruction algorithms are either not implemented or not in a final stage, but a preliminary profiling per layer was performed. The required storage was also estimated assuming the current preliminary reconstruction format.

5.4.1 Reprocessing constraints

The reconstruction will be reprocessed multiple times during the 10 years of data-taking. Dedicated machines will be allocated for the prompt reconstruction of the data as soon as they are taken and for the reprocessing during the optimisation of the algorithms. The computing time taken by each layer was evaluated on an Apple M1 core keeping in mind that it is several times faster than typical CPUs in a data center.

The Bayesian Blocks algorithm makes the layer 0 the slowest with around 3 s of processing per second of data. However it only depends on the γ parameter in the fitness of blocks and on the three parameters related to the chunk definition. It is therefore straight-forward to optimise and will not require many reprocessing.

The layer 1 which includes pulse integration, peak finding, and classification takes 0.6 s per live second. The computation of the integrated energy variables and PSD estimator along with the classification is very fast ($\sim 5\%$ of the total time of the layer). The processing time is dominated by the peak finder ($\sim 35\%$) and the selection for each pulse of the corresponding hits and ZLEs from the raw data ($\sim 60\%$).

The implemented preliminary algorithms of layer 2 are very simple and fast. The foreseen computing time of this layer is higher than what can be evaluated at this stage. We indicatively estimated it at 3 s per live second. Streaming the pulse to dedicated algorithms according to their nature will allow to speed up the reprocessing of this layer. We checked that the time required for the reading and writing on disk is negligible at all layers compared to the reconstruction processing.

In order to speed up the reprocessing, we have developed a modular framework that allows the reconstruction of specific layer. For example, it is possible to access the low-level pulse information from layer 1 and apply the

algorithms of layer 2. It is also possible to apply a high-level algorithm of layer 2 on a particular class of pulses, for instance to work on its fine-tuning. The lowest levels of the reconstruction are typically foreseen to be the fastest to optimise and will therefore require less frequent reprocessing.

5.4.2 Reconstruction level formats

In the first draft of the framework, we defined for each layer (except for the event builder) the format of the binary file output, reported in table 5.4.2. Each layer is concatenated to the next, either through direct data streaming or through intermediate files, written at each step and read by the next.

Layer	Variable	Type	Size (bytes)
0	Block left	int32	4
	Block right	int32	4
	Hit rate in block	float32	4
1	Pulse ID	int16	2
	Pulse left	int32	4
	Pulse right	int32	4
	Number of hits	int32	4
	Number of peaks	int32	4
	Pulse start	int32	4
	Time difference with next pulse	int32	4
	Rise time	int32	4
	Pulse classification	int16	2
	Sum of hit prominences	float32	4
	ZLE charge integral	float32	4
f_{200} computed over $8 \mu s$	float32	4	
Peaks position	int32	$4 \times \text{Number of peaks}$	
2	Pulse ID	int16	2
	X for S1	float32	4
	Y for S1	float32	4
	Z for S1	float32	4
	X for S2	float32	4
	Y for S2	float32	4
	Z for S2	float32	4

Table 5.2: Preliminary reconstruction formats.

The software is developed to ensure maximum flexibility for the user to

add additional variables at any level. The custom tools to define, write, and read these files are the same as those used for all the previously presented softwares.

A preliminary estimation of the storage constraints gave a reconstructed data rate of 60 kB/s, dominated by the Bayesian Blocks information (40 kB/s). This will definitely increase with the variables that will be added to the structure during the development of the framework, but it will very likely remain small compared to the raw slice data rate (24 MB/s).

Chapter 6

Sensitivity to light dark matter particles

In this chapter, I present the results of studies on the search for light dark matter particles using liquid argon. This search relies solely on the ionization signal, which allows the energy threshold to be lowered to a few hundred electron volts through amplification. Without the scintillation signal, pulse shape discrimination cannot be exploited, necessitating highly accurate modelling of ER background. The success of this analysis with DarkSide-50 data motivates the assessment of the expected sensitivity of DarkSide-20k in this mass region.

The low-mass analysis of DarkSide-50 demonstrated the potential of this method by setting the world's best limits on WIMPs below $4 \text{ GeV}/c^2$. I describe here the calibration of the argon ionization response which allowed for the first time to go down to 500 eV_{NR} , equivalent to a few ionization electrons. Then, I present the development of the background model and my contribution to the production of the DarkSide-50 limits.

This work provided the response model and the framework from which the projected sensitivity of DarkSide-20k was established. The latter is described in the last section along with my contributions to the background model and the evaluation of the single electron resolution using the software tools presented in the previous chapters.

6.1 Ionization response

The ionization yield of ER is measured using intrinsic ^{37}Ar and ^{39}Ar contamination in DarkSide-50 and down to few numbers of electrons. For NR, the calibration relies on the AmBe and AmC neutron spectra in DarkSide-50 and on external datasets from beam experiments.

6.1.1 Electronic recoil response

The response is based on the Thomas-Imel model which describes the recombination probability as a function of energy and electric field. It relies on the assumption that the electron-ion pairs are uniformly contained in a box [73]. The properties of the box are described by a single parameter γ which appears in the following expression of the charge yield Q_y^{ER} , namely the number of electrons escaping recombination per unit of energy:

$$Q_y^{ER} = \frac{1}{\gamma} \frac{\ln(1 + \gamma N_i)}{E_{er}} \quad (6.1)$$

Here, N_i is the number of produced electron-ion pairs and E_{er} is the total ER energy. The model was calibrated on ^{39}Ar and ^{37}Ar from DarkSide-50 data. The ^{39}Ar spectrum is obtained from the atmospheric argon run of DarkSide-50 during which it dominates the data rate. ^{37}Ar is a cosmogenic isotope decaying through electron capture with a half-life of 35 days. The electron capture of ^{37}Ar induces cascades emitting electrons and X-ray lines for a total of 2.8 keV in the K-shell and 277 eV in the L1-shell in average.

The fit with the Thomas-Imel model is in very good agreement with the data at low energy, from 0.1 to 3 keV $_{ER}$. However, figure 6.1 shows that an extra term is needed to reproduce the data at energies in the order of 10 keV $_{er}$. Equation 6.1 becomes:

$$Q_y^{ER} = \left(\frac{1}{\gamma} + p_0 (E_{er}/\text{keV}_{ER})^{p_1} \right) \frac{\ln(1 + \gamma N_i)}{E_{er}} \quad (6.2)$$

Here, p_0 and p_1 are the two free parameters of the model extension whose values are obtained from the fit of ^{37}Ar and ^{39}Ar .

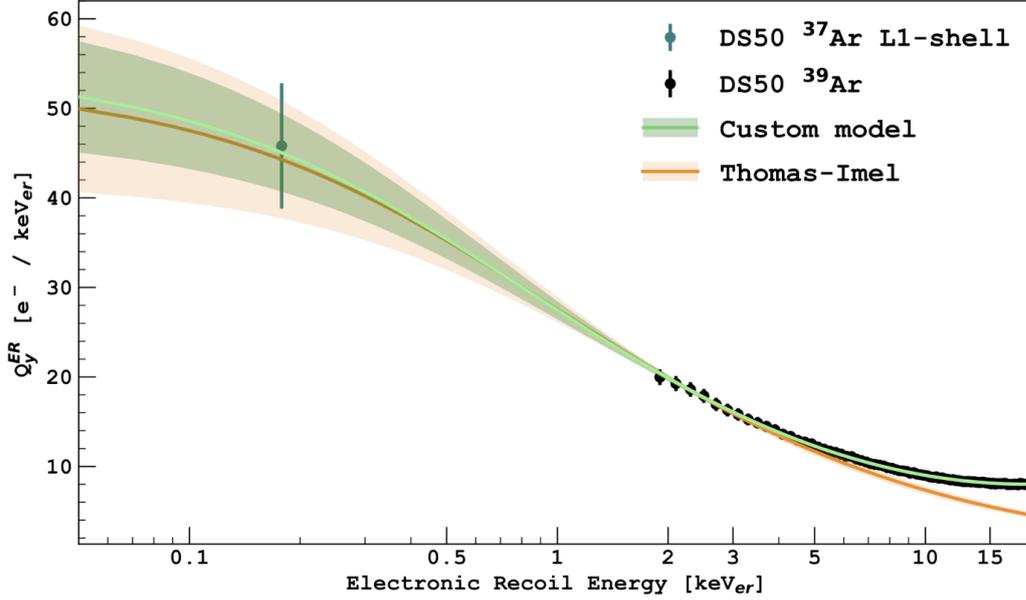


Figure 6.1: Fits of the ER custom response extended from the Thomas-Imel model.

6.1.2 Nuclear recoil response

The same formalism is applied for NR, also relying on the Thomas-Imel box model. In addition, the stopping power of the nucleus s_n and the excitation rate of electrons s_e have to be described as they appear in the expression of N_i :

$$N_i = \beta \frac{\epsilon s_e(\epsilon)}{s_e(\epsilon) + s_n(\epsilon)} \quad (6.3)$$

Here, β is a normalisation constant and ϵ is the energy parameter which depends on the properties of the material. For argon, its value is $\epsilon = 0.0135 E_{nr}/keV$. For $s_e(\epsilon)$, the dependency is described as the square root of ϵ with a pre-factor related to the atomic and mass numbers of the target (0.145 for argon). The nuclear stopping power function used for the analysis is the Ziegler et al. universal screening function:

$$s_n(\epsilon) = \frac{\ln(1 + 1.1383 f_Z \epsilon)}{2(f_Z \epsilon + 0.001321(f_Z \epsilon)^{0.21226} + 0.19593(f_Z \epsilon)^{0.5})} \quad (6.4)$$

A conversion factor f_Z is needed to account for differences in the definition of ϵ in the Ziegler et al. function [74]. Its value in argon is 0.953. The normalisation constant and the box parameters are fitted on neutron sources

calibration data of DS-50 and external data sets from ARIS and SCENE, another dual-phase TPC for argon response characterisation (figure 6.2) [75]. Several screening functions were tested, showing marginal differences on the charge yield.

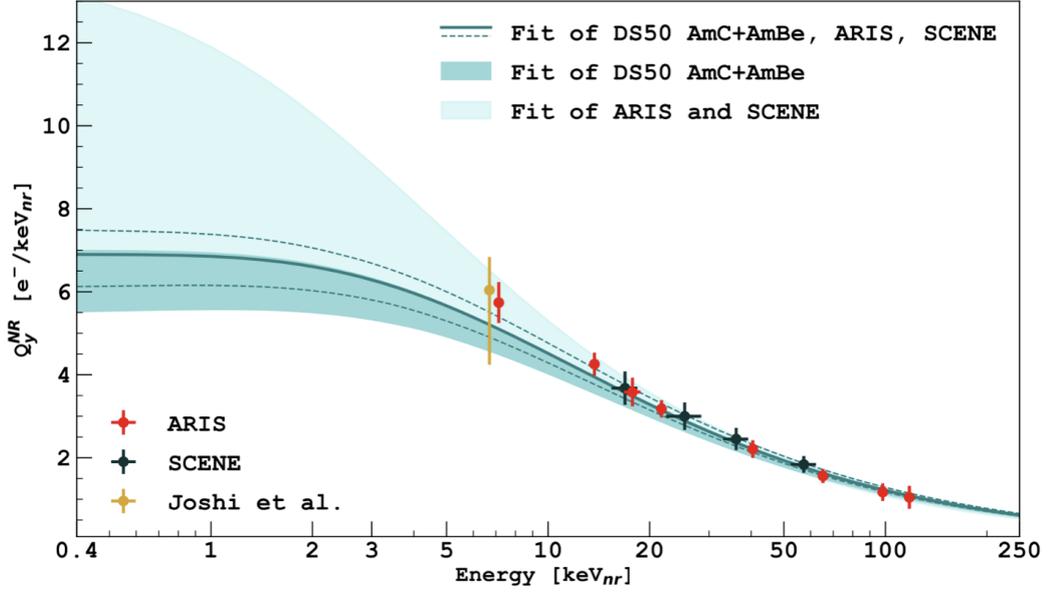


Figure 6.2: NR response model fitted on calibration data. The dark blue area is a fit from AmC and AmBe in DarkSide-50 while the light blue area is an extrapolation from ARIS and SCENE. The dark blue line is a combined fit of the DarkSide-50 neutron spectra and ARIS and SCENE data.

This work provides a complete ionization response model for both ER and NR from 180 eV_{ER} to 500 eV_{NR} with a field of 200 V/cm . This corresponds to 3 ionization electrons which is the lowest calibration ever performed in liquid argon.

6.2 Light dark matter signal models

The goal of the S2-only analysis is to produce limits on light dark matter candidates. The signal models for these candidates is described here, starting from the NR energy distribution of low-mass WIMPs. In the case of NR signal, the stochastic fluctuations of quenching are unknown, therefore two extreme cases will be considered. The inclusion of ER in S2-only data opens the path to competitive sensitivity in the dark matter-electron channel [76].

Signal models were developed for leptophilic dark matter candidates, namely particles inducing ER.

6.2.1 Low-mass WIMPs

The signal from WIMPs was produced for masses in the GeV/c^2 range, assuming the Standard Halo Model described in 1.2.1. The NR energy distributions as a function of the mass are computed and can be given as an input to the detector model to convert energy in number of detected electrons. The expected signal in DarkSide-50 for a given cross-section is shown on figure 6.3.

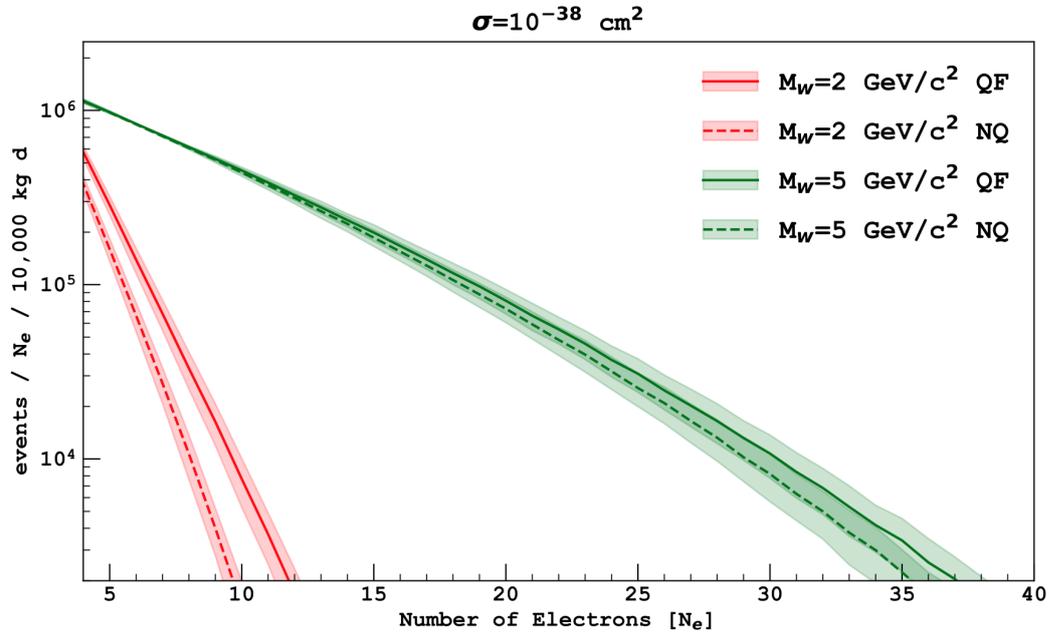


Figure 6.3: Signal model of low-mass WIMPs inducing NR in DarkSide-50, in the absence of quenching fluctuation (NQ) and with binomial quenching fluctuation (QF). The error bands are computed from the uncertainty on the liquid argon ionization response by producing the spectrum with a charge yield 1σ above and below the nominal value.

Upon applying the response model of the detectors, it must be taken into account that the quenching in liquid argon is a stochastic process from which only the expected value is known. Still, the fluctuations of the quenching around this value are unknown and have considerable impact on the response at low energy. The signal model for NR is therefore computed for the

two extreme cases, the case in which there is no fluctuation at all, and the case in which the quenching follows a binomial distribution which maximises the fluctuation.

6.2.2 Leptophilic signal

In the ER channel, the WIMP signal shape strongly depends on a form factor $F_{DM}(q)$ which describes the mass of the mediator of the interaction, q being the momentum-transfer. Two cases are considered, one is the case of a light mediator where the form factor scales as $(\alpha m_e/q)^2$ with α the fine structure constant and m_e the electron mass. In the case of an heavy mediator whose mass is much larger than αm_e , the form factor is constant with the momentum. The signal for the two considered mediators are compared in figure 6.4.

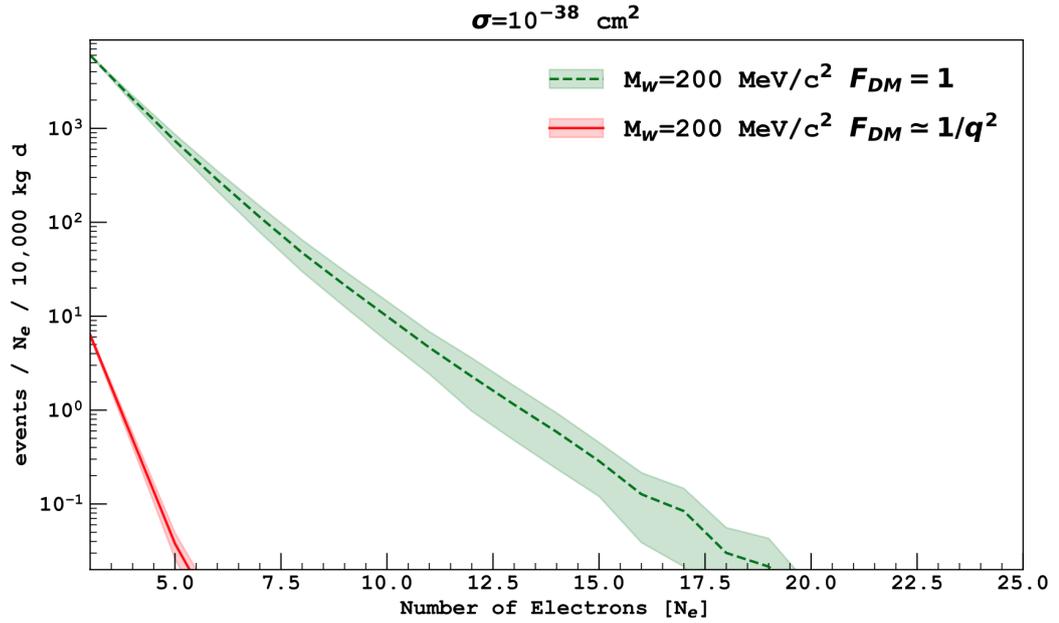


Figure 6.4: Signal model of WIMPs inducing ER in DarkSide-50, with a heavy mediator ($F_{DM} = 1$) and light mediator ($F_{DM} = 1/q^2$).

