Probing new physics underground

Edwards, T.D.P.

Publication date
2019
Document Version
Final published version
License
Other

Probing New Physics Underground

Thomas Edwards

Probing New Physics Underground

Thomas Edwards

Probing New Physics Underground

This work has been accomplished at the Gravity and AstroParticle Physics in Amsterdam (GRAPPA) center of excellence and the Institute for Theoretical Physics (ITFA) of the University of Amsterdam (UvA). This research is funded by NWO through the VIDI research program "Probing the Genesis of Dark Matter" (680-47-532).

Probing New Physics Underground

Academisch Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het college voor promoties

ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op Donderdag 17 oktober 2019, te 10.00 uur

door

Thomas David Powell Edwards

geboren te London

Promotiecommissie

Promotor

prof. dr. M.P. Decowski	Universiteit van Amsterdam

Co-Promotores

dr. C. Weniger	Universiteit van Amsterdam

Overige leden

prof. dr. T. Bringmann	University of Oslo
prof. dr. A.P.Colijn	Universiteit van Utrecht
prof. dr. D.D. Baumann	Universiteit van Amsterdam
dr. S.M. Nissanke	Universiteit van Amsterdam
dr. G. Bertone	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Publications

This thesis is based on the following publications:

[1] Thomas D. P. Edwards and Christoph Weniger, *A Fresh Approach to Forecasting in Astroparticle Physics and Dark Matter Searches*, *JCAP* **02** (2018) 021, arXiv:1704.05458.

Presented in Chapter **4**.
Both authors participated in the planning of the project. CW initiated many of the analytic calculations and TE confirmed the results. The numerical analyses were done by both CW and TE.

[2] Thomas D. P. Edwards, Bradley J. Kavanagh, and Christoph Weniger *Dark Matter Model or Mass, but Not Both: Assessing Near-Future Direct Searches with Benchmark-free Forecasting*, *Phys. Rev. Lett.* **121** (2018) 181101, arXiv:1805.04117.

Presented in Chapter **5**.
All authors participated in the planning of the project. The analysis code was primarily written by TE with input from both BK and CW. The paper was written by all authors.

[3] Thomas D. P. Edwards, Bradley J. Kavanagh, Christoph Weniger, Sebastian Baum, Andrzej K. Drukier, Katherine Freese, Maciej Górski, and Patrick Stengel *Digging for Dark Matter: Spectral Analysis and Discovery Potential of Paleo-Detectors*, *Phys. Rev. D.* **4** (2019) 043541, arXiv:1811.10549

Presented in Chapter **6**.
All authors participated in planning the project. The analysis was primarily done by TE and BK, with input and feedback from all other authors. The paper was written primarily by TE and BK with input towards the end by all authors.

[4] Sebastian Baum, Thomas D. P. Edwards, Bradley J. Kavanagh, Patrick Stengel, Andrzej K. Drukier, Katherine Freese, Maciej Górski, and Christoph Weniger *Paleo-Detectors for Galactic Supernova Neutrinos*, `arXiv:1906.05800`

Presented in Chapter **7**.
All authors participated in planning the project. The analysis was primarily done by TE and BK, with input and feedback from all other authors. The paper was written by SB, TE, BK, and PS with input towards the end by all authors.

Other publications and preprints by the author:

[5] Thomas D. P. Edwards, Marco Chianese, Bradley J. Kavanagh, Samaya M. Nissanke and, Christoph Weniger, *A Unique Multi-Messenger Signal of QCD Axion Dark Matter*, `arXiv:1905.04686`.

[6] Thomas D. P. Edwards and Christoph Weniger, *swordfish: Efficient Forecasting of New Physics Searches without Monte Carlo*, `arXiv:1712.05401`.

[7] Richard T. Bartels, Thomas D. P. Edwards, and Christoph Weniger *Bayesian model comparison and analysis of the Galactic disc population of gamma-ray millisecond pulsars*, *Mon. Not. Roy. Astron. Soc.* **3** (2018) 048, `arXiv:1805.11097`.

[8] Richard T. Bartels and Thomas D. P. Edwards, *Comment on "Understanding the γ-ray emission from the globular cluster 47 Tuc: evidence for dark matter?"*, *Phys. Rev. D. (Accepted)* (2018) 048, `arXiv:1807.08800`.

[9] Sara Algeri, Melissa van Beekveld, Nassim Bozorgnia, Alyson Brooks, J. Alberto Casas, Jessi Cisewski-Kehe, Francis-Yan Cyr-Racine, Thomas D. P. Edwards (Editor), Fabio Iocco, Bradley J. Kavanagh, Judita Mamuzic, Siddharth Mishra-Sharma, Wolfgang Rau, Roberto Ruiz de Austri, Benjamin R. Safdi, Pat Scott, Tracy R. Slatyer, Yue-Lin Sming Tsai, Aaron C. Vincent, Christoph Weniger, Jennifer Rittenhouse West, and Robert L. Wolpert *Statistical challenges in the search for dark matter*, `arXiv:1807.09273`.

Contents

1 Introduction

Understanding the laws that describe the fundamental building blocks of the Universe has been a pursuit of physicist's for many centuries. Simultaneously, terrestrial experiments and astronomical observations have shown the enormity of the unknown. Whilst daunting, these observational channels provide a unique opportunity to learn about the world around us. Making use of this wealth of information requires one to understand both the statistics and theoretical physics that underpins these experiments. Only then can we anticipate where the next discovery will emerge and concentrate our thoughts accordingly. This thesis is an attempt to contribute to the huge ongoing effort to search for signs of new physics. In particular, we focus on how to efficiently search for Dark Matter and neutrino interactions using underground experiments.

The history of the Universe is described by the standard model of cosmology, commonly known as ΛCDM . The name comes from the components that currently dominate the energy density of the Universe, Λ for dark energy (DE) and CDM for Cold Dark Matter. The particles that make up everything we see and interact with are described by the Standard Model (SM) of particle physics which has been confirmed with unprecedented accuracy [10]. The SM is a particular Quantum Field theory that encodes all the fundamental particles that have been discovered to date. Finally, Einstein's General Relativity (GR) describes gravity, the force dictating structure formation throughout the Universe whilst also being responsible for keeping our feet glued to the surface of the Earth [11].

Despite the success of these theories in describing vast swathes of experimental observations both on Earth and throughout the Cosmos, many observational puzzles elude a complete theoretical explanation: The accelerated expansion of the Universe [12, 13, 14], the matter-antimatter asymmetry [15], and Dark Matter (DM) [16] to name just a few. Fortunately, we live in an age of extreme technological innovation and unprecedented experimental precision. We are therefore gifted with the necessary tools to explore these unanswered questions from an observational standpoint.

Searches for Beyond the Standard Model (BSM) physics are primarily concerned with the existence of new particles and understanding their interactions with the SM. There are two primary ways in which these particles could have evaded detection. They could either be too heavy to have been produced in particle collisions, such as those taking place at the LHC. Alternatively, these new particles could have such weak couplings to the SM that their interactions with experimental equipment go unnoticed.

These searches proceed by performing experiments to confront particular theories, either making a detection or ruling out certain regions of the parameter space. When excluding regions, the claim is that if a theoretical model were to have some set of parameters then the experiment would have seen it already. How to clearly define the regions which an experiment will rule out, or will remain *alive*, is subject to much discussion. To find any BSM physics, one must have a careful understanding of the background from the normal SM physics. In addition, making concrete detection statements requires a detailed description of the probability that these SM backgrounds will occur and how they will look. Systematic errors on backgrounds, whether there are many or few background events, the subtleties of finite resolution effects on signal predictions all play an important role. To accurately account for all these affects is a computationally expensive and opaque[1] procedure, which is therefore limited to a set of well known theoretical models. The optimal method to forecast an exclusion limit would have a simple interface for use by all physicists, be trustworthy given that any one experiment could be in the Poissonian or Gaussian background regime, and be able to account for errors in the sometimes uncertain background components. Theorists and experimentalists alike would then be able to examine the sensitivity of future experiments to their favourite model, allowing for a more concrete discussion about which set experiments can probe the largest range of BSM models. Unfortunately, connecting the physical understanding of backgrounds and their associated statistical distributions can be difficult, especially for searches of rare events.

BSM sensitivity forecasts are formulated around a model selection procedure. The goal is to quantify the probability that a particular model provides a better description of the experimental data than that of the SM. For experiments with many background events and a small associated systematic error, there are analytic descriptions of the test statistic used in the model selection [17]. These expressions can be used to great effect in most BSM searches, allowing researchers to quantify the sensitivity of future experimental setups. Unfortunately, these analytic descriptions break down when the number of background events becomes small (we will refer to this as the Poissonian regime), forcing one to charaterise the distribution of the test statistic using Monte Carlo (MC) simulations. The computational expense of MC simulations prevents one

[1] To non-statisticians, such as myself.

from evaluating the sensitivity of future experiments to wide classes of BSM models. We address this issue by computing approximate analytic expressions for the test statistic into the Poissonian regime [1, 6]. These methods allow for a more rigorous and complete discussion of how to design experimental setups to constrain the widest range of physical signatures, maximising the possibility of a discovery.

Of the many motivated BSM particle candidates, Dark Matter is potentially the most striking. Evidence for missing mass throughout the Universe is extremely abundant; from the Cosmic Microwave Background at the largest scales all the way to galaxy rotation curves, particle DM is a necessary and simple explanation of these phenomena. Despite the prevalent observations of DM's gravitational effect on astrophysical objects, little is known about it's particle nature. Currently, there is a huge ongoing effort to observe a non-gravitational signature in astrophysical settings. Simultaneously, terrestrial experiments seek to observe DM particle interactions with the SM directly, for example see Refs. [18, 19, 20, 21, 22].

Direct Detection (DD) is one of the primary methods used to search for DM that is both heavy and weakly interacting. The general tactic is to build an experiment that, for the vast majority of time, does nothing at all but is sensitive enough to detect particles flying in from outer space and interacting with the detector. Unfortunately, there are a variety of naturally occurring backgrounds: cosmic rays interacting in the atmosphere and natural radiation in materials to name just a few of the major culprits. Miraculously, experimentalists have driven these backgrounds down by building detectors in underground laboratories using ultra-pure materials. Today, experiments such as XENON1T [18], LUX[23], and PandaX [20] have reached unprecedented sensitivity to GeV scale Weakly Interacting Massive Particle (WIMP) DM. We will discuss the the complementarity of different DD experiments and how they can be used to identify the characteristics of particle DM using a non-relativisitc effective field theory formalism.

The long-term future of DM DD is uncertain. Experiments will increase their exposure whilst maintaining ultra-low backgrounds for at least another five years. Unfortunately, this progress is difficult to sustain, with each successive development increasing in cost. In addition, DD experiments using nuclear recoils with xenon (and other similar mass nuclei) lack sensitivity to sub-GeV scale DM. Consequently, many new variations upon the traditional DD methods have been proposed [24, 25, 26, 27]. To increase a DD experiments senstivity to GeV DM, the exposure of an experiment must be increased. On the other hand, sensitivity to sub-GeV DM comes from a reduction in the energy threshold of an experiment[2].

We will discuss how ancient minerals from the Earths crust can be used to gain competitive sensitivity to both sub-GeV and GeV scale WIMPs. The DM signature in these minerals would be nano-scale damage tracks left by nuclear recoils from DM-

[2]The energy threshold is the lowest energy interaction to which an experiment is sensitive.

nuclei interactions. The age of the minerals, $\mathcal{O}(10^9)$ yr, provides extreme sensitivity to GeV scale WIMPs. Additionally, the precision to which these tracks can be measured, $\mathcal{O}(1)$ nm, corresponds to an extremely low energy threshold, providing greater sensitivity to sub-GeV DM than any proposed experimental setup to date.

1.1 Outline

Each section requires its own introduction but we begin with a couple of broad overview chapters. In particular, Chapter 2 will focus on the primary theoretical motivations and methods to search for new physics. The observational evidence and theoretical constraints that represent the primary reasons to study particle DM will also be discussed here. We then discuss some particle DM models and the standard methods to search for them.

Chapter 3 will review some of the commonly used methods to calculate the sensitivity of an experimental approach to new physics. To constrain the breadth of this chapter we consider two primary questions asked by theoretical physicists and experimentalists alike:

- Given a new physics model, what set of experiments will provide the maximum probability that we will observe evidence of this model?
- Given a variety of theoretical models, what set of experiments will provide the maximum sensitivity to the largest range of models?