Models considering pseudo-scalar dark-matter such as axion-like particles or vector-boson dark matter like dark photons were tested. Those candidates are expected to provide a mono-energetic signal at the particle's rest mass. The expected signal from sterile neutrinos with mass between 7 keV and 36 keV can also be computed. The detection channel is in two steps with first

the mixing of the sterile neutrino into an active state and then the collision with an electron of the argon cloud. This allows to constrain the mixing angle $|U_{e4}|^2$. The expected signals for a given mixing angle and with different sterile neutrino masses are shown in figure 6.5.

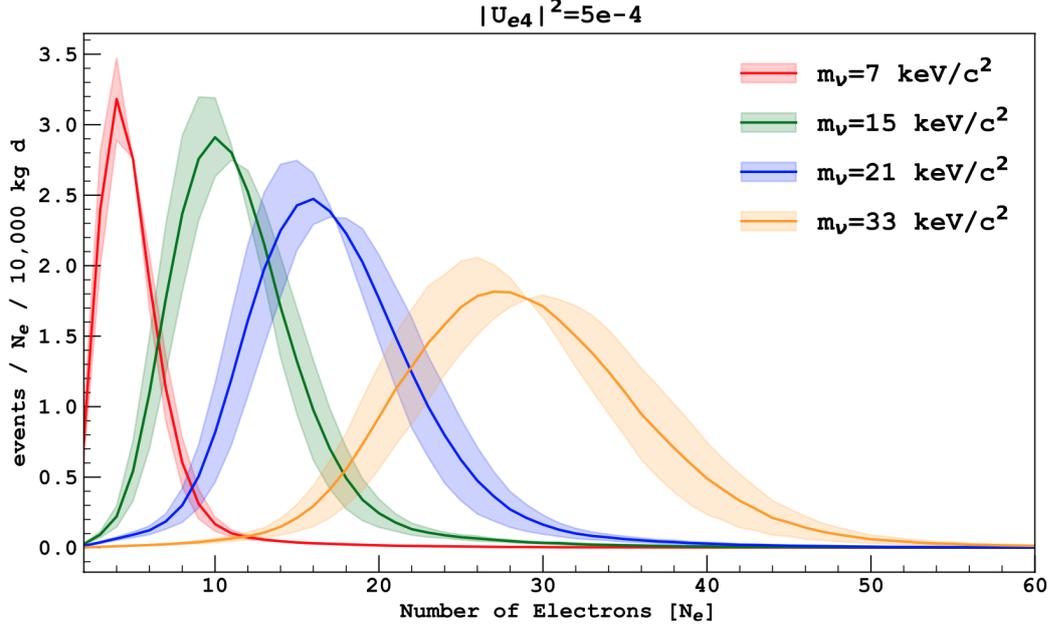


Figure 6.5: Signal model of sterile neutrinos in DarkSide-50. The error bands are computed with a charge yield 1σ above and below the nominal value.

6.2.3 Migdal effect

The low-mass analysis was extended to account for an atomic effect expected to enhance the ionization signal, in particular at low-energy [77]. The so-called Migdal effect describes the possibility for the target atom to be ionised during the nuclear recoil. This can be pictured by considering the inertia of the electronic cloud. The result is a probability of emitting additional electrons for NR events. The impact of the Migdal effect on the signal spectrum is reported on figure 6.6.

The Migdal effect was predicted in 1939 but has still never been observed, although it is strongly supported by atomic physics calculation [78]. Some experiments have tried to demonstrate its existence such as the MIGDAL

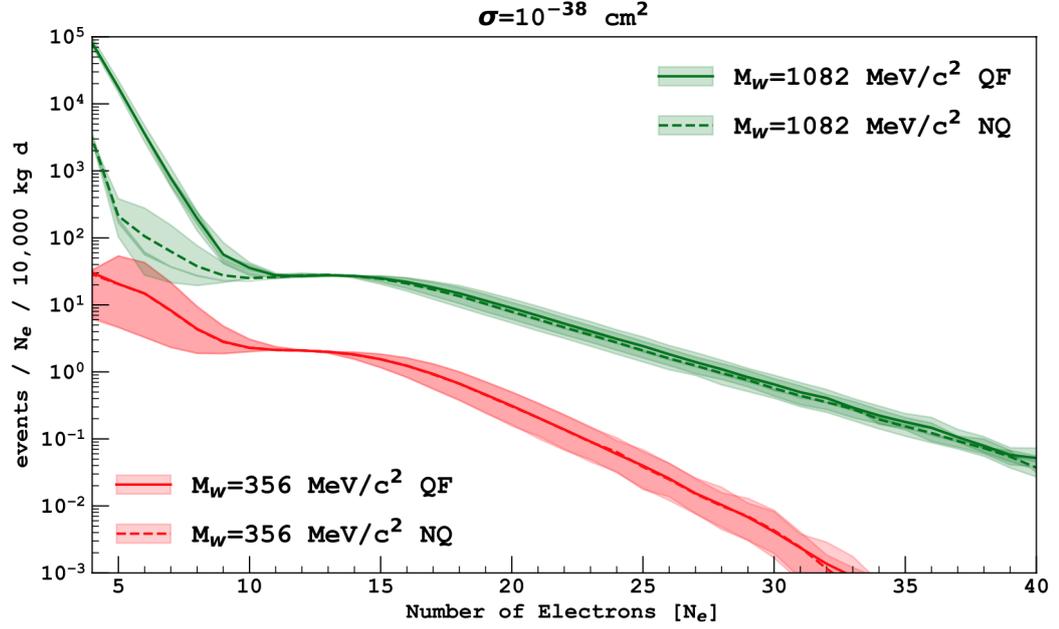


Figure 6.6: Signal model of WIMPs inducing NR when considering Migdal effect, in DarkSide-50. The error bands are computed with a charge yield 1σ above and below the nominal value.

(Migdal In Galactic Dark Matter Exploration) collaboration [79].

6.3 DarkSide-50 low-mass analysis

The analysis based on the ionization signal in DarkSide-50 is described here. The dataset used for this work consists in 653.1 live days. Limits were obtained for all the previously described signals from light dark matter candidates. In this context, I contributed to the evaluation of the impact of each systematic uncertainty source.

6.3.1 Data selection

The selection is made requiring events with a single S2 pulse. Since the detection efficiency for S1 photons is 0.16 ± 0.01 , events below 3 keV_{er} , are mostly S2-only [80]. Because of the low number of photons in this energy region, the reconstruction algorithm fails to identify the xy position of the event. Therefore, the position used for fiducialisation is simply defined as

the position of the top PMT being the largest fraction of light.

Like in the standard analysis described in 2.2.1, the prompt fraction f_{90} is used to discriminate between S1 and S2. S2 pulses are selected by imposing $f_{90} < 0.1$. A set of quality cuts is applied to reject pile-up of multiple S2s or of S1 and S2. To reject the firsts, the S2 pulse must be contained in 100 μs and the peak time of the pulse with respect to the trigger must be less than 6 μs . To reject overlapping S1 and S2, the peak time must be longer than 200 ns and the FWHM of the light distribution must be larger than 100 ns.

α events from the lateral walls of the TPC were observed to induce large S1 pulses with anomalously low S2. This is due to the absorption of the ionization electrons of the α interaction in the walls suppressing the S2 signals from these events. In addition, the scintillation light extracts some electrons from the cathode which are drifted and interpreted as a low energy S2 signal. A specific cut was applied on the subsample of events which include an S1 pulse. These events are rejected if the S2/S1 ratio is 2.58σ away from the mean value, ensuring a 99% acceptance.

Another class of background at low number of electrons arises from the so-called spurious electrons (SE). The mechanism behind this background is still not well understood, but it was demonstrated that they are correlated with previous events and that their probability increases with the amount of impurity in argon. They are partially rejected by imposing a 20 ms minimum time difference between consecutive events.

6.3.2 Background fit and systematics

The events remaining in the region of interest are mostly internal contamination from ^{39}Ar and ^{85}Kr in liquid argon, and radioactive isotopes in the material surrounding the active target. The activities of ^{39}Ar and ^{85}Kr are fitted from S1 distribution. For the latter which is a β^- emitter with a Q-value of 687 keV, it is also possible to validate its activity through the fast coincidence with γ emission from the ^{85m}Kr metastable state.

The background model is built by summing the spectra of internal background with external background from the PMTs and the cryostat obtained from simulations. The model depends on several nuisance parameters listed in table 6.1. These include parameters describing the atomic exchange and screening effects affecting ^{39}Ar and ^{85}Kr forbidden unique β -decay [81]. These

	Name	Source	Affected components
Amplitude	A_{FV}	Uncertainty on the fiducial volume	WIMP, ^{39}Ar , ^{85}Kr , PMTs, Cryostat
	A_{Ar}	14.0% uncertainty on ^{39}Ar activity	^{39}Ar
	A_{Kr}	4.7% uncertainty on ^{85}Kr activity	^{85}Kr
	A_{pmt}	11.5% uncertainty on activity from PMTs	PMT
	A_{cryo}	6.6% uncertainty on activity from the cryostat	Cryostat
Shape	Q_{Kr}	0.4% uncertainty on the ^{85}Kr -decay Q-value	^{85}Kr
	Q_{Ar}	1% uncertainty on the ^{39}Ar -decay Q-value	^{39}Ar
	S_{Kr}	Spectral shape uncertainty on atomic exchange and screening effects	^{85}Kr
	S_{Ar}	Spectral shape uncertainty on atomic exchange and screening effects	^{39}Ar
	Q_y^{er}	Spectral shape systematics from ER ionization response uncertainty	^{39}Ar , ^{85}Kr , PMTs, Cryostat
	Q_y^{nr}	Spectral shape systematics from NR ionization response uncertainty	WIMP

Table 6.1: List of systematics, their sources, and impacted signal and background components included in the binned profile likelihood.

effects impact the spectra at low energy and induce additional systematic uncertainties. The nuisance parameters are fitted with a binned profile likelihood defined as follow:

$$\mathcal{L} = \prod_{i=0}^N \mathcal{P}(n_i | m_i(\Theta)) \prod_{\theta_i \in \Theta} \mathcal{G}(\theta_i^0 | \theta_i, \Delta\theta_i) \prod_{i=0}^N \mathcal{G}(m_i^0 | m_i(\Theta), \delta m_i(\Theta)) \quad (6.5)$$

The first factor is the Poisson probability of observing n_i events in the i^{th} bin of the spectrum where the expected number of event m_i is expressed as a function of the set of nuisance parameters Θ . The second factor is a Gaussian penalty that increases with the distance between a nuisance parameter θ_i and its prior value θ_i^0 in order to decrease the likelihood of fitted parameters values that are too far from the expected ones. The standard deviation of this gaussian terms are the uncertainties on the prior nuisance parameter $\Delta\theta_i$. The last factor accounts for statistical uncertainties of the Monte Carlo simulated samples from which the model is derived.

Figure 6.7 shows the fit of the data with the background model only. An excess below 4 electrons appears due to the inefficiency of the SE cut. The threshold for the analysis is therefore set at $4 e^-$ corresponding to about 0.05 keV_{er} and 0.4 keV_{nr} . The background data are produced with a binning of $0.25 e^-$. Figure 6.8 shows that the choice of binning has minor impact on the fitted values of the parameters. All fitted values are compatible within 1σ with their prior values. The distribution is in very good agreement with the Monte Carlo spectrum. This demonstrates the robustness of the background model.

In order to understand the individual impact of the systematics, I evaluated the contribution of each group of uncertainties to the expected limit.

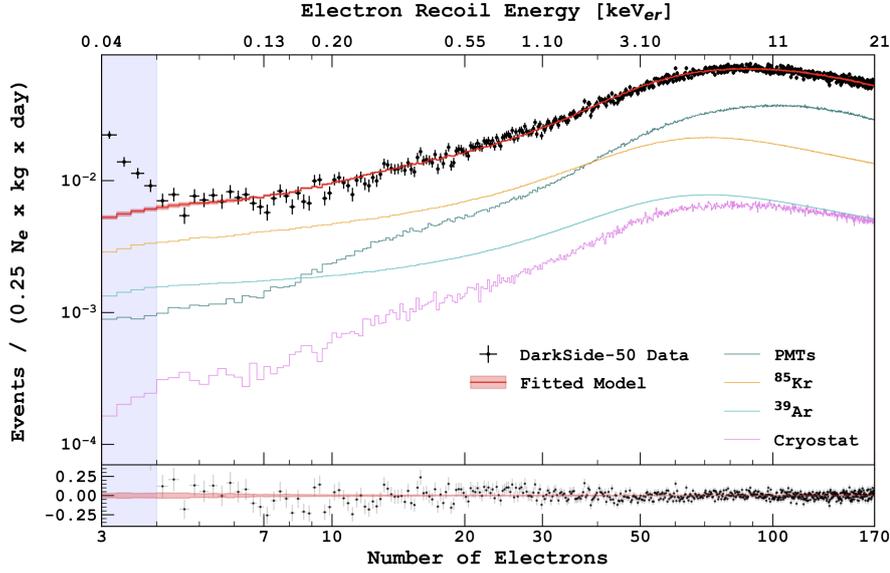


Figure 6.7: Background-only fit between 4 and 170 e^- with contributions of PMTs, cryostat and internal contaminations. The bottom plot shows the residuals defined as the difference between observed and expected number of events, normalised by the expected number of events.

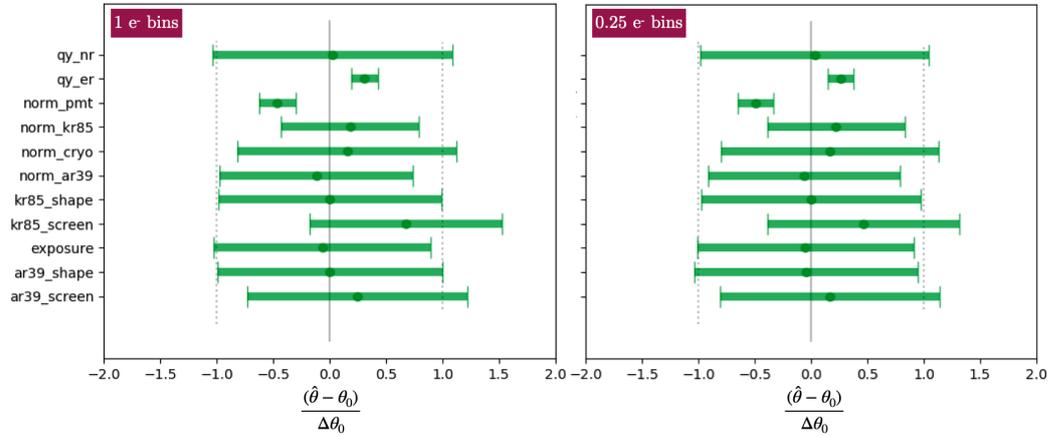


Figure 6.8: For two different binnings, distances between the fitted nuisance parameters $\hat{\theta}$ and their prior values θ_0 normalised by the uncertainty on the latter.

As an example, this is reported on figure 6.9 in the case of the limit for a WIMP of $0.3 \text{ GeV}/c^2$.

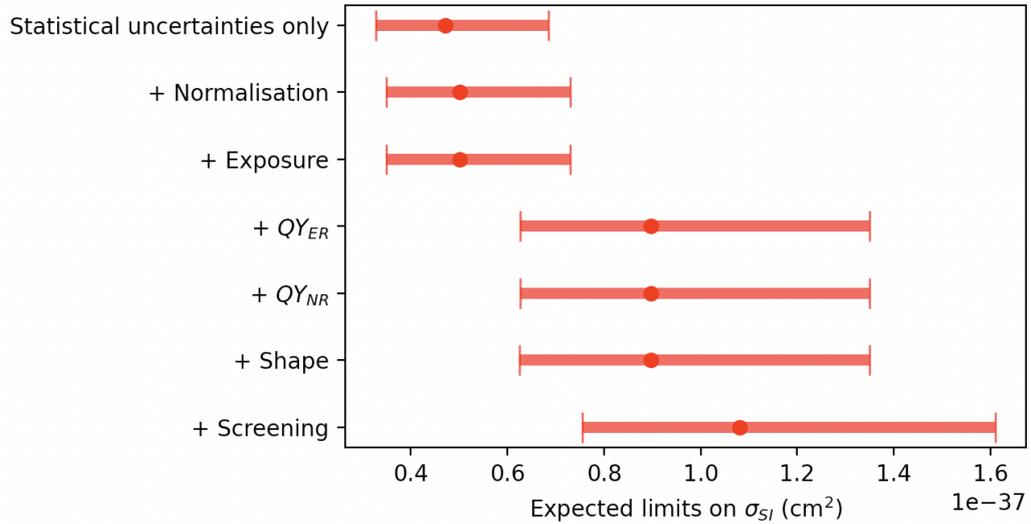


Figure 6.9: Impact of each category of systematic to the expected limit for a WIMP mass of $3 \text{ GeV}/c^2$ and including Migdal effect. "Normalisation" denotes the activities of the PMTs, cryostat, and internal contaminants. "Shape" regroups the uncertainties of the Q-values and shape parameters of the internal contaminants.