The answers to these questions also serve as the introduction for Chapters 4 and 5. In addition we hope this chapter will serve as a useful introduction to the statistics of new physics searches where the methods used are often opaque and not well documented for someone new to the field.

The rest of the thesis contains the main results. To begin with we focus on novel statistical methods with applications to Dark Matter direct detection searches in Chapters 4 and 5 respectively. In particular, we develop a suit of methods to allow for correct upper limits to be set without resorting to computationally expensive Monte-Carlo simulations. These methods are built around the Fisher information matrix, using the geometrical information of the likelihood space for the efficient computation of the likelihood ratio test statistic. They are of particular interest in new physics searches since many of these experiments can vary between the Poissonian and Gaussian statistical regimes. Our approach effectively and precisely interpolates between these regimes with little computational overhead, allowing for the efficient computation of sensitivities for large classes of new physics signals. In addition, we

liberate one from the restriction of resorting to benchmark scenarios when considering how well the parameters of a model from a future detection could be constrained. For this we developed the *euclideanised signal* technique, an effective mapping of the model parameters to a (typically) higher dimensional space that has the convenient property that the likelihood distance between points in a parameter space is approximated by the Euclidean distance in the higher dimensional space. Chapter 5 is the first direct application of the above methods to future experimental setups, namely DM direct detection experiments. We study the non-relativistic effective field theory of dark matter direct detection, showing for the first time that even with multiple detector materials it is unlikely that a future detection will allow for a simultaneous constraint on the mass and interaction type.

We then turn to novel search strategies by looking for the affects of DM and neutrinos on ancient minerals buried deep in the Earth's crust, presented in Chapters 6 and 7. These minerals, called *paleo-detectors*, act similarly to solid state direct detection methods; they record the recoils of nuclei from interactions between astrophysical particles and the minerals constituent nuclei. The interactions leave nano-scale defects in the otherwise crystalline structure of the mineral. These minerals, although only small ($\mathcal{O}(1)\,\mathrm{mm}^3$) compared to typical direct detection experiments ($\mathcal{O}(1)\,\mathrm{m}^3$), provide superior sensitivity to DM due to their extreme age; akin to running an experiment for one billion years. We show that paleo-detectors gain almost two orders of magnitude in sensitivity for GeV scale DM and many more for sub-GeV DM. Additionally, we show that these sensitivity predictions are robust to unknown systematic uncertainties in the backgrounds and that, given a detection, they are able to constrain the mass of the DM particle to much larger values than traditional direct detection experiments.

Neutrinos are the most mysterious sector of the standard model. Although light and extremely weakly interacting, their huge abundance can still cause visible affects in direct detection experiments. Even worse, the mountains and layers of Earth under which we build, and extract in the case of paleo-detectors, experiments are not effective shields to these particles. For DM detection in paleo-detectors neutrinos therefore represent a primary background. On the other hand, significant numbers of neutrino tracks should be observed and can be studied to extract interesting astrophysical information. We demonstrate this by studying neutrino emission from core-collapse supernovae (SN), explosions that occur at the end of a very massive stars' life, that have occurred in our own galaxy. We show that the time variation of the galactic star formation history over the past one billion years can be constrained using paleo-detectors. SN play an important role in galaxy formation and evolution. Understanding the local SN rate will therefore provide key insights into the history of our own Milky Way galaxy.

2 Finding New Physics

The purpose of this chapter is to lay the foundations of the standard models of particle physics and cosmology, whilst emphasising the two main ways in which BSM physics can evade detection; being too heavy or too weakly coupled. We then turn to the vast swathes of observational evidence for missing mass in the Universe and its interpretation as particle DM, emphasising that long range modifications to GR carries less motivation than certain classes of particle DM candidate. Finally, we present a couple the most well studied DM particle candidates and review direct detection methods since they are a main point of interest throughout this thesis. We use natural units throughout with $\hbar \equiv c \equiv 1$.

2.1 Particle Physics and Cosmology

2.1.1 Standard Model of Particle Physics

Poincaré invariance together with locality and unitarity constrain the space of possible SM particles and interactions. The Poincaré transformations encode the full symmetries of special relativity; translations, rotations, and boosts. The local gauge symmetry of the standard model is $SU(3) \otimes SU(2) \otimes U(1)$. The standard model of Particle Physics is made up of:

- spin-1/2 matter fields.
- spin-1 gauge fields to mediate the electroweak force.
- spin-1 gluon fields to mediate the strong force.
- A spin-0 Higgs boson to generate the masses of the particles through the Higgs mechanism [28].

A schematic diagram of the SM and their relative interaction strengths can be found in Fig. 2.1.

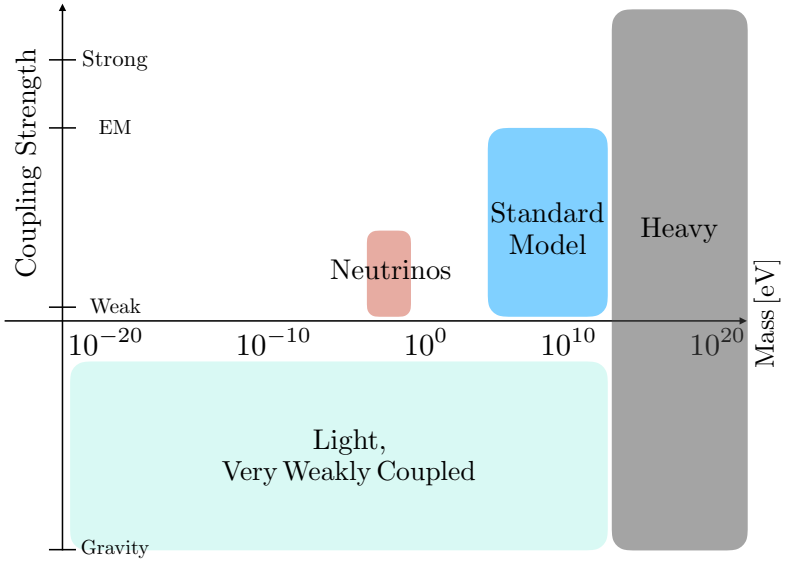

Figure 2.1: **Schematic diagram of the SM and potential avenues to search for new physics**. There are two primary ways in which particles can evade detection. Either they are too heavy for production at the LHC or too weakly couple to the SM to produce an effect. These scenarios are shown in grey and light blue respectively. Credit goes to Ref. [29]

The various fields can be split in different ways; Bosons are integer spin particles, whereas fermions have half-integer spin. The fundamental fermions can be further split into quarks and leptons. Quarks are particles that interact through the weak, electromagnetic, and strong forces. Leptons on the other hand do not take part in the strong force. Within the SM there are six quarks, up and down types being the most common in the Universe as they are the constituents of protons and neutrons. There are three charged fundamental leptons, the most famous being the electron, and three neutral ones. The neutral leptons are also the lightest massive particles in the SM and are known as neutrinos.