6.3.3 Limits

The signal from spin-independent WIMP-nucleon scattering is added to the model and fitted for different WIMP masses. The maximum cross-section with which each mass point is compatible within 90% confidence level is obtained from the data.

As it appears in figure 6.10, with higher quenching fluctuation, the limits are improved. This is because quenching fluctuations induce an additional smearing of the energy spectrum. In particular, at low energy, some physically below threshold events would be detected above the threshold and some above threshold events would fall below. Since the signal spectrum decreases with energy, more events pass from below to above threshold than the opposite, increasing the expected number of events.

Figure 6.11 shows the improvement of the limits when including Migdal effect in the signal model. In the absence of direct observation of the Migdal effect, its impact on the sensitivity to low-mass WIMPs is put aside of the analysis, as an extension.

The limits produced for leptophilic candidates are reported in figure 6.12. Some regions of the phase space were for the first time excluded for axion-like particles and for electron interacting WIMPs. The obtained limits for sterile neutrinos are not competitive with indirect detection limits, but still constitute the first limits derived from a direct-detection experiment.

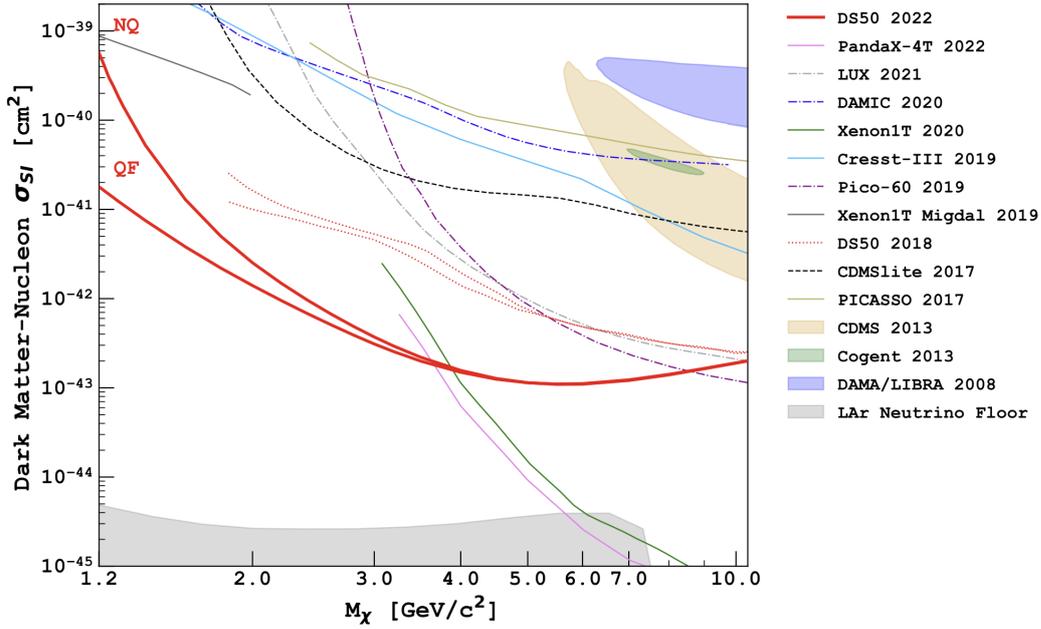


Figure 6.10: Spin-independent DM-nucleon cross-section 90 % confidence level in the low-mass region for DarkSide-50 in the case of binomial quenching fluctuation (QF) and in the absence of quenching fluctuation (NQ) compared with other experiments. The coloured areas are claimed discoveries. The grey area is the neutrino floor in liquid argon inside which CE ν NS background from solar neutrino cannot be neglected anymore.

6.4 DarkSide-20k low-mass projected sensitivity

The projected sensitivity of DarkSide-20k on light dark matter was obtained on the basis of the techniques developed for DarkSide-50 [82]. Here, I describe the analysis and my contributions to the development of the background model.

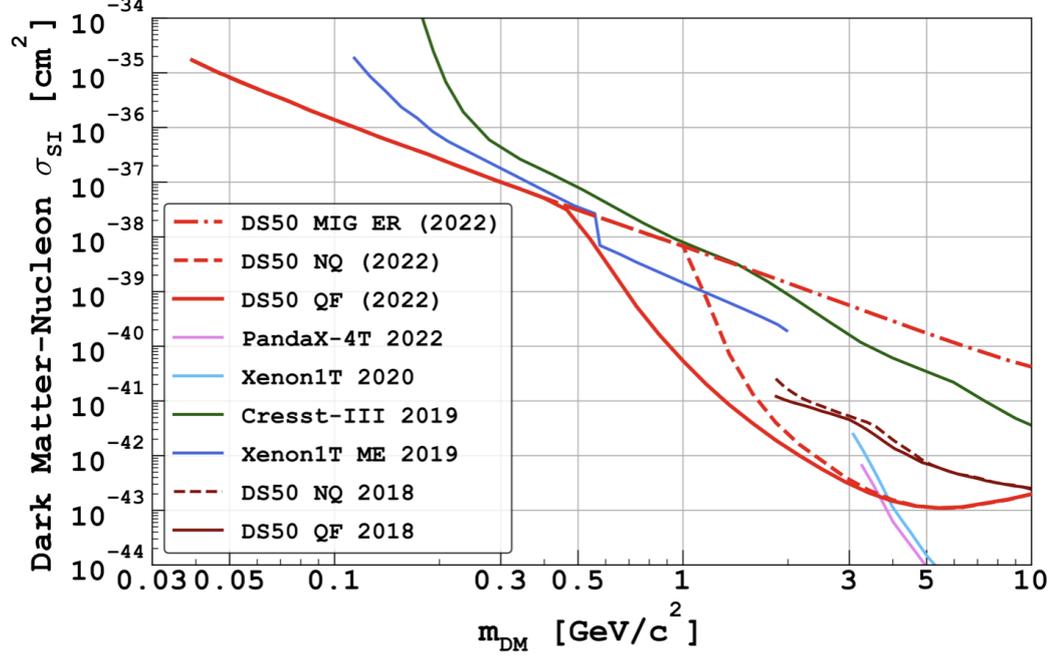


Figure 6.11: Improvement of the limits on low-mass WIMPs when including Migdal effect in the signal model.

6.4.1 Expected background

The response model of the detector is based on the previously described ionization response. The fluctuations of the number of electrons for ER are implemented following the description in section 3.2.1 and the two extreme cases of binomial fluctuation and no fluctuation are assumed for the quenching of NR. The electron loss along the drift is applied assuming a 16 ms electron lifetime as measured in DarkSide-50 data. The electron lifetime is the average time an electron drifts before being absorbed by some impurity.

The background spectra were produced with G4DS. For each isotope in each detector location, I simulated 10^8 events. The activities used for the normalisation of the rate are quoted in table 3.1. No fiducialisation is performed along the vertical drift as the energy is too low to detect the S1 pulse. Therefore, the activity from the SiPMs on the top and bottom edges of the TPC cannot be rejected and dominates the background. Ten times more events (10^9 instead of 10^8) were simulated in the optical planes to reduce statistical uncertainty from this background.

With respect to DarkSide-50, the external background level per unit of

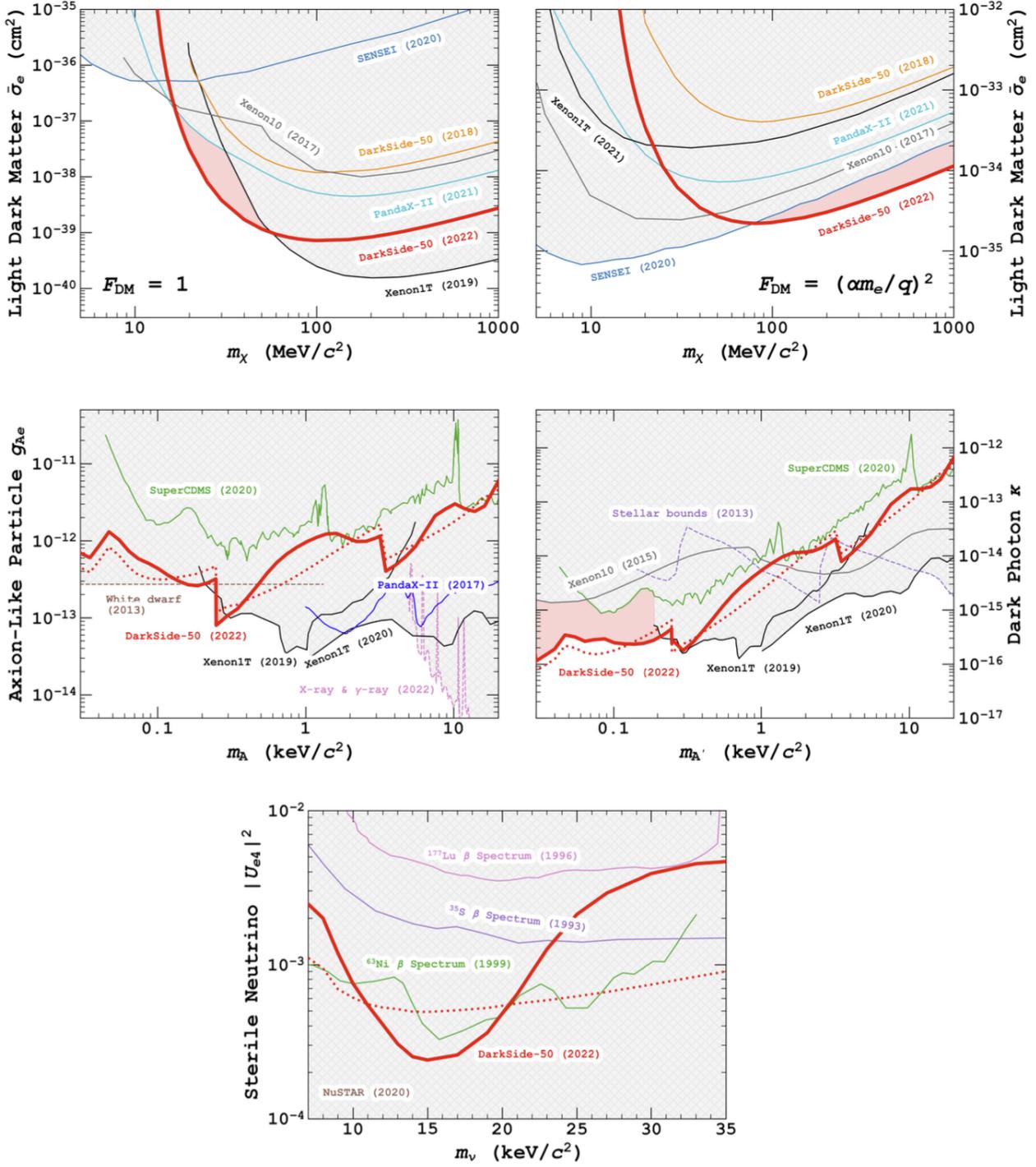


Figure 6.12: Limits from DarkSide-50 on light dark matter cross-section with heavy mediator (top left) and light mediator (top right), axion-electron coupling strength g_{Ae} (center left), photon and dark-photon kinetic mixing κ (center right), and sterile to electronic neutrino mixing angle $|U_{e4}|^2$ (bottom).

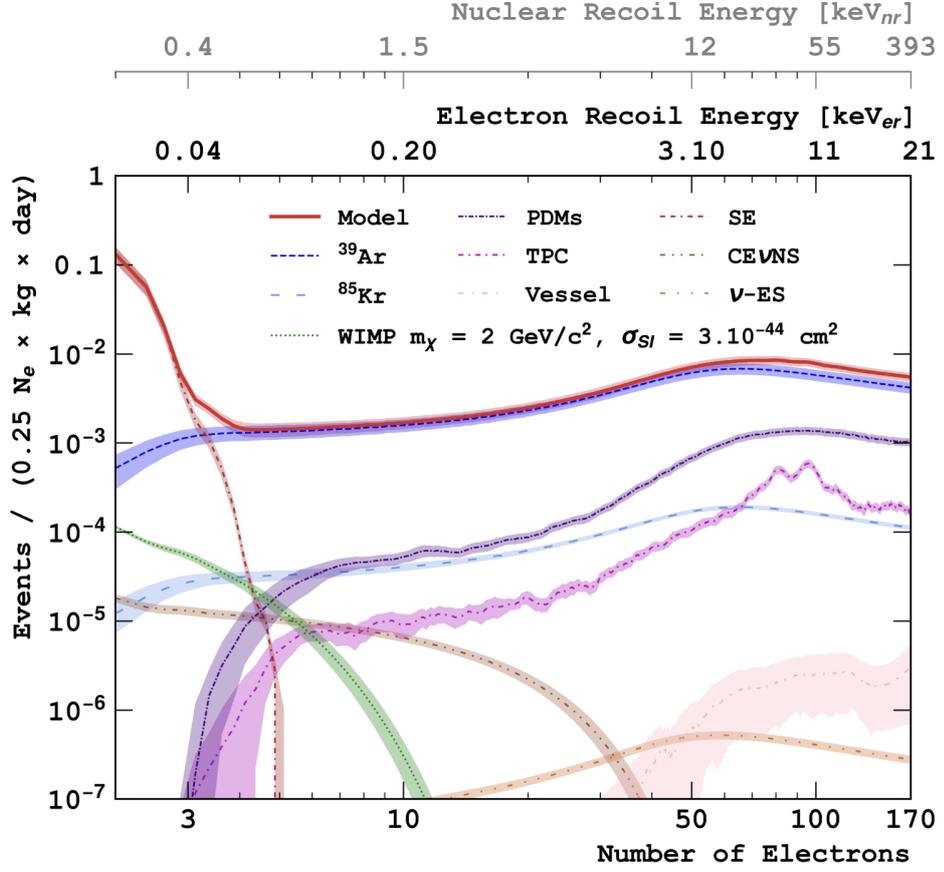


Figure 6.13: Background distribution summed with a signal corresponding to the highest cross-section excluded by DarkSide-50 at $2 \text{ GeV}/c^2$.

fiducial argon mass is reduced by a factor 22. This is due to the lower surface-to-volume ratio and to the purity of the material in DarkSide-20k. The neutrino background from $\text{CE}\nu\text{NS}$ and neutrino-electron interactions are also included in this analysis, although their contributions are sub-dominant. The contribution of neutrons from (α, n) reactions and spontaneous fission is negligible and therefore not included in the background model.

A model of the SE background from DarkSide-50 data was developed. The SE spectrum of DarkSide-50 was fitted with a Poisson convoluted with a Gaussian to account for the resolution effects of DarkSide-50. The amplitude and resolution are then rescaled to the response of DarkSide-20k. We here assume two cases: the SE is well assessed and the threshold is lowered down to 2 electrons to constrain it; the threshold is set at 4 electrons above

which spurious electrons become sub-dominant.

The fiducialisation in the xy plane is done by removing events that are less than 30 cm away from the lateral walls, assuming a 3 cm resolution from the position reconstruction algorithm. The fiducial mass after this cut is 34.2 tons. The final background spectrum is shown in figure 6.13.

6.4.2 Single electron resolution

Instrumental effects on the sensitivity are embedded in the S2 resolution. This was evaluated by means of the electronic simulation and the reconstruction tools presented in sections 4.1.3 and 5.2.3. I measured the expected resolution by generating single electron pulse uniformly distributed in the xy plane as a function of g_2 , the mean number of photo-electron produced per electron extracted in gas. Its nominal value is 25 PE/e⁻ at the centre of the TPC but a different value would impact the resolution, and therefore the shape of the energy spectrum. I simulated pulses for different values of g_2 and computed the integrated charge for the ZLE integrals. The relative resolution of the charge distribution is presented in figure 6.14. In the response model, the single electron resolution is fixed at 23%.

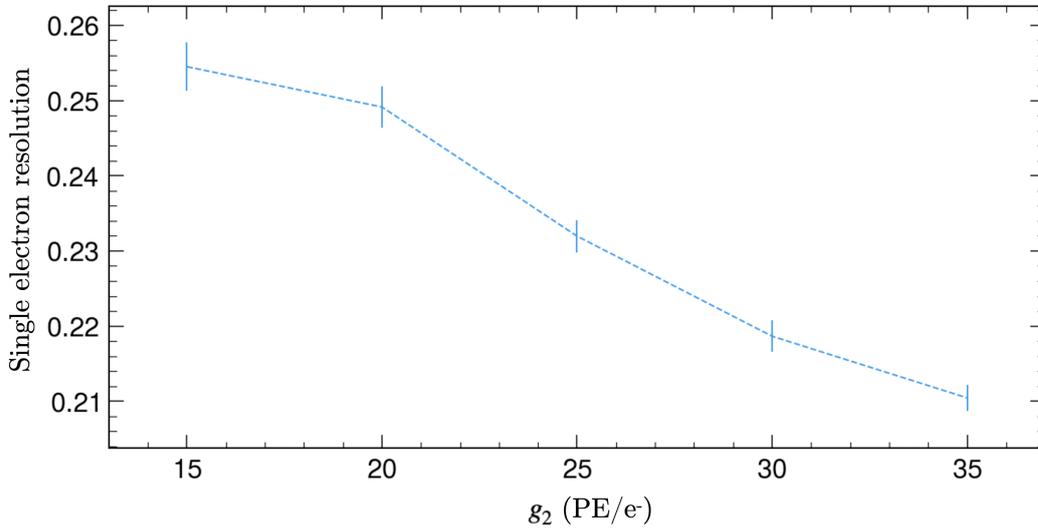


Figure 6.14: Single electron relative resolution from charge variable as a function of g_2 the amplification factor for S2. The relative resolution is defined as the ratio between the standard deviation and the mean from a gaussian fit.

6.4.3 Live time estimation

The live time is reduced by the cut applied to reject multiple scatter events. In order to reject them, we veto a window around each S2 pulse backward and forward in time of the maximum drift time (3.7 ms). This guarantees that the S2 pulse that triggers the vetoed window is isolated and hence single scatter. For this study, we also assume that S2 pulses within $2 \mu\text{s}$ in time are indistinguishable.

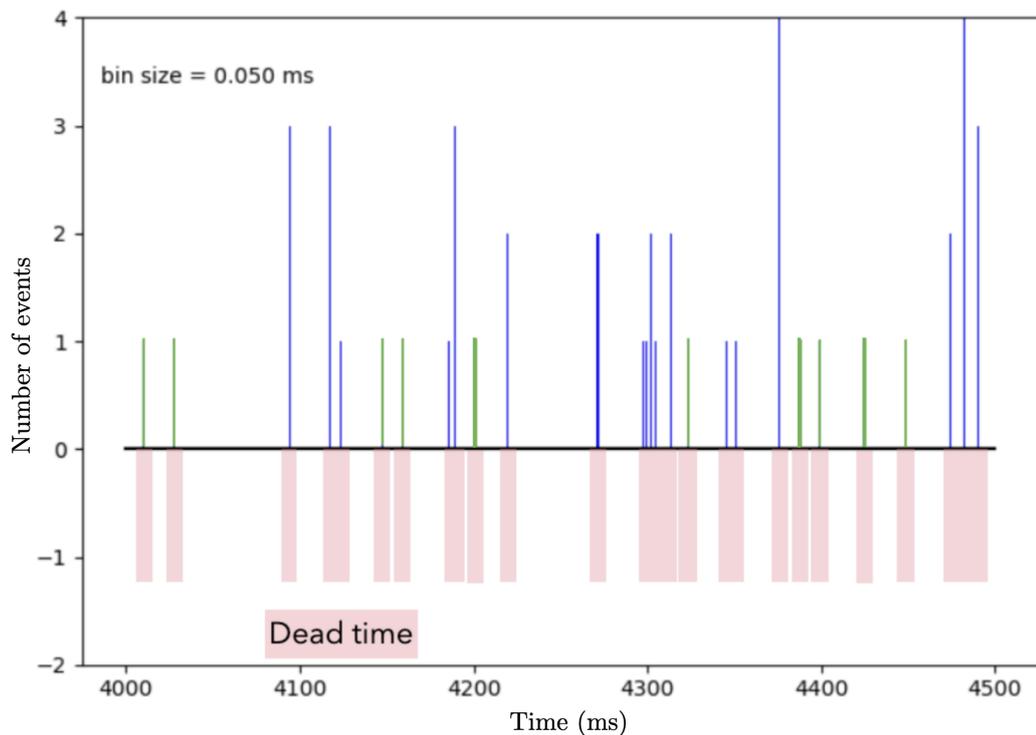


Figure 6.15: 500 ms time slice with S2 pulses simulated from the background model. Red area are the time windows opened before and after each pulse. Green peaks are isolated S2 pulses, selected for the analysis. Blue peaks are pulses less than 3.7 ms away from another, discarded by the cut.

To calculate the live time, we used the background model as input of the event mixer, presented in 3.1.2. This produced time series of S2 pulses selected in the region of interest as shown in figure 6.15. This cut is made under the approximation that the dark matter signal rate is very small with respect to the background rate and hence does not contribute to the dead time. The obtained live time is 51% of the total data-taking time. In the

future, the z position of the events will be evaluated from diffusion estimators of the S2 pulse. This will allow to perform a less stringent selection and increase the effective live time.

6.4.4 Expected limits

The fitting procedure is based on the binned profile likelihood of equation 6.5. The same systematics as in table 6.1 were used. In addition, the normalisation on the neutrino background is included with a 10% systematic uncertainty. The impact of the single electron resolution on the expected limits was tested showing sub-percent differences for g_2 between 20 and 30 PE/ e^- .

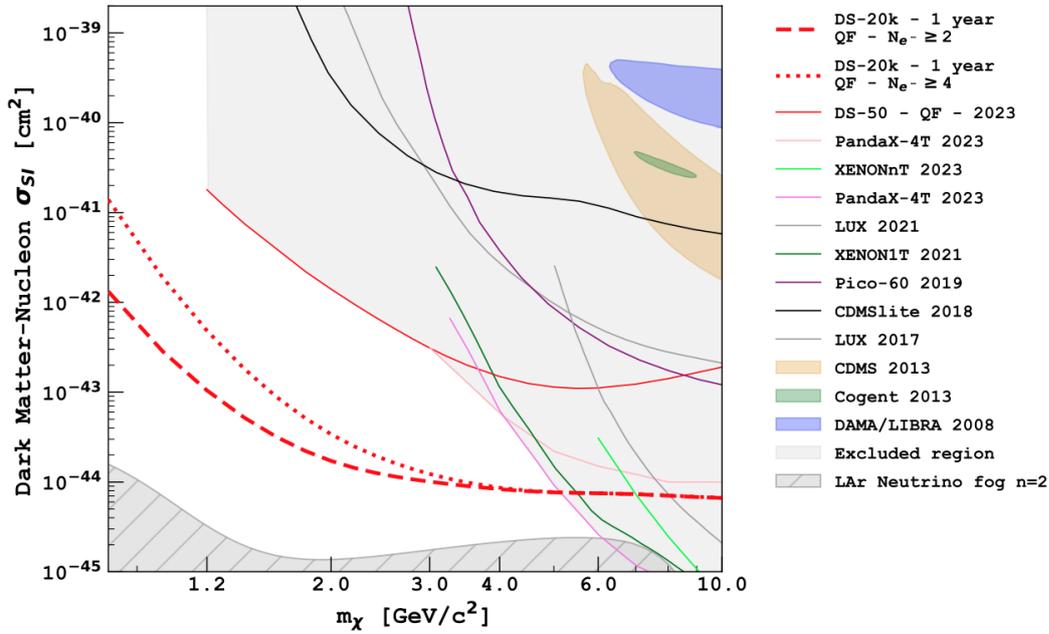


Figure 6.16: DarkSide-20k expected limits at 90% confidence level after 1 year of exposure with the spurious electron models (dashed red) and without (dotted red). Binomial quenching (QF) fluctuation is assumed.

Similarly to the previous analysis, the WIMP signal is tested for different masses. The limits are derived with and without the spurious electrons model with the threshold at 2 electrons and 4 electrons respectively. The results are showed on figure 6.16. In only one year of data-taking, DarkSide-20k will reach 20 to 40 times stronger limits than DarkSide-50. With the signal

models previously described, the limits to leptophilic dark matter and with NR WIMPs with Migdal effect were produced and are showed on figure 6.17.

By adapting the strategy of DarkSide-50, this work demonstrates that DarkSide-20k will very likely be a world leading experiment in the light dark matter region. This is mostly due to the 1000 times larger target volume, the high light yield, and the material radio-purity. It also highlights that lowering the threshold to constrain the spurious electron background would provide an improvement by a factor ten of the limits for WIMPs of $1 \text{ GeV}/c^2$.

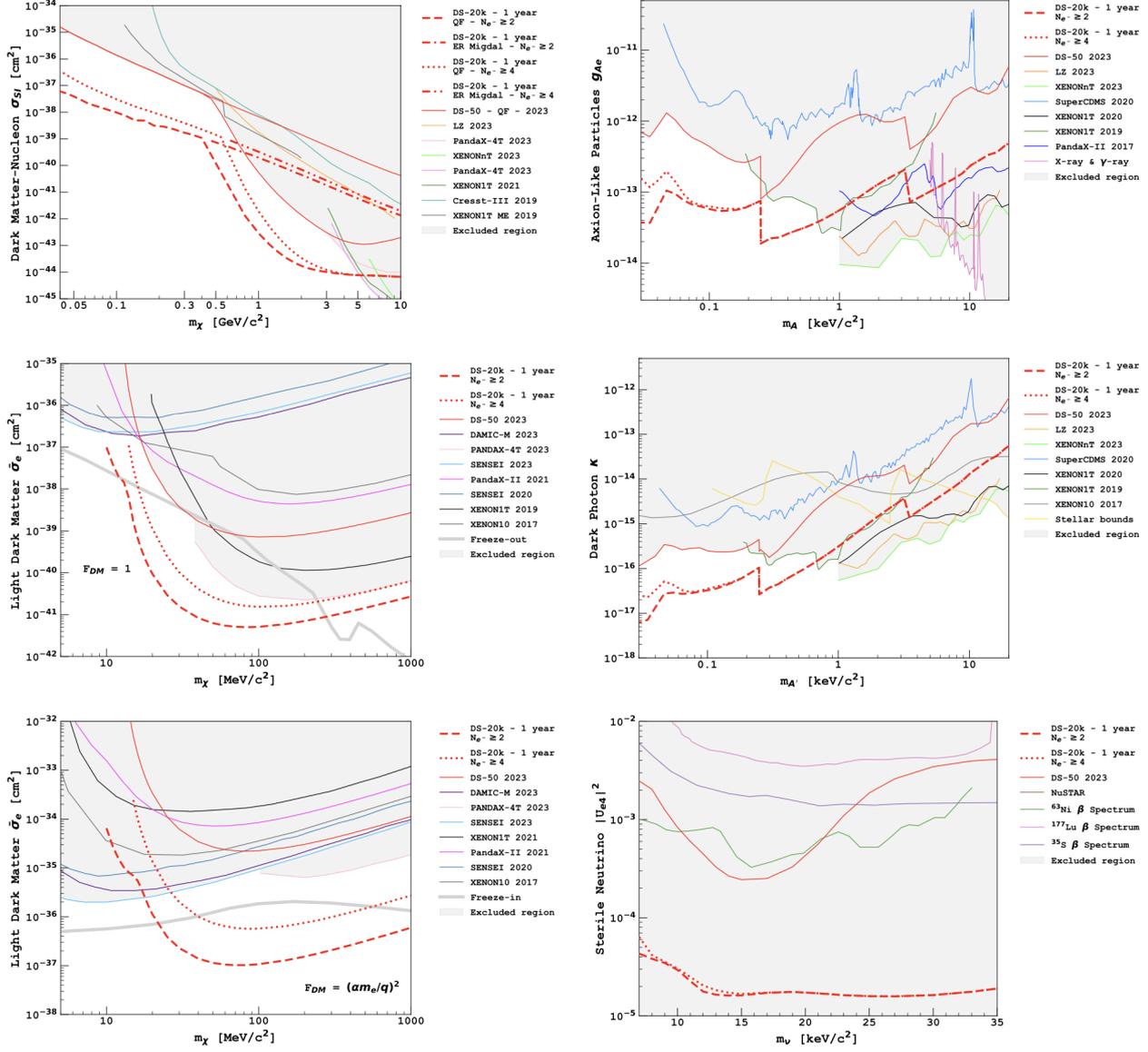


Figure 6.17: DarkSide-20k sensitivity to WIMP inducing NR considering Migdal effect (top left) and to leptophilic dark matter candidates.

Chapter 7

Argon response to low energy nuclear recoil

The success of the analysis of DarkSide-50 and the projected sensitivity of DarkSide-20k in the light dark matter search suggest to deepen the understanding of the target responses to NR in the sub-keV region for improving the sensitivity. In particular, two main caveats remain in the NR ionization model described in the previous chapter: the indetermination of the quenching fluctuation and the large uncertainties in the sub-keV region. A measurement of the argon response to NR at energies below 1 keV_{NR} with monochromatic lines would allow to develop a more robust model. This can be achieved using a $\mathcal{O}(100 \text{ keV})$ neutron beam. Figure 7.1 shows that this will give access to NR energies below 1 keV .

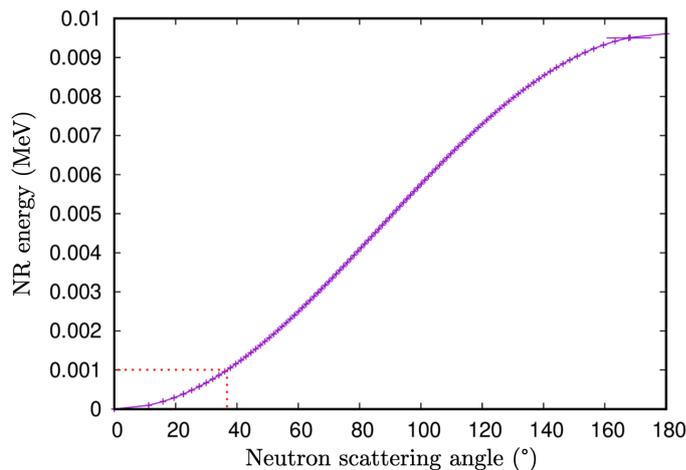


Figure 7.1: Accessible NR energy in argon from a 100 keV neutron beam.

In this chapter, I describe the first steps toward such challenging measurement. Similarly to ARIS or SCENE, it will consist in a liquid argon TPC exposed to a mono-chromatic neutron beam inducing NR interactions. The TPC will be surrounded by detectors aiming at measuring the scattering angle of the neutron as illustrated in figure 7.2. From the kinematics, the energy of the NR in the TPC can be computed. No scintillation signal is expected to be detected in the TPC at so low energy. Therefore, the TPC events must be identified through the time coincidence between the beam and the neutron detector, requiring it to provide $\mathcal{O}(1 \text{ ns})$ time resolution.

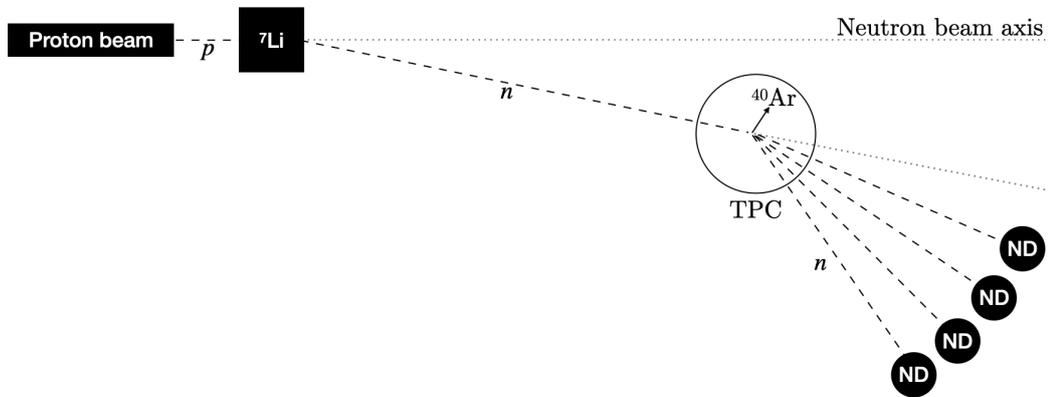


Figure 7.2: Scheme of the low energy NR measurement with neutron beam from the ${}^7\text{Li}(p, n){}^7\text{Be}$ reaction.

Three days of data-taking were allocated in December 2023 for two purposes. The first is the characterisation of the ALTO neutron beam based on the ${}^7\text{Li}(p, n){}^7\text{Be}$ operating at $\mathcal{O}(100 \text{ keV})$. This characterisation consists in measuring the angular profile of the neutron flux and the energy spectrum, as well as the background induced by the beam. The second goal is to characterise the performance of neutron detectors sensitive to $\mathcal{O}(100 \text{ keV})$ neutrons. Plastic scintillators, like those used in ARIS, cannot detect neutrons at such energy. Two types of neutron detectors based on neutron capture were tested for this purpose.

The time resolution and background rates of two designs of neutron detectors were evaluated during these three days. From the results, I established a preliminary projection of the sensitivity that could be reached with the future measurement.

7.1 Setup for beam and detector characterisations

The ALTO (Accélérateur Linéaire et Tandem à Orsay) platform, located at the Nuclear Physics Institute of Orsay, provides a directional proton source from a Tandem accelerator. A Lithium-Fluoride (LiF) target converts it into the neutron beam that will be used for the future measurement.

At the low energy involved in this measurement, most of the neutron detectors rely on thermalisation before capture. However, this process takes $\mathcal{O}(10\text{-}100\ \mu\text{s})$ which is not suitable in the case of a time-of-flight measurement. Therefore, we propose neutron detectors based on capturing neutron on the fly, avoiding thermalisation. We used two neutron capture targets, ^{10}B and ^6Li , to evaluate their performances and to characterise the beam.

7.1.1 Low energy neutron beam

The production of the neutron beam is based on the $^7\text{Li}(p, n)^7\text{Be}$ reaction. For the first time, this reaction was tested with neutron energies down to 150 keV. The Tandem accelerator, an electro-static Van de Graaf particle accelerator, is shot on a fixed ^7Li target which consists in a $4\ \mu\text{m}$ foil made of an alloy of LiF and copper. The threshold of the $^7\text{Li}(p, n)^7\text{Be}$ reaction is around 1.8 MeV. So a proton beam around 2 MeV is expected to provide a $\mathcal{O}(100\ \text{keV})$ neutron beam, considering the energy loss of the proton in the LiF target.

The ARIS measurement was also done in ALTO but the reaction was inverted. The fixed target was a gaseous hydrogen cell shot by a ^7Li beam accelerated in Tandem. The advantage of this $p(^7\text{Li}, n)^7\text{Be}$ reaction is that the neutron beam is very collimated. However the energy threshold of the reaction does not allow to produce neutrons below $\sim 1.5\ \text{MeV}$. While the $^7\text{Li}(p, n)^7\text{Be}$ reaction gives access to lower energies, the collimation of the beam is not as good, it is expected to be contained in a 45° cone in the forward direction as shown in figure 7.3.

Upon the reaction, protons interaction with oxygen in the target cell induce the emission of γ s with energies up to $\sim 10\ \text{MeV}$. This so-called γ flash induce background in the neutron detectors but is also used to compute the time-of-flight as it gives a precise information on the reaction time.