Neutrinos are the least understood sector of the SM, with many properties still unknown. Firstly, they appear with only left handed chirality, unlike the rest of the SM. Secondly, they are the only particle whose absolute mass remains unknown. Cosmology constrains their combined masses to be sub-eV [30], much smaller than that of the other SM particles[1]. Thirdly, a neutrinos flavour[2] can also oscillate whilst

[1]The other standard model particles range from $\mathcal{O}(1)$ MeV for the electron to $\mathcal{O}(100)$ GeV for the top-quark [10].

[2]The flavour of a neutrino is dictated by its associated charged lepton. They are called electron

propagating in space. For example, excess electron neutrinos produced in some astrophysical process can oscillate into other flavours such that we see equal ratios of each flavour on Earth. Oscillations have been observed in solar, atmospheric, and reactor neutrinos [31, 32, 33, 34]. Finally, whether they are their own anti-particle or not is still unknown, although some experiments may probe this soon through the potential observation of neutrinoless double beta decay [35, 36, 37].

The spin-1 particles mediate forces. At low energies there is the photon as the mediator of electromagnetism, the W and Z bosons for the weak force, and eight gluons that mediate the strong nuclear force. All particles have been confirmed with dramatic precision at experiments such as the LHC, its predacessors, and various table-top experiments [10]. In addition, the 19 free parameters of the SM have been measured now to a high degree of accuracy[3].

2.1.2 Standard Model of Cosmology

The *cosmological principle* states that the Universe is both homogeneous and isotropic when viewed on large enough scales. Synonymously, the laws of nature are the same for all observers in all directions throughout the Universe. This simple principle is supported by observations of large scale structure (LSS). When smoothed over scales of $\mathcal{O}(100)$ Mpc, the Universe appears to be approximately homogeneous [38, 39]. Similarly, observations of LSS look the same in all directions on the sky, thus confirming the isotropic nature of the Universe.

The geometry of the Universe is dictated by the cosmological principle. The spatial geometries must be embedded into the four dimensional spacetime. The cosmological principle restricts these embeddings to be spatially flat, positively curved, or negatively curved. The spacetime is described the Friedmann-Lemaître- Robertson-Walker metric given in spherical polar coordinates by

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right], \tag{2.1}$$

where $a(t)$ is the scale factor that encodes the cosmological expansion history and k is the parameter encoding the curvature. Observationally, the curvature parameter is constrained to be very close to zero, suggesting with live in a spatially flat Universe [40]. This fact can be encoded into the dimensionless energy density quantities as $\Omega_r + \Omega_m + \Omega_\Lambda = 1$.

Today, the energy desntiy is dominated by the cosmological constant Λ followed closely by that of cold dark matter (CDM). In the past however, this was not the case.

neutrinos, muon neutrinos, and tau neutrinos.

[3]Note that this does not include the neutrino masses which are not well understood.

The evolution of the Universe can be split into three distinct regimes: the radiation dominated era, the matter dominated era, and the dark energy dominated era we are in today. In cosmology, radiation generally refers to the energy density of photons and neutrinos[4], the most abundant and light particles. These regimes can be easily understood by their relation to the scale factor:

$$\Omega_r \propto a(t)^{-4}, \qquad \Omega_m \propto a(t)^{-3}, \qquad \Omega_\Lambda \propto a(t)^0 \,. \tag{2.2}$$

Since $a(t)$ is defined to be one today and monotonically decreases into the past, the three eras become clear due to the different rates at which their energy density contributions are diluted.

ΛCDM contains six free parameters; two for the initial conditions of the Universe encoding the primodial curvature perturbations, two encoding the DM and baryon densities, and two associated with parameters of the cosmic microwave background (CMB) [41].

2.1.3 searching for New Physics

There are two primary ways in which new particle physics could have evaded detection from experiments so far, both shown in Fig. 2.1 as the grey and light blue boxes. The first is that new particles could be too heavy to have been efficiently produced in a collider such as the LHC. The second is that the couplings to the SM are so weak that their effects are too small to measure. We briefly explain the logic behind searches for heavy particles and potential paths to search for very weakly interacting particles. DM searches will be the focus of this thesis. We therefore indicate the regions in which typical DM searches occur, giving some of the main examples.

Energy Frontier

There are two primary ways to search for heavy particles. The first is referred to as the energy frontier. The basic principle is that by colliding particles with high enough energies it becomes possible that a new, unknown heavy particle will be produced during the collision. The second is an indirect approach, utilising astrophysical observations of highly energetic environments.

At the Large Hadron Collider (LHC), protons are collided with a centre-of-mass energy up to $\sim 13\,\text{TeV}$. The LHC can detect many of the SM particles produced in the collision, allowing for a careful characterisation of the interaction that took place. It can therefore find new physics by looking for deviations from the SM background that it expects. One particular search pipeline looks for missing energy in the output

[4] In the early universe this is not true since many particles are relativistic and therefore contribute to the radiation energy density.