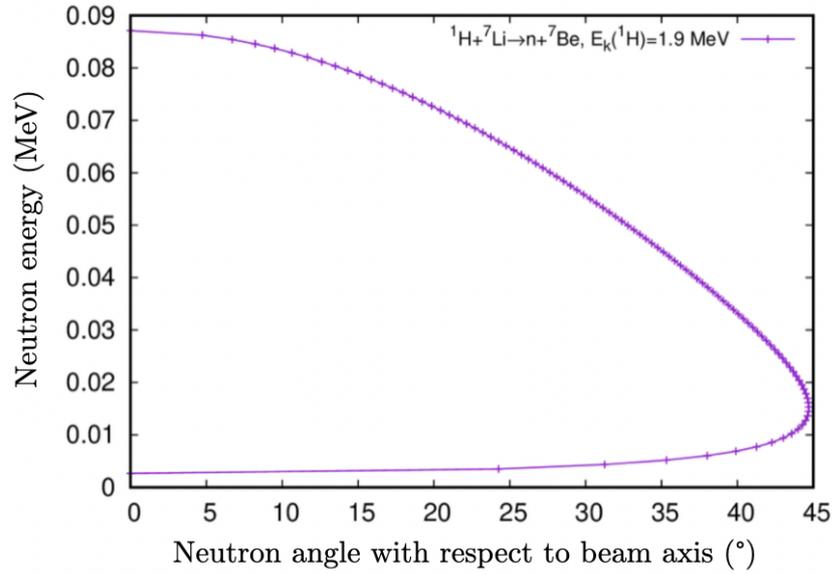


Figure 7.3: Expected neutron beam profile from ${}^7\text{Li}(p, n){}^7\text{Be}$ reaction with 1.9 MeV proton.

7.1.2 ${}^{10}\text{B}$ -based neutron detector

A ${}^{10}\text{B}$ based design was developed and tested for the neutron detection. ${}^{10}\text{B}$ is a stable isotope that has a 10% abundance in natural boron and a very high cross section for neutron capture of 2 barn at 100 keV. The signature for the neutron capture is the emission of an α and a 478 keV γ :



The detector consists in an aluminum capsule filled with ${}^{10}\text{B}$ enriched powder. The capsules are 2 cm long cylinders with a diameter of 14 mm and 1 mm aluminum thickness. They are filled with approximately 3 g of ${}^{10}\text{B}$. The principle is to detect the emitted γ while the α will be absorbed in the target. To do that, the capsule is surrounded by an inorganic crystals in which the light is read by a photo-multiplier tube (PMT). The choice of inorganic crystal is driven by the need for the scintillator to be transparent to neutrons.

Two scintillation crystals were tested, barium fluoride (BaF_2) which has a very fast scintillation component with a characteristic time of 1 ns. The BaF_2 detector is shown on figure 7.4. The other one is a sodium-iodine crystal doped with Thallium ($\text{NaI}(\text{Tl})$). This one has a slower scintillation (250 ns) but a light yield of 38 photon/keV, higher than the BaF_2 (10 pho-

ton/keV). Not only a higher light yield ensures better energy resolution but it also improve timing reconstruction as the rise time is easier to identify when the distribution is more populated. The two detectors were calibrated using radioactive sources to measure their gains and energy resolutions.

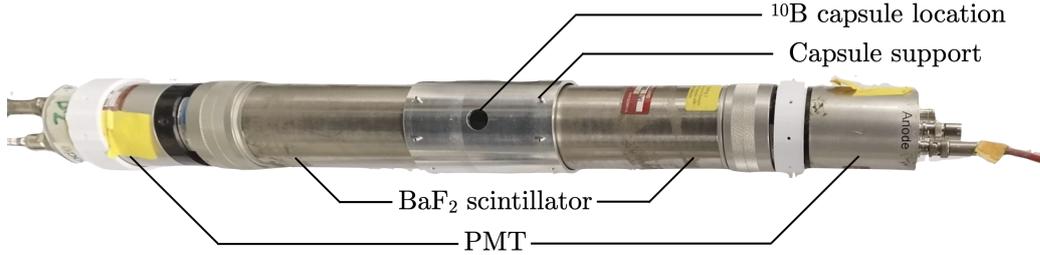


Figure 7.4: Picture of the ^{10}B based neutron detector with BaF_2 scintillators. Two scintillators are placed around the ^{10}B capsule and each of them is read-out with a PMT.

An online energy trigger selection is applied around the 478 keV band corresponding to the γ energy of the neutron capture reaction. This was designed to reject background that was dominating the trigger rate coming from environmental and beam correlated γ s and internal contamination.

7.1.3 ^6Li -Glass neutron detector

The Li-glass detector is the state-of-the-art technology for this kind of purpose. It consists in a scintillating glass doped with ^6Li , aiming to detect the following reaction:



The signature of this reaction is the scintillation produced by the α particle in the glass. The detector is a cylinder with a 6 cm diameter and 1 cm thickness coupled to a PMT. Another identical detector, doped with ^7Li in which there is no neutron capture reaction is also used to measure the background contamination.

Due to a resonance, the neutron capture cross-section of ^6Li shows a peak centred around 240 keV [83]. This peak is due to a P-wave resonance. The collision is the superposition of the plane wave of the incoming neutron and a fixed spherical potential from the ^6Li target. The plane wave and the target are characterised together by a quantum angular momentum l , which is

analogous to the impact parameter in the classical case, in the sense that if it is too large considering the range of the target potential, the collision has no effect. When l is equal to the angular momentum of the product of the reaction, the probability of interaction is strongly enhanced. In that case it happens for $l = 1$, so the second term of the partial wave expansion, which is known as P-wave scattering.

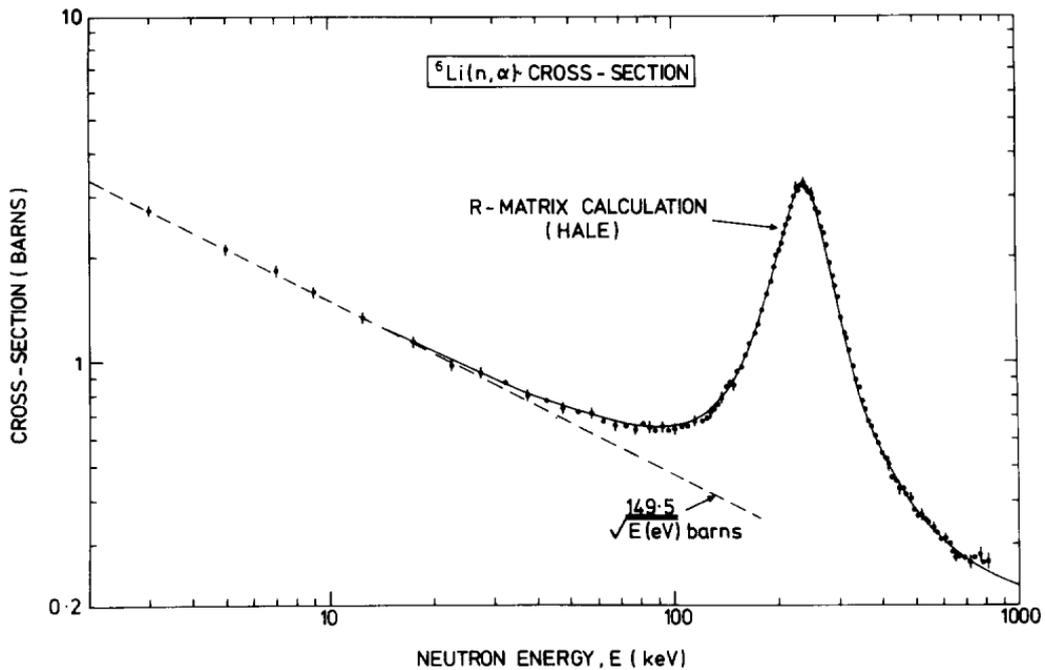


Figure 7.5: Cross-section of neutron capture with alpha emission on ${}^6\text{Li}$ [83].

The P-wave resonance appears on top of the exponentially decreasing behaviour of the cross-section with the neutron energy, as shown in figure 7.5. The effect of this resonance is a distortion of the observed spectrum with respect to the incoming neutron flux. When computing the energy distribution of the neutrons, this has to be taken into account by re-scaling by the inverse of the cross-section convoluted with the energy resolution of the detector.

7.1.4 Position and angle reconstruction

To infer the beam profile from the time-of-flight, the distance between the LiF target and the detector as well as the angle of the detector with respect to the beam axis must be known. The neutron detectors are placed on a

table that is 21 cm lower than the beam. While the table stays at the same position, the detector is moved on it to measure energies at different angles.

Two distinct points on the floor are defined forming with the target a plane perpendicular to the beam axis. The distances between each of these points and the front legs of the detector table are measured with a telemeter laser. This allows to reconstruct the table position with respect to the beam.

Knowing the coordinates of the detector on the table, and the height difference between the table and the beam, the distance and angle are computed. The impact of a 1 cm deviation on each distance measurement is propagated to estimate the resolution of this method, giving a 1.3 cm and 2.3° uncertainties on the target-detector distance and angle respectively. For redundancy and to prevent operator mistakes, two other similar measurements are performed, one with the back legs of the table and one from the front corners of the top of the table.

7.2 Characterisation results

The different neutron detectors were characterised with radioactive sources and with the beam. The time resolution was measured to infer the energy resolution that can be reached. This was then used to measure the beam energy profile.

7.2.1 Time-of-flight resolution

The acquisition systems triggers on the neutron detectors and acquires the waveforms for 6 μ s with a 2 ns sampling. A periodic signal at the beam pulse frequency (5 MHz) is also recorded. It is used to synchronise each event with the beam time in order to calculate the time-of-flight.

A dedicated software was developed to reconstruct and monitor the data. It includes an algorithm to measure the start time of the scintillation pulse that was specifically tuned to the Li-glass detector response. The start time is defined as the first sample after which the waveform is higher than 15 times the standard deviation of the baseline. The accuracy of this algorithm can be measured on the γ flash. The standard deviation of the reconstructed times of γ flash events gives the time resolution taking into consideration every

effect of the experiment from the beam to the reconstruction. The obtained resolution is 2.4 ns (figure 7.6). For the ^{10}B detectors, we obtained 0.8 ns resolution with the NaI(Tl) design and 0.3 ns with the BaF_2 .

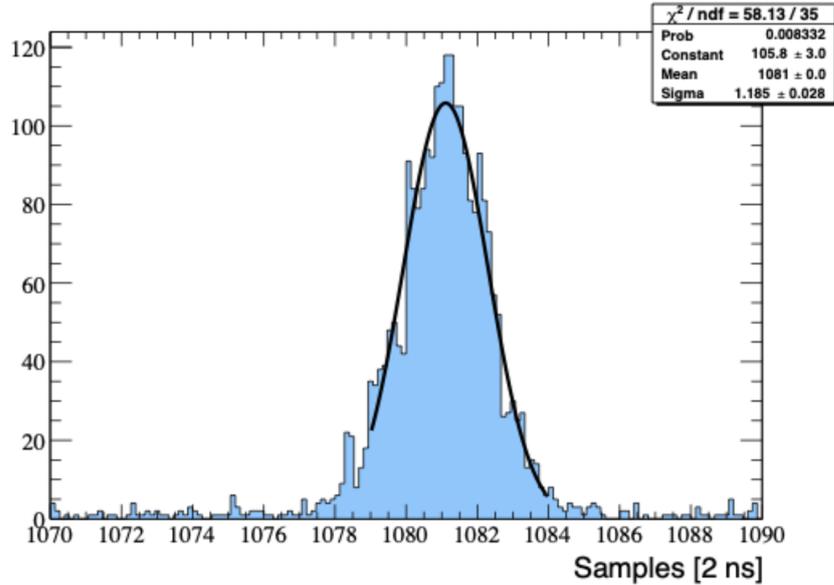


Figure 7.6: Reconstructed pulse time distribution for gamma flash events in Li-glass.

In the Li-glass detector, the scintillation threshold is around 5 MeV. Therefore, environmental γ s do not have enough energy to produce any signal, allowing the detector to be almost background free outside of the γ flash. On the contrary, in the ~ 500 keV region, the ^{10}B detector rates are dominated by background events which makes them unsuitable for the future measurement. The comparison of background and neutron rates can be seen in figure 7.7 for the two types of detector. The first outcome of this test is a strong preference toward the Li-glass detector.

The Li-glass scintillation can provide further discrimination between γ s and neutrons to reject background. Similarly to liquid argon, the estimator for the pulse shape discrimination is the fraction of light f_{prompt} contained in a prompt window after the trigger. The duration of the prompt window was searched by fitting with Gaussians the γ and the neutron f_{prompt} distributions. The optimal gate is the one that maximises this ratio:

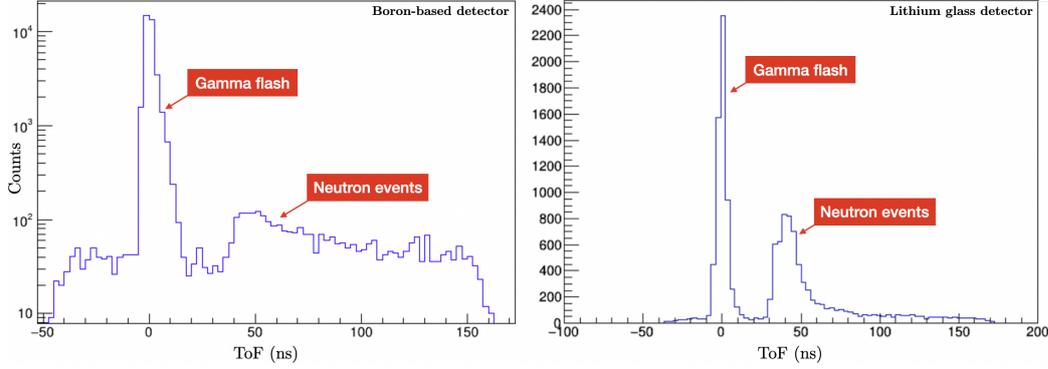


Figure 7.7: Histogram of the detection time modulo beam period for ^{10}B (left) and ^6Li (right) detectors. The first peak corresponds to gamma flash events occurring with a constant phase with respect to the beam pulse. The second peak are neutrons detected with some delay from which time-of-flight is computed.

$$\delta = \frac{\mu_n - \mu_\gamma}{\sqrt{\sigma_n^2 + \sigma_\gamma^2}} \quad (7.3)$$

Here, $\mu_{n,\gamma}$ and $\sigma_{n,\gamma}$ are the mean and standard deviation of the fitted Gaussians for the neutron and γ distributions respectively. This gives an optimal value for the gate of 300 ns. With this definition of f_{prompt} , the neutron and γ distributions are the ones shown in figure 7.8. Although the separation is not enough to safely remove a large amount of gamma events, it still allows to validate the nature of the detected flux.

7.2.2 Beam characterisation

We varied the proton beam between 2.1 MeV and 2.4 MeV. The difference between this energy and the threshold of the reaction is there to compensate the energy loss of the proton in the LiF target. The Li-glass detector which showed better performance was placed at different positions with respect to the beam to characterise its energy profile through time-of-flight measurement. With energies in the order of 100 keV, the neutrons are non-relativistic:

$$E = \frac{1}{2}m_n \times \left(\frac{d}{t}\right)^2 \quad (7.4)$$

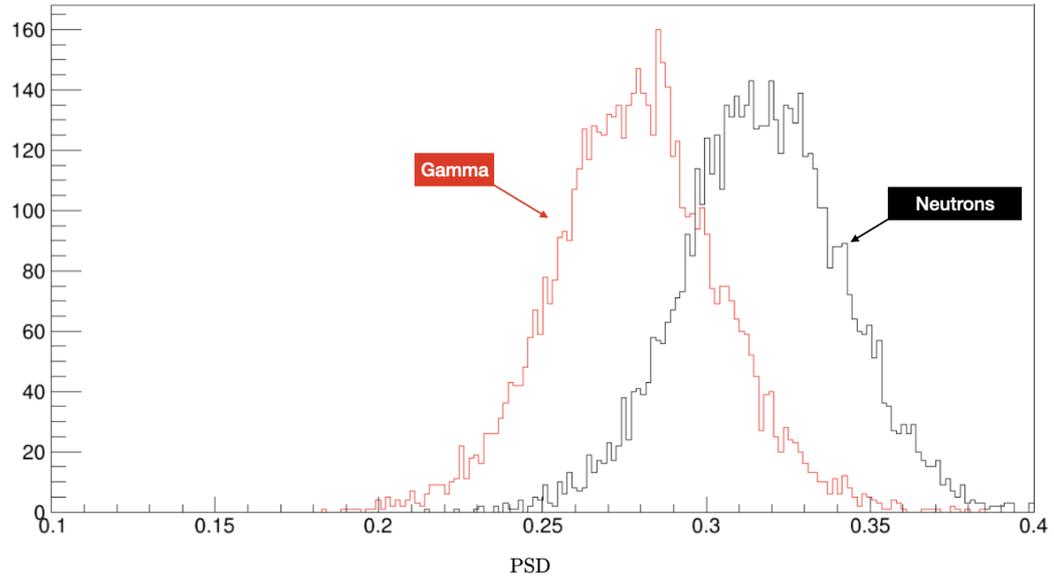


Figure 7.8: PSD estimator distribution between gamma and neutron events in ${}^6\text{Li}$ -glass detector.

For each position, the distance and angle are computed. The time-of-flight is then converted into energy and the obtained spectrum is corrected by the cross-section. We measure the peak of the neutron energy distribution as a function of the angle with respect to the beam.

With a beam energy fixed at 2.2 MeV, the obtained kinematics showed in (figure 7.9) is compared with the expected kinematics. The data matches with the theoretical energy profile below 30° although the energy loss in the LiF target is not included. At higher angles, the measured energy drastically drops. This behaviour is not understood and suggests that some effect is not properly taken into account such as the energy loss or unconsidered resolution effects. Despite this, we demonstrated the good performances of the Li-glass detector for the future measurement.

7.3 Nuclear recoil low energy sensitivity

In this section, I estimate the resolution of the future setup in term of NR response and energy. This preliminary projection is based on the results of the beam and neutron detector characterisations. The uncertainties associated to the beam profile and from geometrical considerations are propagated to

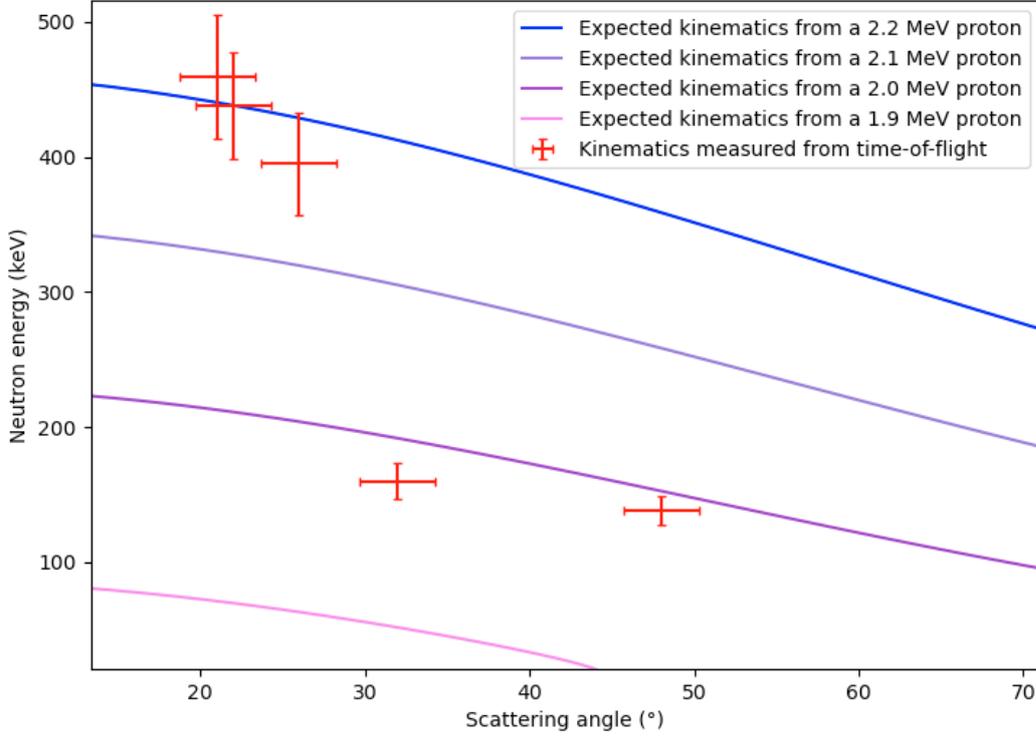


Figure 7.9: Neutron energy as a function of the scattering angle for a 2.2 MeV proton beam on ${}^7\text{Li}$ target.

the NR recoil energy in the future TPC. The sensitivity that can be expected from this study is compared to the results of ARIS on the measurement of the L_{eff} , the quenched quantum yield of NR with respect to ER.

7.3.1 Experimental setup

We assume a TPC of 5 cm height to minimise the drift time and the uncertainty associated to the location of the NR. From the beam profile showed in figure 7.9, we assume that a $160 \text{ keV} \pm 13 \text{ keV}$ neutron flux can be obtained by placing the TPC 32° away from the beam axis. The neutron detectors will be placed behind the TPC at different angles to give access to different NR energies as shown in figure 7.2. We assume 5 positions of the neutron detectors from 20° to 60° with respect to the beam-TPC axis giving access to NR energies between 0.5 keV and 4 keV.

The event selection will rely on the coincidence between the beam and any of the neutron detectors accounting for the time-of-flight of the neutron.

When the trigger condition is filled, an acquisition window as long as the drift time of the TPC ($\mathcal{O}(50 \mu\text{s})$) is opened to record the ionisation signal. The 2.4 ns resolution of the Li-glass is therefore crucial to mitigate background.

The distance between the TPC and neutron detectors is a compromise between the rate of coincidence events and the precision of the scattering angle. I tested two values of this distance: 18 cm and 36 cm.

7.3.2 NR energy uncertainty

The NR energy is inferred from the neutron detector angle. The spatial distribution of both the TPC and the neutron detector is therefore a major source of uncertainty. The position at which the neutron scatters and at which it is captured in the detector affects the scattering angle and therefore the deposited energy.

To give an estimation of this uncertainty, I assume that the interaction position is uniformly distributed between $-l_{TPC}/2$ and $+l_{TPC}/2$ in an axis perpendicular to the TPC-detector axis with l_{TPC} the characteristic size of the TPC. Similarly, the neutron is captured between $-l_{ND}/2$ and $+l_{ND}/2$, with l_{ND} the diameter of the Li-glass detector. Figure 7.10 is a scheme of the maximum deviation of the angle from this effect. The amplitude of the deviation induced by the spatial distributions can be expressed as follow:

$$\tan(\theta) = \frac{l_{TPC} + l_{ND}}{2d} \quad (7.5)$$

With d the distance between the centres of the TPC and the detector, the scattering angle varies between $-\theta$ and $+\theta$ with respect to the average. Assuming this variation to be uniform, the error is $\delta_{\theta} = \theta/\sqrt{3}$.

This spatial uncertainty is combined to the uncertainty on the beam energy. If the TPC is placed such that its solid angle with respect to the beam is the same as the neutron detector during the characterisation, the error on the beam energy is the same (13 keV uncertainty at a neutron energy of 160 keV). In a first order approximation, the relative error on the nuclear recoil energy from this contribution is equal to the relative error on the projectile neutron.

The NR energy resolution is computed from these two sources of uncertainties, for the given angles of the neutron detectors, and for the two considered distances between them and the TPC (18 cm and 36 cm). At

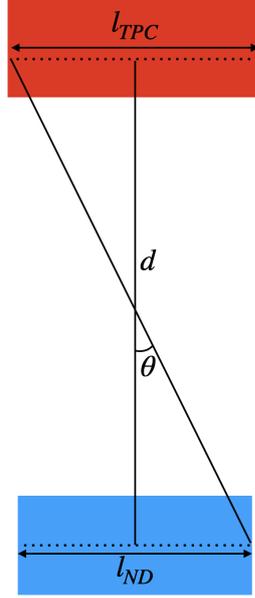


Figure 7.10: Scheme of the evaluation of the uncertainty from spatial distribution of the TPC in red and the neutron detector in blue.

low angle (low NR energy), the geometrical component dominates the error, while at higher energy, it is compatible with the beam energy uncertainty (table 7.1).

7.3.3 Quantum yield uncertainty

The error on the L_{eff} measurement is propagated from an estimation of the statistical and systematic uncertainties. In the absence of a TPC design, I assume that the systematic uncertainty will be similar to the one reached in ARIS and fix it to 1% [84].

The statistical uncertainties are computed by estimating the number of events that will be acquired in each neutron detector. We assume 10 days of data-taking with the beam frequency set at 10 MHz. During the beam characterisation, we recorded a neutron rate of 2 Hz. This rate is the product of the number of neutron crossing the detector and its efficiency. The latter can be computed from the dimension of the detector and the properties of the Li-glass. The density ρ of the Li-glass is 2.5 g/cm³, the weight composition w of Li₂O is 18%, and the concentration of ⁶Li is 95% [85]. The efficiency is then computed as:

$$Eff_{Li}(E_n = 160\text{keV}) = \frac{\sigma_n(160\text{keV}) \times \rho \times 2w \times 95\% \times t}{6g/\mathcal{N}_A} \quad (7.6)$$

Here, $t = 1\text{cm}$ is the thickness of the Li-glass detector and \mathcal{N}_A is the Avogadro constant. This gives an efficiency of 1.7%. The efficiency of the TPC and the angular distribution of the scattered neutrons were evaluated by simulating collisions of 160 keV neutrons in a 5 cm thick liquid argon volume with G4DS. The efficiency of the TPC is around 3%. I used this to compute that the expected neutron rate in the TPC is 7 Hz, assuming the same solid angle than the Li-glass during the characterisation and a pulse rate twice higher (10 MHz instead of 5 MHz).

The angular distribution of the simulated neutron sample is integrated over the solid angle of each Li-glass detector. This gives the probability for a scattered neutron to reach the detector. This number is multiplied by the efficiency of the Li-glass, the neutron rate in the TPC, and the 10 days of data-taking to give the expected number of events per detector.

Angle	E_{NR} (keV)	18 cm				36 cm			
		Geometry	NR error	Stat.	L_{eff} error	Geometry	NR error	Stat.	L_{eff} error
20°	0.49	0.54	0.55	0.072	0.073	0.26	0.27	0.141	0.141
30°	1.08	0.34	0.35	0.075	0.076	0.17	0.19	0.148	0.149
40°	1.88	0.25	0.26	0.076	0.077	0.12	0.15	0.148	0.149
50°	2.86	0.19	0.20	0.078	0.079	0.09	0.12	0.154	0.154
60°	3.99	0.15	0.17	0.079	0.080	0.07	0.11	0.151	0.151

Table 7.1: Accessed NR energy and relative uncertainties as a function of the Li-glass detector angle with respect to beam-TPC axis, considering them 18 cm and 36 cm away from the TPC. The relative beam uncertainty and the systematic uncertainty on the quantum yield are fixed at 0.08 and 0.01 respectively.

The uncertainty on the L_{eff} is defined as the quadratic sum of the systematic and the statistical uncertainties. The contribution of each uncertainty source is shown for all Li-glass position in table 7.1. Figure 7.11 shows a comparison of the expected results from this measurement and the results of ARIS.

This study is at a preliminary stage and a better compromise between the angular resolution and the statistical uncertainty can be found. The design of the TPC and the details of the setup will be guided by these results to refine this projection and to optimise the quality of the measurement. For

instance, the time-of-flight measurement, which is so far only considered for data selection, could improve the NR energy resolution. This work will be continued in view of the measurement that will take place in 2 or 3 years from now on.

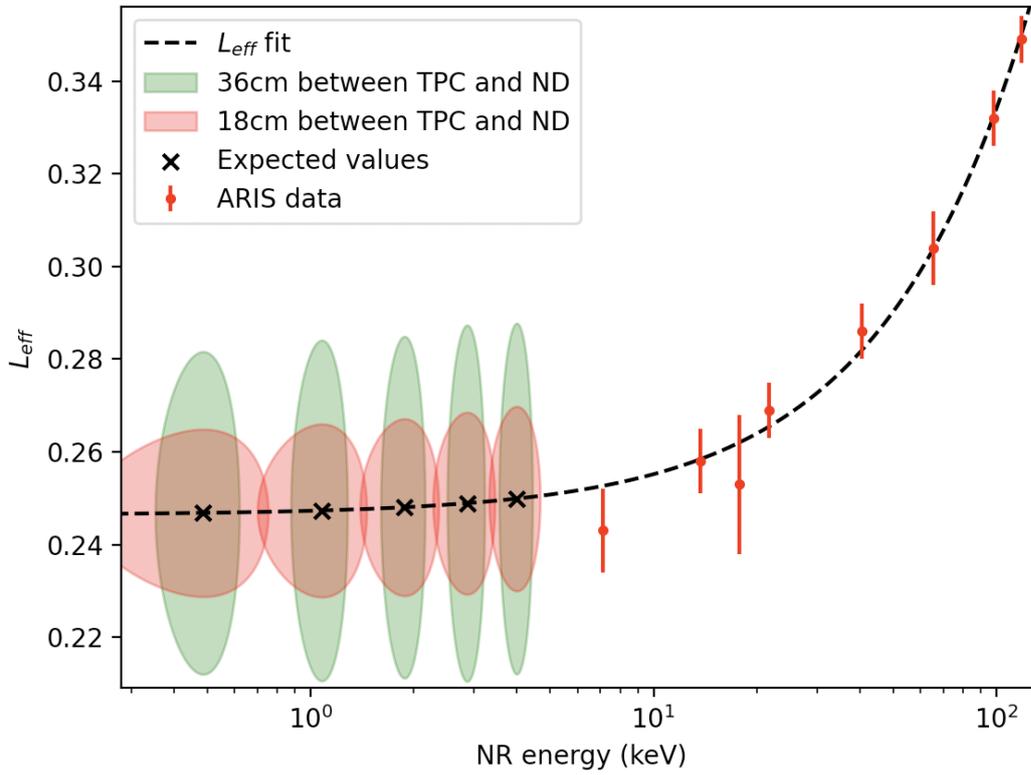


Figure 7.11: Expected NR energy accessed by the experimental setup compared to ARIS L_{eff} measurement. The projected data points are aligned on a linear fit of the L_{eff} just to provide an idea of what results are expected.

Conclusion

DarkSide-20k is expected to provide yet unachieved sensitivity on WIMP-nucleon interaction in the mass range between $0.1 \text{ TeV}/c^2$ and $10 \text{ TeV}/c^2$. Data-taking will start in 2027 and a cross-section sensitivity around 10^{-48} cm^2 will be reached at $0.1 \text{ TeV}/c^2$ in five years. Moreover, we demonstrated that world leading sensitivity will also be obtained for WIMPs below $5 \text{ GeV}/c^2$.

The success of such experiment relies, among other things, on the robustness of the simulation and reconstruction softwares. During these three years, I contributed to the finalisation of the full simulation chain of DarkSide-20k. We delivered to the collaboration packages that represent the foundations for the future reconstruction and analysis. The definition of the Event Data Model allowed us to reach a crucial milestone and to develop a reconstruction framework that ensures efficient reprocessing and parallel refinement of dedicated algorithms.

Despite the impressive technological progresses that have been made, discovering dark matter with the current generation of detectors is far from being guaranteed. In case of an excess of events, the complementarity of the results, in particular between argon and xenon experiments, is an absolute requirement for the community to claim for a discovery.

An ultimate generation of noble liquid detectors is under discussion in the community. The proposed XLZD and ARGO detectors, with xenon and argon targets respectively, would aim at enhancing the sensitivity by increasing the target mass by about one order of magnitude. However, this requires to develop powerful methods to deal with the extreme challenges of the event pile-ups caused by longer drift times. In view of the incoming difficulties, it is essential to broaden our knowledge on the target response and to explore alternative technologies in order to pursue the quest for dark matter. From this perspective, the results obtained in these three years may be a step toward the discovery of dark matter.

Bibliography

- [1] F. Zwicky. “Die Rotverschiebung von extragalaktischen Nebeln”. In: *Helv. Phys. Acta* 6 (1933), pp. 110–127. DOI: 10.1007/s10714-008-0707-4.
- [2] G. Bertone and D. Hooper. “A History of Dark Matter”. In: *Reviews of Modern Physics* 90 (2018).
- [3] K. Lundmark. “Über die Bestimmung der Entfernungen, Dimensionen, Massen und Dichtigkeit für die nächstgelegenen anagalactischen Sternsysteme.” In: *Meddelanden från Lunds Astronomiska Observatorium Serie I* 125 (1930), pp. 1–13.
- [4] K. Freeman. “On the Disks of Spiral and S0 Galaxies”. In: *Astrophysical Journal* 160 (1970), p. 811.
- [5] M. Fich et al. “The Rotation Curve of the Milky Way to 2R₀”. In: *Astrophysical Journal* 342 (1989), p. 272.
- [6] T. Treu and L. V. E. Koopmans. “Massive dark - matter halos and evolution of early - type galaxies to $z = 1$ ”. In: *Astrophys. J.* 611 (2004), pp. 739–760. DOI: 10.1086/422245. arXiv: astro-ph/0401373.
- [7] M. Milgrom. “A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis”. In: *Astrophysical Journal* 270 (1983), pp. 365–370.
- [8] D. Clowe et al. “A Direct Empirical Proof of the Existence of Dark Matter”. In: *Astrophysical Journal* 648.2 (2006).
- [9] Z. Chacko et al. “Cosmological Limits on the Neutrino Mass and Lifetime”. In: *JHEP* 04 (2020), p. 020. DOI: 10.1007/JHEP04(2020)020. arXiv: 1909.05275 [hep-ph].
- [10] J. L. Cervantes-Cota and G. Smoot. “Cosmology today-A brief review”. In: *AIP Conf. Proc.* 1396.1 (2011). Ed. by Luis Arturo Urena-Lopez et al., pp. 28–52. DOI: 10.1063/1.3647524. arXiv: 1107.1789 [astro-ph.CO].

- [11] H. Silverwood et al. “A non-parametric method for measuring the local dark matter density”. In: *Royal Astronomical Society* 459.4 (2016), pp. 4191–4208.
- [12] Michael Kuhlen et al. “Dark Matter Direct Detection with Non-Maxwellian Velocity Structure”. In: *JCAP* 02 (2010), p. 030. DOI: 10.1088/1475-7516/2010/02/030. arXiv: 0912.2358 [astro-ph.GA].
- [13] C. Alcock et al. “MACHO project limits on black hole dark matter in the 1-30 solar mass range”. In: *Astrophys. J. Lett.* 550 (2001), p. L169. DOI: 10.1086/319636. arXiv: astro-ph/0011506.
- [14] V. Atal, A. Sanglas, and N. Triantafyllou. “LIGO/Virgo black holes and dark matter: the effect of spatial clustering”. In: *Journal of Cosmology and Astroparticle Physics* 2020.11 (2020), p. 036. DOI: 10.1088/1475-7516/2020/11/036. URL: <https://dx.doi.org/10.1088/1475-7516/2020/11/036>.
- [15] K. Bechtol et al. “Dark Matter Science in the Era of LSST”. In: (Mar. 2019). arXiv: 1903.04425 [astro-ph.CO].
- [16] The ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- [17] Q. R. Ahmad et al. “Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory”. In: *Phys. Rev. Lett.* 87 (2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015.
- [18] R. N. Mohapatra and P. B. Pal. *Massive Neutrinos In Physics and Astrophysics*. 2004.
- [19] L. Di Luzio et al. “The landscape of QCD axion models”. In: *Phys. Rept.* 870 (2020), pp. 1–117. DOI: 10.1016/j.physrep.2020.06.002. arXiv: 2003.01100 [hep-ph].
- [20] R. D. Peccei. “The Strong CP problem and axions”. In: *Lect. Notes Phys.* 741 (2008). Ed. by Markus Kuster, Georg Raffelt, and Berta Beltran, pp. 3–17. DOI: 10.1007/978-3-540-73518-2_1. arXiv: hep-ph/0607268.
- [21] P. S. Bhupal Dev, A. Mazumdar, and S. Qutub. “Constraining Non-thermal and Thermal properties of Dark Matter”. In: *Front. in Phys.* 2 (2014), p. 26. DOI: 10.3389/fphy.2014.00026. arXiv: 1311.5297 [hep-ph].