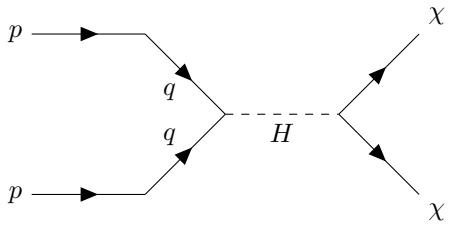

Figure 2.2: **Schematic Feynman diagram for Higgs portal DM production at the LHC**. p indicates protons made up of quarks that can interact and produce DM through a higgs portal coupling. The DM escapes the detector but its prescence can be inferred from missing translational energy.

of a collision. For example, DM with a Higgs portal coupling can be produced via the diagram in Fig. 2.2. Unlike the majority of the SM particles, DM will leave the detector without any interaction. Using translational energy conservation, one can equate the input energy of the proton collision to the measured output energy. If there is an imbalance between these energies, we infer that particles must have escaped the detector without interaction. By characterising the SM background of missing energy, one can search for the existence of unknown particles. Similar signatures exist for different DM models, some involving partial SM outputs in the form of jets [42]. Collider searches have the benefit of being a controlled setting where the input is known, allowing for a careful charaterisation of the SM background. For a review of DM searches at the LHC, see Ref. [43].

Although this approach has been extremely successful at finding the various particles of the SM (for example, see Ref. [44, 45, 46]), there is an obvious limitation set by the collider's scale. Any Particle with mass higher than the collisional energy of the protons cannot be produced and will therefore remain a mystery with collider search techniques.

Astrophysical probes can circumvent these constraints in two ways. If the new heavy particle has a coupling to SM photons (not necessarily directly) and a high enough astrophysical abundance we can search for the annihilation or decay products of the particle. These methods have been used to great affect in the search for DM annihilation products from dwarf galaxies by the Fermi Large Area Telescope [19]. The second method utilises the fact that ultra-high energy particles exist in nature up to 10^{20} eV [47], much greater in energy than anything produced on Earth. By modelling the possible astrophysical sources of these particles, we can search for observed deviations to the energy spectrum of cosmic rays [48], potentially finding new physics.

Very Weakly Coupled and Light Particles

Particles could have also evaded detection by being too weakly coupled to the SM. If so, even if the mass of the new particle was below the energy scale of the LHC, it would be produced at such a small rate that it would be impossible to observe.

There are a few ways to search for such particles; on the intensity frontier, by using low background experiments with large exposures, and by using indirect astrophysical or cosmological probes.

The intensity frontier is essentially the process of increasing the rate at which collisions occur in a collider or beam experiment. By observing larger numbers of events, we can search for extremely rare physical processes or small deviations from SM predictions. Low-background experiments with large volumes are one of the primary techniques to search for DM[5] but are also used to search for neutrinos. Super-Kamiokande [49] and IceCube [50] are two of the largest neutrino detectors, both utilising the very weak interactions of neutrinos with other particles to occasionally produce high energy charged leptons which can emit Cherenkov radiation [51]. In addition Super-Kamiokande can search for other rare processes such as proton decay [52], providing information about grand unified theories and perhaps probing energy scales well beyond that of colliders [53].

Astrophysical and cosmological probes utilise the high density environments that exist throughout the Universe. For example, observations of stellar cooling rates can provide constraints on additional light particles. These particles would be readily produced in the core of the star and escape out to space without interaction, thus providing an additional channel to remove energy [54]. Cosmological measurements provide another probe of light relics, for example see Refs. [55, 56]. More recently, observations of neutron stars have been proposed as an indirect way to observe the effects of light fields with couplings to the electromagnetic sector. This is particularly powerful when the these light fields make up all of the dark matter in the Universe, often called axion DM [57, 58, 5]. Finally, the recent observation of gravitational waves has provided a new view into light fields in the Universe. In particular, the phenomena known as *superraddiance* [59] allows for the classical generation of a cloud around a highly spinning black hole. This cloud can generate gravitational waves by itself [60] but is also dense enough to cause distortions to the binary inspiral that can be observed in current and next generation detectors [61].

2.2 Motivation for Particle Dark Matter

Dark Matter is perhaps the most well motivated BSM particle candidate. Here we review the wide variety of observational evidence for missing mass throughout the

[5] We will return to this subject in more depth in § 2.4.

Universe followed by a discussion on the theoretical boundaries of modifying gravity on large scales imposed by locality and unitarity.

2.2.1 Observational Evidence for Missing Mass

Various forms of the missing mass problem have been used to postulate the existence of new astronomical objects throughout history [62], in our galaxy and beyond. For example, in 1846, the astronomer Urbain Le Verrier proposed the existence of a new planet to explain peculiarities in the motion of Uranus [62]. This planet turned out to be Neptune which was discovered the same day it was postulated in 1846. The very same astronomer proposed the existence of another planet called *Vulcan* to explain the precession of the perihelion of Mercury [63]. Unlike Uranus though, an explanation of Mercury's movements had to wait until Newtons laws of gravitation were replaced with Einstein's GR [64].

These proposals all rely on the principle that gravitational interactions are universal whereas some objects may not emit enough light for us to observe, or at all. It is easy to find astrophysical objects which do not emit much light, brown dwarfs for example. It is even possible to imagine objects that emit no light at all, things we call black holes (BHs). Today we think that most spiral galaxies contain supermassive BHs ($> 10^6\,\mathrm{M}_\odot$) in their central region. We can precisely search for the gravitational affect of these BHs by monitoring the motions of stars close to the BH surface. The Event Horizon Telescope has recently been able to image the *shadow* created by the total absorption of light from background sources [65], constituting the first direct imaging evidence of BHs.

One of the most modern and compelling missing mass problems is that of DM. Unlike other instances of missing mass, the evidence for DM does not stem from any one observation. Rather, it is the accumulation of multiple probes pointing towards a common origin that constitute the global evidence of DM dominating the matter budget of the Universe. Here we discuss the main sources of evidence, proceeding from the smallest scales (galaxies) to large scale structure and the Cosmic Microwave Background (CMB)[6].

Galaxies

Galaxies resemble the smallest scale evidence that there is significant mass not accounted for in astrophysical objects. To infer the enclosed mass of a system we study the motions of stars; objects that are easy to see due to their emission of light at a variety of wavelengths. Since stars are gravitationally bound objects in a galaxy and

[6] We will include some historical perspective in each section although the sections do not proceed in chronological order.

approximately collisionless, they tend to follow circular orbits around the galactic centre [66]. We can predict the circular velocities of these stars by equating centripetal and gravitational forces to give

$$v_c = \sqrt{\frac{GM(\leq r)}{r}}\,, \tag{2.3}$$

where G is Newton's gravitational constant and $M(\leq r)$ is the mass enclosed by the orbit. Observations of luminous matter within these galaxies suggest that mass is concentrated towards their central region [67, 68]. The enclosed mass should therefore be constant at large radii, predicting that the circular velocity will fall as $v_c \propto r^{-1/2}$. Instead we observe that these velocities remain constant out to large radii [69], suggesting $M(\leq r) \propto r$. We therefore infer the existence of a dark *halo* which extends to significantly larger radii than visible matter.

We cannot directly measure the circular velocities of stars, instead these can be inferred from the spectra of galaxies. Vera Rubin [69] was the first to notice the flattening of rotation curves in the 1970's and 80's. Albert Bosma found similar results in 1985 [70], confirming Rubin's original study. Since then, huge observational campaigns have cemented galaxy rotation curves as the most famous evidence for DM [71]. DM also exists locally in our own galaxy. We can constrain the local density to be approximately $\rho_\odot \sim 0.3\,\mathrm{GeV\,cm^{-3}}$ using stellar kinematicss, see Ref. [72] for a review.

Galaxy Clusters

The next scale up is galaxy clusters; gravitationally bound structures formed from hundreds to thousands of individual galaxies. Evidence on this scale has two independent sources. The first comes from the velocities of the galaxies that make up these clusters. Like stars in rotation curves, we can measure the overall motions of galaxies within a cluster. This gives us a measure of the velocity dispersion, the statistical spread of velocities from the mean. By applying the virial theorem[7] we can infer the total mass of the system [66] directly from these velocities. Separately, an estimate of the mass of a cluster can be obtained by summing the masses of each galaxy and any intracluster medium. For this we require a *mass-to-light* ratio to translate between the observed luminosity and inferred mass [73]. For these mass measurements to agree, an unrealistically large mass-to-light ratio is required. Again we must infer an extended distribution of missing mass for observations to agree. Evidence of missing mass in galaxy clusters came first from Fritz Zwicky in 1933 and represents some of the first concrete evidence for DM [62].

In addition to these well separated and distinct galaxy clusters, we observe colli-

[7] Note that this applies to bound systems only.

Figure 2.3: **The collision of two galaxy clusters, called the *Bullet Cluster*.** The coloured regions show the X-ray emission and reconstructed mass distribution in pink and blue respectively. Knowing that the pink intracluster gas dominates the mass budget for baryonic matter we can see that the total mass is dominated by a collisionless dark component. Image credit - [74, 75]

sions of galaxy clusters. The most famous, named the Bullet cluster [74, 75], is shown in Fig. 2.3. The pink colour represents the X-rays emitted by hot gas known as intracluster medium. By studying the optical emission from the component galaxies as well as this X-ray emission one can conclude that this hot intracluster gas is the dominant form of visible, or *baryonic*, matter in the system. Mass distributions can also be probed directly by studying the gravitational distortion of light from background sources by foreground masses. This technique is called *weak lensing* [76, 77] with the result shown in blue in Fig. 2.3. The separation of the dominant baryonic mass (red) and total mass (blue) shows that there must be some additional dark component that is approximately collisionless, unlike the hot gas. For this reason the bullet cluster is commonly cited as the best evidence for *collisionless* particle DM.

Large Scale Structure

Structure formation in the Universe is a hierarchical procedure. Small clumps form first which then gradually combine and grow to form ever larger objects. The timescale for these procedures is crucial for the motivation of missing mass. If the only sources of gravitational potentials were that of normal baryonic matter, clumping into galaxies and galaxy clusters would proceed at a much slower rate than what is observed. These comparisons are performed with respect to numerical N-body cosmological simulations, which were made possible by advancements in computer technologies throughout the early 20th century [78]. These simulations found that a Universe with ten times more matter other than the observed baryonic matter matches well with observations. By the end of 70's this missing mass problem was collectively referred to as DM, without limiting its interpretation to particle DM or any other solution [79]. Throughout the rest of chapter 2 we use the term DM to refer to the missing mass problem associated with all observations from this section.

In addition to providing evidence for large abundances of DM, we can learn more about the required properties of this matter. As the dominant source of gravity throughout the Universe, the particular sizes and number densities of the structures we observe can be used to constrain DM. Crucially though, the simulations must be reliable enough to compare directly with observations. It was therefore not until the 70's and 80's that simulations became a useful tool to constrain DM. The dominant, and perhaps only, interaction for DM is gravity. Simulations are therefore insensitive to its particular particle properties. Luckily they are highly sensitive to the initial velocity distribution [80, 81]. For instance, the larger the average velocity of DM particles, the greater the suppression of small structures. This is important since DM cannot cool through radiative processes like baryonic matter. Assuming ΛCDM cosmology, this then forces DM to be cold (or warm) such that structures on small scales can form. Simulations are therefore the primary tool for ruling out SM neutrinos as DM in the Universe [78]. More recently, small scale structure, as inferred from lyman-α observations have placed further constraints on other BSM candidates [82, 83] and how *warm* they can be.

Cosmic Microwave Background

The CMB comes from light at the beginning of the Universe, approximately 300,000 years after the Big Bang. The Universe is observed to be expanding, therefore going backwards in time reduces the volume of the Universe. The expansion of spacetime causes the energy density of a system to decrease, also decreasing the temperature. We can therefore view going back in time as an increasing temperature scale. In particular, the temperature of the Univers is related to the scale factor as $T \propto a(t)^{-1}$.

After the Big Bang the Universe is a hot soup of fundamental particles. These

particles gain their mass during the Electroweak phase transition and continue to cool until the strong interaction becomes important, allowing baryons and mesons to form. The formation of baryons and mesons in the early Universe is known as the QCD phase transition. At around 1 keV, the Universe is filled primarily with radiation, with sub-dominant baryonic and DM components. The baryonic matter and radiation are tightly coupled through Thompson scattering ($e^- + \gamma \rightarrow e^- + \gamma$), whereas the DM only interacts gravitationally here i.e. it has already *frozen out* of the system. The amount of neutral hydrogen is moderated by $e^- + p^+ \rightarrow H + \gamma$ as well as the reverse process caused by high energy photons. As the Universe expands and the temperature cools, there is not sufficient energy to disintegrate hydrogen and it's formation becomes energetically preferred. This process is called *recombination*. In conjunction, the dramatic reduction in the abundance of free electrons causes Thompson scattering to become inefficient, allowing photons to free stream. These are the photons we observe today as the CMB[8].