- [22] A. Bottino et al. “Neutralinos as dark matter candidates”. In: (Feb. 1994). arXiv: [hep-ph/9405417](https://arxiv.org/abs/hep-ph/9405417).
- [23] L3 Collaboration. “Determination of the number of light neutrino species from single photon production at LEP”. In: *Phys. Lett. B* 431 (1998), pp. 199–208. DOI: [10.1016/S0370-2693\(98\)00519-X](https://doi.org/10.1016/S0370-2693(98)00519-X).
- [24] J. Kopp. “Sterile neutrinos as dark matter candidates”. In: *SciPost Phys. Lect. Notes* 36 (2022), p. 1. DOI: [10.21468/SciPostPhysLectNotes.36](https://doi.org/10.21468/SciPostPhysLectNotes.36). arXiv: [2109.00767](https://arxiv.org/abs/2109.00767) [[hep-ph](#)].
- [25] A. Boyarksy et al. “Sterile neutrino Dark Matter”. In: *Prog. Part. Nucl. Phys.* 104 (2019), pp. 1–45. DOI: [10.1016/j.pnpnp.2018.07.004](https://doi.org/10.1016/j.pnpnp.2018.07.004). arXiv: [1807.07938](https://arxiv.org/abs/1807.07938) [[hep-ph](#)].
- [26] V. Barinov and D. Gorbunov. “BEST impact on sterile neutrino hypothesis”. In: *Phys. Rev. D* 105.5 (2022), p. L051703. DOI: [10.1103/PhysRevD.105.L051703](https://doi.org/10.1103/PhysRevD.105.L051703). arXiv: [2109.14654](https://arxiv.org/abs/2109.14654) [[hep-ph](#)].
- [27] D. J. H. Chung et al. “On the Gravitational Production of Superheavy Dark Matter”. In: *Phys. Rev. D* 64 (2001), p. 043503. DOI: [10.1103/PhysRevD.64.043503](https://doi.org/10.1103/PhysRevD.64.043503). arXiv: [hep-ph/0104100](https://arxiv.org/abs/hep-ph/0104100).
- [28] S. Bhattacharya, P. Ghosh, and S. Verma. “SIMPLer realisation of Scalar Dark Matter”. In: *JCAP* 01 (2020), p. 040. DOI: [10.1088/1475-7516/2020/01/040](https://doi.org/10.1088/1475-7516/2020/01/040). arXiv: [1904.07562](https://arxiv.org/abs/1904.07562) [[hep-ph](#)].
- [29] C. Pérez de los Heros. “Status, Challenges and Directions in Indirect Dark Matter Searches”. In: *Symmetry* 12.10 (2020), p. 1648. DOI: [10.3390/sym12101648](https://doi.org/10.3390/sym12101648). arXiv: [2008.11561](https://arxiv.org/abs/2008.11561) [[astro-ph.HE](#)].
- [30] T. Aramaki et al. “Snowmass2021 Cosmic Frontier: The landscape of cosmic-ray and high-energy photon probes of particle dark matter”. In: (Mar. 2022). arXiv: [2203.06894](https://arxiv.org/abs/2203.06894) [[hep-ex](#)].
- [31] P. Bargassa et al. “Quantum algorithm for the classification of supersymmetric top quark events”. In: *Phys. Rev. D* 104.9 (2021), p. 096004. DOI: [10.1103/PhysRevD.104.096004](https://doi.org/10.1103/PhysRevD.104.096004). arXiv: [2106.00051](https://arxiv.org/abs/2106.00051) [[quant-ph](#)].
- [32] A. Askew et al. “Searching for Dark Matter at Hadron Colliders”. In: *Int. J. Mod. Phys. A* 29 (2014), p. 1430041. DOI: [10.1142/S0217751X14300415](https://doi.org/10.1142/S0217751X14300415). arXiv: [1406.5662](https://arxiv.org/abs/1406.5662) [[hep-ph](#)].
- [33] D.S. Akerib et al. “Snowmass2021 Cosmic Frontier Dark Matter Direct Detection to the Neutrino Fog”. In: (Mar. 2022). arXiv: [2203.08084](https://arxiv.org/abs/2203.08084) [[hep-ex](#)].

- [34] R. Bernabei et al. “Results from the DAMA/LIBRA experiment”. In: *J. Phys. Conf. Ser.* 203 (2010). Ed. by Eugenio Coccia et al., p. 012003. DOI: 10.1088/1742-6596/203/1/012003.
- [35] I Arnquist et al. “Search for Daily Modulation of MeV Dark Matter Signals with DAMIC-M”. In: *Phys. Rev. Lett.* 132.10 (2024), p. 101006. DOI: 10.1103/PhysRevLett.132.101006. arXiv: 2307.07251 [hep-ex].
- [36] Daniel Santos. “Dark Matter Directional Detection with MIMAC”. In: *J. Phys. Conf. Ser.* 1029.1 (2018), p. 012005. DOI: 10.1088/1742-6596/1029/1/012005.
- [37] P. Agnes et al. “Directionality for nuclear recoils in a liquid argon Time Projection Chamber”. In: *PoS PANIC2021* (2022), p. 067. DOI: 10.22323/1.380.0067.
- [38] M. Schumann. “Dark Matter Search with liquid Noble Gases”. In: (June 2012). arXiv: 1206.2169 [astro-ph.IM].
- [39] J. S. Adams et al. “Scintillation and Quantum Evaporation Generated by Single Monoenergetic Electrons Stopped in Superfluid Helium”. In: *Journal of Low Temperature Physics* 113 (1998), pp. 1121–1128.
- [40] C. Galbiati et al. “Pulse shape study of the fast scintillation light emitted from xenon-doped liquid argon using silicon photomultipliers”. In: *JINST* 16.02 (2021), P02015. DOI: 10.1088/1748-0221/16/02/P02015. arXiv: 2009.06238 [physics.ins-det].
- [41] L. Baudis. “WIMP Dark Matter Direct-Detection Searches in Noble Gases”. In: *Phys. Dark Univ.* 4 (2014). Ed. by Frank Avignone and Wick Haxton, pp. 50–59. DOI: 10.1016/j.dark.2014.07.001. arXiv: 1408.4371 [astro-ph.IM].
- [42] S. Kubota, T. Takahashi, and T. Doke. “Mechanism of Scintillation of Helium, Helium-Argon, and Helium-Neon Gas Mixtures Excited by Alpha Particles”. In: *Phys. Rev.* 165 (1 1968), pp. 225–230. DOI: 10.1103/PhysRev.165.225. URL: <https://link.aps.org/doi/10.1103/PhysRev.165.225>.
- [43] A. Hitachi. “Properties of liquid xenon scintillation for dark matter searches”. In: *Astropart. Phys.* 24 (2005), pp. 247–256. DOI: 10.1016/j.astropartphys.2005.07.002.
- [44] E. Borisova and A. Buzulutskov. “Neutral bremsstrahlung and excimer electroluminescence in noble gases and its relevance to two-phase dark matter detectors”. In: *Eur. Phys. J. C* 81.12 (2021), p. 1128. DOI: 10.1140/epjc/s10052-021-09913-z. arXiv: 2107.06497 [physics.ins-det].

- [45] F. Merkt et al. “High Rydberg states of argon: Stark effect and field-ionization properties”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 31.8 (1998).
- [46] K. Abe et al. “A measurement of the scintillation decay time constant of nuclear recoils in liquid xenon with the XMASS-I detector”. In: *Journal of Instrumentation* 13.12 (2018).
- [47] K. Likharev. *Part QM: Quantum Mechanics. Essential Graduate Physics*, 2013.
- [48] T. Doke et al. “Let Dependence of Scintillation Yields in Liquid Argon”. In: *Nucl. Instrum. Meth. A* 269 (1988), pp. 291–296. DOI: 10.1016/0168-9002(88)90892-3.
- [49] T. Alexander et al. “Light Yield in DarkSide-10: A Prototype Two-Phase Argon TPC for Dark Matter Searches”. In: *Astropart. Phys.* 49 (2013), pp. 44–51. DOI: 10.1016/j.astropartphys.2013.08.004. arXiv: 1204.6218 [astro-ph.IM].
- [50] P. Agnes et al. “DarkSide-20k Technical Design Report”. In: *Submitted to INFN* (2021).
- [51] P. Agnes et al. “Calibration of the liquid argon ionization response to low energy electronic and nuclear recoils with DarkSide-50”. In: *Phys. Rev. D* 104.8 (2021), p. 082005. DOI: 10.1103/PhysRevD.104.082005. arXiv: 2107.08087 [physics.ins-det].
- [52] P. Agnes et al. “DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon”. In: *Phys. Rev. D* 98.10 (2018), p. 102006. DOI: 10.1103/PhysRevD.98.102006. arXiv: 1802.07198 [astro-ph.CO].
- [53] C. Stanford et al. “Surface background suppression in liquid argon dark matter detectors using a newly discovered time component of tetraphenyl-butadiene scintillation”. In: *Phys. Rev. D* 98.6 (2018), p. 062002. DOI: 10.1103/PhysRevD.98.062002. arXiv: 1804.06895 [physics.ins-det].
- [54] A. Abed Abud et al. “Scintillation light detection in the 6-m drift-length ProtoDUNE Dual Phase liquid argon TPC”. In: *Eur. Phys. J. C* 82.7 (2022), p. 618. DOI: 10.1140/epjc/s10052-022-10549-w. arXiv: 2203.16134 [physics.ins-det].
- [55] P.-A. Amaudruz et al. “First results from the DEAP-3600 dark matter search with argon at SNOLAB”. In: *Phys. Rev. Lett.* 121.7 (2018), p. 071801. DOI: 10.1103/PhysRevLett.121.071801. arXiv: 1707.08042 [astro-ph.CO].

- [56] F. Carnesecchi. “Light detection in DarkSide-20k”. In: *JINST* 15.03 (2020), p. C03038. DOI: 10.1088/1748-0221/15/03/C03038.
- [57] P. Agnes et al. “Simulation of argon response and light detection in the DarkSide-50 dual phase TPC”. In: *JINST* 12.10 (2017), P10015. DOI: 10.1088/1748-0221/12/10/P10015. arXiv: 1707.05630 [physics.ins-det].
- [58] N. T. Fourches and M. Zielińska. *High Purity Germanium: from gamma-ray detection to dark matter Subterranean detectors*. Gamma-rays, 2019.
- [59] S. Walsh et al. “A sensitive method to determine ^{210}Po and ^{210}Pb in environmental samples by alpha spectrometry using CuS micro-precipitation”. In: *Scientific Reports* 13.1 (2023), p. 19754.
- [60] T. Van Acker et al. “Inductively coupled plasma mass spectrometry”. In: *Nature reviews Methods primers* 3.1 (2023), p. 52.
- [61] S. Amoruso et al. “Study of electron recombination in liquid argon with the ICARUS TPC”. In: *Nucl. Instrum. Meth. A* 523 (2004), pp. 275–286. DOI: 10.1016/j.nima.2003.11.423.
- [62] M. Miyajima et al. “Average energy expended per ion pair in liquid argon”. In: *Phys. Rev. A* 9 (1974), pp. 1438–1443. DOI: 10.1103/PhysRevA.9.1438.
- [63] T. Doke et al. “Estimation of Fano factors in liquid argon, krypton, xenon and xenon-doped liquid argon”. In: *Nucl. Instrum. Meth.* 134 (1976), pp. 353–357. DOI: 10.1016/0029-554X(76)90292-5.
- [64] R. L. Platzman. “Total Ionization in Gases by High-Energy Particles: An Appraisal of Our Understanding”. In: *International Journal of Applied Radiation and Isotopes* 10.2-3 (1961), pp. 116–127.
- [65] T. Doke et al. “Estimation of the fraction of electrons escaping from recombination in the ionization of liquid argon with relativistic electrons and heavy ions”. In: *Chemical Physics Letters* 115.2 (1985), pp. 164–166.
- [66] M. Kara et al. “The SNEWS 2.0 Alert Software for the Coincident Detection of Neutrinos from Core-Collapse Supernovae”. In: (June 2024). arXiv: 2406.17743 [astro-ph.IM].
- [67] Siu Kwan Lam, Antoine Pitrou, and Stanley Seibert. “Numba: a LLVM-based Python JIT compiler”. In: *International Conference for High Performance Computing, Networking, Storage and Analysis: HPC Transforms*. 2015. DOI: 10.1145/2833157.2833162.

- [68] S. Vinogradov. “Analytical model of SiPM time resolution and order statistics with crosstalk”. In: *Nuclear Instruments and Methods in Physics Research A* 787 (2014), pp. 229–233.
- [69] J. D. Scargle. “Studies in astronomical time series analyses. V. Bayesian blocks, a new method to analyze structure in photon counting data”. In: *The Astrophysical Journal* 504 (1998), p. 405.
- [70] J. D. Scargle et al. “Studies in astronomical time series analyses. VI. Bayesian blocks representations”. In: *The Astrophysical Journal* 764 (2013), p. 167.
- [71] W. Tompkins. “Applications of Likelihood Analysis in Gamma-Ray Astronomy”. PhD thesis. Stanford University, 1999.
- [72] J. Rode. “DarkSide : à la recherche de la matière noire avec l’argon liquide”. PhD thesis. Sorbonne université, 2022.
- [73] J. Thomas and D. A. Imel. “Recombination of electron-ion pairs in liquid argon and liquid xenon”. In: *Phys. Rev. A* 36 (1987), pp. 614–616. DOI: 10.1103/PhysRevA.36.614.
- [74] Heiner Ryssel and Hans Glawischnig, eds. *The Stopping and Range of Ions in Solids*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1982, pp. 122–156. ISBN: 978-3-642-68779-2.
- [75] H. Cao et al. “Measurement of Scintillation and Ionization Yield and Scintillation Pulse Shape from Nuclear Recoils in Liquid Argon”. In: *Phys. Rev. D* 91 (2015), p. 092007. DOI: 10.1103/PhysRevD.91.092007. arXiv: 1406.4825 [physics.ins-det].
- [76] P. Agnes et al. “Search for Dark Matter Particle Interactions with Electron Final States with DarkSide-50”. In: *Phys. Rev. Lett.* 130.10 (2023), p. 101002. DOI: 10.1103/PhysRevLett.130.101002. arXiv: 2207.11968 [hep-ex].
- [77] P. Agnes et al. “Search for Dark-Matter–Nucleon Interactions via Migdal Effect with DarkSide-50”. In: *Phys. Rev. Lett.* 130.10 (2023), p. 101001. DOI: 10.1103/PhysRevLett.130.101001. arXiv: 2207.11967 [hep-ex].
- [78] A. Migdal. “Ионизация атомов при ядерных реакциях (Ionization of atoms in nuclear reactions)”. In: *Soviet Physics JETP* 9 (1939), pp. 1163–1165.
- [79] H.M. Araújo et al. “The MIGDAL experiment: Measuring a rare atomic process to aid the search for dark matter”. In: *Astropart. Phys.* 151 (2023), p. 102853. DOI: 10.1016/j.astropartphys.2023.102853. arXiv: 2207.08284 [hep-ex].

- [80] P. Agnes et al. “Search for low-mass dark matter WIMPs with 12 ton-day exposure of DarkSide-50”. In: *Phys. Rev. D* 107.6 (2023), p. 063001. DOI: 10.1103/PhysRevD.107.063001. arXiv: 2207.11966 [hep-ex].
- [81] X. Mougeot and C. Bisch. “Consistent calculation of the screening and exchange effects in allowed β^- transitions”. In: *Phys. Rev. A* 90 (1 2014), p. 012501. DOI: 10.1103/PhysRevA.90.012501. URL: <https://link.aps.org/doi/10.1103/PhysRevA.90.012501>.
- [82] 20k Collaboration et al. *DarkSide-20k sensitivity to light dark matter particles*. 2024. arXiv: 2407.05813 [hep-ex]. URL: <https://arxiv.org/abs/2407.05813>.
- [83] D. B. Gayther. “A measurement of the ${}^6\text{Li}(n, \alpha)$ cross-section”. In: *Annals of Nuclear Energy* 4.11-12 (1977), pp. 515–526.
- [84] P. Agnes et al. “Measurement of the liquid argon energy response to nuclear and electronic recoils”. In: *Phys. Rev. D* 97.11 (2018), p. 112005. DOI: 10.1103/PhysRevD.97.112005. arXiv: 1801.06653 [physics.ins-det].
- [85] Y. Song et al. “Monte Carlo simulation of a very high resolution thermal neutron detector composed of glass scintillator microfibers”. In: *Applied Radiation and Isotopes* 108 (Dec. 2015). DOI: 10.1016/j.apradiso.2015.12.035.