The success of DM in the CMB comes from studying its statistical anisotropies over the sky. The average temperature of these photons is observed to be extremely uniform, but deviations of $10^{-5}K$ can be found. Gaussian anisotropies are characterised completely by their angular power spectrum, a measure of the power at different angular scales. The structure of these anisotropies is dependent on a number of processes, the most important being *baryonic acoustic oscillations* (BAO). BAO are caused by the time varying density perturbations of the tightly coupled baryon-photon plasma[9]. This plasma collapses into the gravitational potential, established by the DM, until enough radiation pressure is built up, causing a rebound and expansion of the plasma. These structures can be decomposed into their spherical harmonic modes which vary in time. The oscillations continue until recombination, at which point the state of the system is *frozen in*. The structure of the angular power spectrum is dictated by the oscillations scales that happened to be at their maxima or minima, reflected as peaks and troughs in Fig. 2.4.

This is reflected in observations by slightly hotter and cooler CMB photons corresponding to over and under dense regions respectively[10]. The angular power spectrum reflects this structure with peaks and troughs at different angular scales. Crucially, the ratio of the peaks heights are highly dependent on the abundance ratio of baryonic matter to DM [85]. We therefore use the CMB to estimate the abundance of DM in the Universe. Miraculously this turns out to be precisely the amount of DM required to explain observations at smaller scales and later times. The standard cosmological evolution used to predict the CMB anisotropies requires that DM behave like a pres-

[8] They are not emitted in at microwave energies but have been subject to gravitational redshift throughout the expansion of the Universe.

[9] Remember that these are coupled through Thompson scattering.

[10] Note that this did not need to be the case since the photons in more dense regions need to climb out of a larger potential well, losing energy in the process.

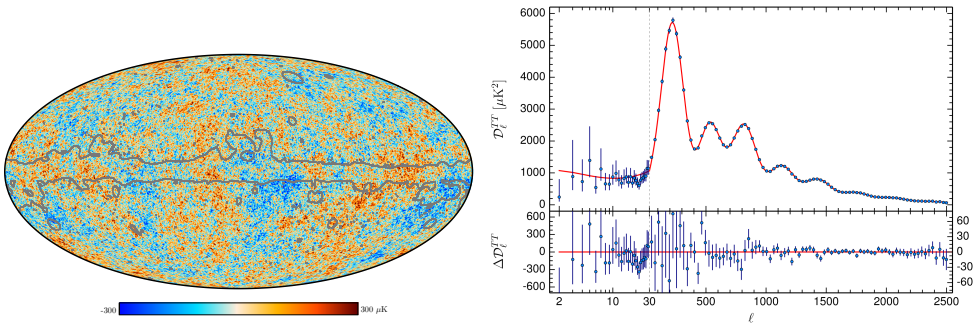

Figure 2.4: **Planck CMB observation (left) and angular power spectrum (right)**. *Left:* 2018 observations of the Cosmic Microwave Background performed by the Planck Satellite. The grey lines indicate the region masked by the galactic plane. *Right:* Red points indicate the angular power spectrum as calculated from the left image. The red line indicates the best fit model from ΛCDM cosmology and the bottom panel indicates the residuals. Image credit goes to Ref. [30, 84]

sureless collisionless fluid that does not substantially interact with normal matter. By itself the CMB therefore also represents a missing mass problem, but in combination with all other observations provides very strong evidence for particle DM outside of the SM.

2.2.2 Gravitational Lensing

All the evidence for missing mass can be accounted for by macroscopic compact objects such as primordial BHs or Brown dwarfs, collectively known as Massive Compact Halo Objects (MACHOs). If MACHOs are to make up all the missing mass, they must be extremely abundant and not emit much radiation. Luckily, their abundance makes them easily testable though gravitational *microlensing*, another form of light distortion from background sources similar to weak lensing. Unlike weak lensing, the effect is simply an observed luminosity amplification of the background source as the lense transits the line of sight. Observing campaigns in the 90's found that the rate at which these microlensing events occur is too low to be consistent with DM [86, 87, 88].

Interestingly MACHOs have had a recent revival of interest following the discovery of gravitational waves by the LIGO/VIRGO collaboration [89]. The first event, GW150914, was consistent with the merger of two $\mathcal{O}(10)\,\mathrm{M}_\odot$ BHs, precisely in the range of parameter space that would allow primordial BHs to account for all of the DM [90] whilst evading all current constraints. Subsequent efforts have been made to exclude this possibility [91, 92, 93]. Though many of these have caveats and assumptions, the general picture that these primordial BHs cannot be *all* of the DM in the

Universe is slowly taking shape.

2.2.3 Modifications to Gravity

The laws of gravity are governed by Einstein's theory of General Relativity (GR). In many of the astrophysical and cosmological scenarios mentioned above it is possible to add modifications to these laws that match the observations. Indeed it may seem more tempting to attempt modifications to GR than introduce new particles. Here I argue that modifications to gravity for the sole purpose of explaining DM should carry less motivation than some classes of DM candidates. In particular these DM particle candidates should be associated with known issues within particle physics, such as the hierarchy or strong CP problem.

The structure of gravitational laws are tightly constrained by basic physical principles such as locality and unitarity. In particular, the infrared (IR)[11] behaviour of any theory governed by a massless spin-2 particle (such as the graviton) is uniquely determined to be GR [94, 95, 96, 97, 98]. There are several ways to see this, with a variety of assumptions. We will sketch the simplest approach here (discussed in [94]) by imposing physical assumptions onto the degrees of freedom of a spin-2 massless particle.