List of Figures

1.1	Example of rotation curves. Dashed curve A describes a situation in which all matter is contained in the disk of radius r_0 . Curve B is what is typically measured.	21
1.2	Strong lensing from point-like (left) and extended (right) sources seen by the Hubble telescope. The left image shows a quasar seen 4 times as it is distorted by the central galaxy. The right image shows the arc image of a galaxy formed through gravitational lensing.	22
1.3	Composite image of the galaxy cluster 1E 0657-56, known as Bullet cluster. The pink area is the X-Ray distribution corresponding to baryonic (ordinary) matter. The purple area is the mass distribution reconstructed with gravitational lensing.	24
1.4	Maps of the anisotropies in the CMB	25
1.5	By microlensing effect, light rays emitted are converged by a massive object between the source and the observer.	28
1.6	Current 95% confidence level limits on MACHO contribution to halo as a function of its mass along with projected sensitivity of the Large Synoptic Survey Telescope [15].	29
1.7	The Standard Model of particle physics. Indicated neutrino masses are those of mass eigenstates which differ from the represented flavour eigenstates.	30
1.8	Abundance evolution of dark matter in case of freeze-out (continuous lines) and in case of freeze-in (dotted lines) for a mass $m_\chi = 100 \text{ GeV}/c^2$. The horizontal line is the relic density observed by the Planck telescope [21].	34
1.9	Summary of the main dark matter particle candidates in the cross-section-mass phase space[29].	36
1.10	Schematic diagrams for the three methods of dark matter search.	37
1.11	Stop pair production at the LHC with four-body decays	38

1.12	90% confidence level limit on spin-independent (left) and spin-dependent (right) WIMP-Nucleon cross-section in ATLAS and CMS from the mono-photon channel compared to direct search experiments in 2014.	39
1.13	Muon flux and depth of underground laboratories.	41
1.14	Local velocity distribution of dark matter in the halo and Earth restframes from three different simulations (Via Lactea 2, GALO and GALO _S). Dotted lines are the expected distribution from Maxwell-Boltzmann distribution. [12]	43
1.15	DAMA/LIBRA residual rate of single-hit scintillation events between 2 and 6 keV [34].	44
2.1	Light yields, normalised to the mean value, from each radioactive background gamma peak identified in the DarkSide-50 null-field spectrum. The light yield, namely the number of detected photons divided by the deposited energy, is constant within 2%, so it is with the quenching. The deviation at low energy may be explained by the increase of the fraction of energy deposits that are lower than w	50
2.2	Detection principle of a dual phase time-projection chamber.	52
2.3	Pulse shape discrimination in DS-50 from atmospheric argon run showing a rejection power higher than $1 : 10^7$. The red lines are different acceptances of the NR band.	54
2.4	DarkSide-50 TPC.	56
2.5	Event distribution in the (S1, f_{90}) space after all analysis cut. Blue area indicates the WIMP search region. the dashed line are the acceptance contours for NR derived from $^{241}\text{ArBe}$ calibration.	57
2.6	Spin-independent DM-nucleon cross-section 90 % confidence level in 2018.	58
2.7	DarkSide-20k TPC, inner veto, and stainless steel vessel.	59
2.8	Geometry of the TPC walls including ESR and TPB coating. The shape of the TPB layer indicates expected non-uniformity of the coating.	60
2.9	Schematic of the fast Cherenkov light produced in the walls for an ER event. The prompt fraction of the S1 pulse increases up to the NR band.	63
2.10	Accidental coincidence between S1-only event due to α events from the wall and S2-only reconstructed top event.	64
2.11	DarkSide-20k vetoes and cryostat.	65

2.12	Disposal of the photo-detection modules for the inner veto, on the outside of the TPC.	66
2.13	Mounting of a PDU in LNGS clean room.	67
2.14	Working principle of the p-n junction of an APD. p and n regions are populated with free holes and free electrons respectively. The p ⁺ and n ⁺ area on the top indicates the active region in which the photo-electric effect occurs. The p region around is the one in which the avalanche occurs as the electric field is the strongest.	68
2.15	Equivalent circuit of the 24 SiPM tile.	68
2.16	From SiPM APDs to the full DarkSide-20k optical planes. . .	69
3.1	Effective recombination probability fitted on DarkSide-50 data [57].	73
3.2	Full ²³⁸ U decay chain. Dashed lines represents the two breaks of the secular equilibrium before ²²² Rn and ²¹⁰ Pb.	76
3.3	Electronic recoil spectrum of DarkSide-20k obtained with G4DS and the realistic sampling of background events. The spectrum corresponds to 500s of exposure.	78
3.4	Gaussian fit of single electrons extracted from echoes in DarkSide-50.	80
3.5	Channel probability distribution at three different positions. .	81
3.6	Electron drift time as a function of the z-coordinate of the deposit, which is 0 at the centre of the TPC. The dashed line indicates the grid above which the electric field is stronger. The red area indicates the fluctuation of the drift time due to the diffusion effect.	82
3.7	Time-of-flight distribution for top and bottom optical plane (left) and example of an S2 pulse generated with light maps (right).	83
3.8	Comparison of an S2 pulse simulated with the full optical model in G4DS and with the light maps approach.	83
3.9	Time-of-flight photon distribution in low and high occupancy channels fitted with equation 3.7.	84
3.10	Power law fit of τ_{decay} and δt	85
3.11	Time-of-flight PDF computed for different channel occupancies.	86
3.12	Time-of-flight distributions for high occupancy channel and low occupancy channel from full optics simulation and from effective generation. Produced in a different location as the one from which the parameters were extracted.	87

3.13	Photo-electron time distribution in the inner veto with and without time-of-flight.	87
4.1	Overall DarkSide20k DAQ. The GDM (1) is connected to the CDMs (2) to distribute clock and commands to the veto (3) and TPC (4) digitisers. The digitisers are connected through a network switch (5) to the FEPs (6). The FEPs are connected to a network data switch (7) all located in the counting room away from the detector. Then, the overall data is distributed to the TSP (8) cluster. The Midas DAQ server (9) manages the data traffic from digitisers to TSPs via separated network switches (10, 11, 12). External network is provided via the gateway server (14). Room for additional processors is anticipated (15).	91
4.2	Example of zero-length encoding	92
4.3	Templates for old and new FEB, with the two capacitor values, used in the DAQ emulation software.	94
4.4	Noise power spectra for different FEB configurations	95
4.5	Gaussian noise waveform (left) and realistic noise waveform from power spectrum (right), both with a single hit shot. . . .	95
4.6	AR (left) and AR-MA (right) filtered waveforms. The horizontal dashed line is the threshold applied on the subtracted filtered waveform and the vertical line is the abscissa of the so-found hit.	96
4.7	Prominence distribution of simulated hits for two different values of MA gate. The true hits are identified with time matching the true PE time.	97
4.8	Single hit efficiency, fake hit rate, and δ_{50} evaluated along MA gate and threshold with 68 pF capacitor under 2x2 grouping configuration operating at 7 VoV.	98
4.9	Hit finder over a waveform with two PEs separated by a 512 ns delay.	99
4.10	Two hits efficiency curves for 68 pF (on the left) and 33 pF (on the right) both under 2x2 grouping operating at 7 VoV. .	100
4.11	Single photo-electron resolutions of the HFA for 68 pF, 2x2 grouping, and 7 VoV for prominence (left) and charge (right). .	101
4.12	Hit prominence (left) and charge (right) finger plots from one channel of Setup-1 laser calibration. The prominence cut is not applied at this stage, hence the fake hit peaks at low prominence and charge.	101

- 4.13 Finger-plot from the charge computed around the trigger region (left) with gaussian fit of the first 4 peaks and linear fit (right) of the mean value of the identified peaks. 102
- 4.14 PE distribution for a cross-talk probability of 0%, 5%, and 10%. The true PE occupancy is set at 3.5. The lines are computed with the Vinogradov model and the crosses are obtained from a toy Monte Carlo simulation. 104
- 4.15 Hit time histogram in the 5000 sample acquisition window. The external trigger is at 1220 samples. The well between the trigger peak and the after-pulse tail correspond to after-pulses that are too close to the primary hit and with too low amplitude to be detected by the HFA. The peaks in the beginning and the end of the window are populated by artefacts of the filtering at the edge of the waveform, so they are excluded from the analysis. 105
- 4.16 Summed waveforms before and after time-over-threshold cut and downsampling for an S1 and an S2. The waveforms are here normalised by the amplitude of one PE assuming ideal gain calibration. 108
- 4.17 A simulated waveform in which a ZLE containing 3 hits is identified. 111
- 4.18 Expected data rates of 100 ms slices with nominal dark count rate (left) and 10 times increased dark count rate (right). . . . 112
- 4.19 Distribution of ZLEs, hits and waveform rates with nominal dark count rate (left) and 10 times increased dark count rate (right). 112
- 4.20 Expected transfer rate per digitiser for 100 ms slices (blue) and 1 s slices (red) with nominal dark count rate (left) and 10 times increased dark count rate (right). 113
- 4.21 Finger plots from hit prominence (left) and ZLE integral (right) from S1 pulses. The truncation of the prominence is due to the 0.6 PE cut of the HFA. 114
- 4.22 Ratio between energy estimator and true number of PEs as a function of energy for S1 NR (left) and ER (right). 115
- 4.23 Ratio between standard deviations of the ZLE reconstructed number of PEs and the true number of PEs as a function of the energy. 115

- 4.24 f_{200} computed from hits without (left) and with prominence weights (right) as a function of S1 computed from the ZLE charge variable. The red line is the maximum leakage curve for ER of 0.1 event in 100 ton year. The black and dark red lines are the NR acceptance curve at 90% and 50% respectively. 116
- 4.25 Channel distribution of an S2 pulse for the 3 observables. The red stars indicates the true position of the event. 117
- 4.26 Hit time distribution (left) and summed waveforms (right) of 3 different S2 pulses, one from the top, one in the center and one on the bottom of the TPC. 117
- 5.1 Modular reconstruction framework in 4 layers. It ensures the possibility to replace a set of algorithms by another while keeping compatibility with previous and next layers. Raw data refers to the time slices as defined in the Event Data Model. . 120
- 5.2 Example of a preliminary cluster reconstructed from the Bayesian Blocks for two S2 pulses. The blue points are hits with a height proportional to their prominence. The red segments are the Bayesian Blocks. The dotted vertical lines are the true MC times and the horizontal green dashed line is the rate threshold. 122
- 5.3 Loops in the bayesian blocks algorithm in an example with 4 cells. In the primary loop, the log-likelihood is computed for a block starting from every previous cell to the last one considered. The so-computed log-likelihood from cell i to N is added to the one of the best segmentation previously found between 1 and i . The green square shows the optimal segmentation of the 4 cells with a total log-likelihood of 9. 123
- 5.4 50% and 100% chunk overlap on which bayesian blocks are computed. The dotted lines represent the limits before which change points are taken from the previous chunk and after which they are taken from the next chunk. 124
- 5.5 Computing time needed to process bayesian blocks algorithm without chunk optimisation and with 50% and 100% chunk overlap. This test is done interactively on a single core, before other optimisations. 125
- 5.6 Bayesian block based pulse finder on a hit distribution from multiple scatter gamma. The height of the blocks is proportional to their hit rates. 126
- 5.7 Hit rate and filtered rate of an isolated S1 pulse (left) and consecutive S2 from multiple scatter (right). The dashed lines are the error bands of the filtered rate. 128

5.8	Peak finder algorithm performance on a complex pulse.	129
5.9	Probability of reconstructing more than one peak in multiple S2 pulses as a function of the standard deviation of the times of the true pulses.	130
5.10	An S1-S2 event unresolved by the bayesian blocks but in which two peaks are reconstructed and indicated by the red vertical lines. The blue area is the pulse.	131
5.11	Probability to distinguish S1 and S2 as a function of the depth of the interaction, for different energies, with the Bayesian Blocks clustering and with the peak finder. Here, the higher field above the grid is not applied.	132
5.12	Difference between reconstructed start time and true MC time for S1 (left) and S2 (right) pulses.	133
5.13	Pulse classification over 10 s of full ER background. For tagging, truncated pulses are given a -1 value for f_{200}	134
5.14	TBA as a function of z computed on S1 pulses for the three energy observables.	136
5.15	Standard deviation of the S2 hit subset, between third and tenth percentiles, as a function of the z position of the interaction with and without prominence weights. This was obtained for ER events uniformly distributed in the TPC, with energy in the ROI.	137
6.1	Fits of the ER custom response extended from the Thomas-Imel model.	143
6.2	NR response model fitted on calibration data. The dark blue area is a fit from AmC and AmBe in DarkSide-50 while the light blue area is an extrapolation from ARIS and SCENE. The dark blue line is a combined fit of the DarkSide-50 neutron spectra and ARIS and SCENE data.	144
6.3	Signal model of low-mass WIMPs inducing NR in DarkSide-50, in the absence of quenching fluctuation (NQ) and with binomial quenching fluctuation (QF). The error bands are computed from the uncertainty on the liquid argon ionization response by producing the spectrum with a charge yield 1σ above and below the nominal value.	145
6.4	Signal model of WIMPs inducing ER in DarkSide-50, with a heavy mediator ($F_{DM} = 1$) and light mediator ($F_{DM} = 1/q^2$).	146
6.5	Signal model of sterile neutrinos in DarkSide-50. The error bands are computed with a charge yield 1σ above and below the nominal value.	147

- 6.6 Signal model of WIMPs inducing NR when considering Migdal effect, in DarkSide-50. The error bands are computed with a charge yield 1σ above and below the nominal value. 148
- 6.7 Background-only fit between 4 and 170 e^- with contributions of PMTs, cryostat and internal contaminations. The bottom plot shows the residuals defined as the difference between observed and expected number of events, normalised by the expected number of events. 151
- 6.8 For two different binnings, distances between the fitted nuisance parameters $\hat{\theta}$ and their prior values θ_0 normalised by the uncertainty on the latter. 151
- 6.9 Impact of each category of systematic to the expected limit for a WIMP mass of $3 \text{ GeV}/c^2$ and including Migdal effect. "Normalisation" denotes the activities of the PMTs, cryostat, and internal contaminants. "Shape" regroups the uncertainties of the Q-values and shape parameters of the internal contaminants. 152
- 6.10 Spin-independent DM-nucleon cross-section 90 % confidence level in the low-mass region for DarkSide-50 in the case of binomial quenching fluctuation (QF) and in the absence of quenching fluctuation (NQ) compared with other experiments. The coloured area are claimed discoveries. The grey area is the neutrino floor in liquid argon inside which $\text{CE}\nu\text{NS}$ background from solar neutrino cannot be neglected anymore. 153
- 6.11 Improvement of the limits on low-mass WIMPs when including Migdal effect in the signal model. 154
- 6.12 Limits from DarkSide-50 on light dark matter cross-section with heavy mediator (top left) and light mediator (top right), axion-electron coupling strength g_{Ae} (center left), photon and dark-photon kinetic mixing κ (center right), and sterile to electronic neutrino mixing angle $|U_{e4}|^2$ (bottom). 155
- 6.13 Background distribution summed with a signal corresponding to the highest cross-section excluded by DarkSide-50 at $2 \text{ GeV}/c^2$ 156
- 6.14 Single electron relative resolution from charge variable as a function of g_2 the amplification factor for S2. The relative resolution is defined as the ratio between the standard deviation and the mean from a gaussian fit. 157

6.15	500 ms time slice with S2 pulses simulated from the background model. Red area are the time windows opened before and after each pulse. Green peaks are isolated S2 pulses, selected for the analysis. Blue peaks are pulses less than 3.7 ms away from another, discarded by the cut.	158
6.16	DarkSide-20k expected limits at 90% confidence level after 1 year of exposure with the spurious electron models (dashed red) and without (dotted red). Binomial quenching (QF) fluctuation is assumed.	159
6.17	DarkSide-20k sensitivity to WIMP inducing NR considering Migdal effect (top left) and to leptophilic dark matter candidates.	161
7.1	Accessible NR energy in argon from a 100 keV neutron beam.	163
7.2	Scheme of the low energy NR measurement with neutron beam from the ${}^7\text{Li}(p, n){}^7\text{Be}$ reaction.	164
7.3	Expected neutron beam profile from ${}^7\text{Li}(p, n){}^7\text{Be}$ reaction with 1.9 MeV proton.	166
7.4	Picture of the ${}^{10}\text{B}$ based neutron detector with BaF_2 scintillators. Two scintillators are placed around the ${}^{10}\text{B}$ capsule and each of them is read-out with a PMT.	167
7.5	Cross-section of neutron capture with alpha emission on ${}^6\text{Li}$ [83].	168
7.6	Reconstructed pulse time distribution for gamma flash events in Li-glass.	170
7.7	Histogram of the detection time modulo beam period for ${}^{10}\text{B}$ (left) and ${}^6\text{Li}$ (right) detectors. The first peak corresponds to gamma flash events occurring with a constant phase with respect to the beam pulse. The second peak are neutrons detected with some delay from which time-of-flight is computed.	171
7.8	PSD estimator distribution between gamma and neutron events in ${}^6\text{Li}$ -glass detector.	172
7.9	Neutron energy as a function of the scattering angle for a 2.2 MeV proton beam on ${}^7\text{Li}$ target.	173
7.10	Scheme of the evaluation of the uncertainty from spatial distribution of the TPC in red and the neutron detector in blue.	175
7.11	Expected NR energy accessed by the experimental setup compared to ARIS L_{eff} measurement. The projected data points are aligned on a linear fit of the L_{eff} just to provide an idea of what results are expected.	177

List of Tables

2.1	Physical and scintillation properties of noble liquids [39]. . . .	48
2.2	Properties of the two TPB scintillation components, the total light yield is 0.41 photons/keV.	63
2.3	Properties of the four ESR scintillation components, the total light yield is 0.35 photons/keV.	63
3.1	Background activity table of DarkSide-20k materials provided by the radioactive assays.	75
3.2	Expected rate in TPC and inner veto consider efficiencies of each process.	75
4.1	Optimisation results for 2x2 grouping configurations.	99
4.2	The time slice data structure.	110
5.1	Inefficiency and false positive probability of bayesian blocks pulse finder with and without downsampling.	127
5.2	Preliminary reconstruction formats.	139
6.1	List of systematics, their sources, and impacted signal and background components included in the binned profile likelihood.	150
7.1	Accessed NR energy and relative uncertainties as a function of the Li-glass detector angle with respect to beam-TPC axis, considering them 18 cm and 36 cm away from the TPC. The relative beam uncertainty and the systematic uncertainty on the quantum yield are fixed at 0.08 and 0.01 respectively. . . .	176