A massless spin-2 particle has *two* independent degrees of freedom (dof) associated with its chirality [99]. We want to encode these two dof into the symmetric tensor $h_{\mu\nu}$, for which there are ten. This field will encode the dynamics of the graviton. To reduce the dof to the required two, we first introduce linearized diffeomorphisms

$$h_{\mu\nu} \sim h_{\mu\nu} + \alpha_\mu p_\nu + \alpha_\nu p_\mu \,, \tag{2.4}$$

where we also impose $\alpha \cdot p = 0$ and p^μ is the four-momenta of the graviton. Equation 2.4 provides four constraints on $h_{\mu\nu}$, reducing the dof to six. This must be further reduced by introducing the manifestly Lorentz invariant constraint

$$p^\mu h_{\mu\nu} = 0 \,. \tag{2.5}$$

Equation 2.5 reduces the dof by four, leaving two remaining. Next we want to find a scalar quantity which is invariant under these transformations from which we can build a Lagrangian. The invariant quantity turns out to be the Riemann tensor [94],

$$R_{\mu\nu\alpha\beta} = \frac{1}{2}\left(p_\mu p_\alpha h_{\nu\beta} - p_\mu p_\beta h_{\nu\alpha} - p_\nu p_\alpha h_{\mu\beta} + p_\nu p_\beta h_{\mu\alpha}\right) \,. \tag{2.6}$$

[11] Here I will refer to long distance/low energy scales and infrared (IR), and small distances/high energies as ultraviolet (UV)

Finally we can make the invariant quantity known as the Ricci scalar $R = \eta^{\nu\beta}\eta^{\alpha\nu}R_{\mu\nu\alpha\beta}$. We can use this scalar as the most naive starting point for a Lagrangian formulation of GR given by the *Einstein-Hilbert* action

$$S = \frac{1}{16\pi}\int d^4x\sqrt{-g}R\,, \tag{2.7}$$

where g is the determinant of the metric. Equation 2.7 can be shown to produce Einstein's field equations when coupled to the matter sector. Any generalisations of this will most likely be introducing new degrees of freedom in the form of a scalar field [95, 96, 97, 100]. Note though that higher order derivative terms can appear but should be suppressed by powers of the Planck mass and are therefore irrelevant for IR interactions [101]. In fact, this is important since naively one could construct a scalar by instead contracting two Reimann tensors which would give you four derivative terms instead. The most common incarnation of the uniqueness of Einstein's GR is called Lovelock's theorem [102], although there are now many ways to see this [94, 95].

Modifications to Newtons Laws have gained traction due to some suspicious numerical coincidences and observed trends. The most famous of these is the connection between the Hubble horizon and scale at which DM is required to explain galactic rotation curves. These theories are often collectively called Modified Newtonian Dynamics (MOND) which are characterised by a modification to the Newtonian potential $\Phi_{\rm N}$ at some acceleration scale a_0,

$$\frac{a^2}{a_0} = \nabla\Phi_{\rm N}\,, \quad a_0 \sim \frac{1}{6}cH_0, \tag{2.8}$$

where c is the speed of light in a vacuum and H_0 is the Hubble constant today. Additionally, the empirical trend between the observed mass/luminosity of a galaxy and its asymptotic rotational velocity $v^4 = a_0 G M_b$ (M_b is the baryonic mass of a galaxy) has led to further speculation that there is an underlying law to relate these systems more universally. This scaling law is known as the Baryonic Tully-Fisher relation [103]. It should be noted that although particle DM candidates do not generically predict this relation, modern simulations have found that it may be a natural consequence of structure formation [104, 105]. These theories can be appealing since they explain the dynamics without appealing to additional unobserved particles *i.e.* the observed baryon abundance can explain entirely the dynamics of the system. Unfortunately, these same successes have not been matched for larger scale cosmological observations such as the CMB. Additionally, the most naive incarnations may violate Lovelock's theorem by modifying gravity in the IR.

One of the viable alternative theories of gravity that maintains the structure of MOND on galactic scales is known as scalar-vector-tensor (TeVeS) gravity. By adding at least a new degree of freedom in the form of a scalar field it is not bound by

the normal constraints on gravitational interactions. These new degrees of freedom though are adhoc, purely introduced to account for the effects of DM. Although technically this is fine, the most well motivated particle DM candidates naturally arise from solutions to fundamental particle physics issues. It is for this reason that they deserve a more credible foothold in the space of theories. In addition, there is no convincing theory of non-particle DM that can accurately model all peaks of the CMB power spectrum shown in Fig. 2.4.

2.3 Particle Dark Matter Candidates

Little is known about the particle nature of DM. Here we review the necessary features of a particle to become a DM candidate. We then discuss two broad classes of candidates with natural motivations from other areas of particle physics that also serve as viable DM candidates.

The most constraining observation of DM is its relic abundance in the early Universe. To satisfy current observations it should also:

- Be stable over cosmological timescales.
- Be almost electrically neutral.
- Interact only weakly with the SM.
- Have self interactions small enough to satisfy constraints from the Bullet cluster.
- Be non-relativistic at the point of decoupling from the SM in the early Universe.
- Have a viable production mechanism whilst preserving the observed structures of the CMB.

For a theory to provide a good particle candidate it must therefore have a viable production mechanism whilst preserving the observed structures of the CMB. There are a wide variety of postulated DM particle models in the literature [106, 107, 108] although their detailed descriptions are not important here.

A simple way to classify DM particle candidates is by their production mechanism:

- Thermal production (Freeze-out) [107, 109]
- Non-thermal production (Freeze-in) [110, 111].
- Misalignment Mechanism [112, 113, 114].

Here, we discuss two of the most popular DM candidates, WIMPs and axion DM. For each, we summarise its particle physics motivation and production mechanism.

2.3.1 WIMPs

The *hierarchy problem* is one main arguments to expect new physics at the weak scale. It deals primarily with the smallness of the observed Higgs mass, $m_h \sim \mathcal{O}(100)\,\text{GeV}$, compared to the Planck mass, M_{pl}. The Planck mass is set by the scale at which gravitational interactions and SM interactions become comparable, therefore requiring a new physical description. The observed mass of the Higgs boson contains contributions from tree level diagrams plus quantum loop corrections,

$$m_h^2 = m_{h,0}^2 + m_{h,1}^2 + \cdots \sim m_{h,0}^2 + \frac{\lambda}{16\pi^2}\Lambda^2 + \dots\,, \tag{2.9}$$

where Λ is the high energy cut-off of the effective theory, and λ is an $\mathcal{O}(1)$ dimensionless coupling constant. Luckily, it is possible to remove the cut-off dependence with an appropriate choice of regularization. The real essence of the problem is simply the quadratic dependence of the Higgs mass to higher mass scales. Imagine instead a theory with an additional heavy scalar field with mass M. Integrating this out of the theory will produce a different correction to the Higgs mass given by,

$$m_h^2 \sim m_{h,0}^2 + \frac{\lambda}{32\pi^2}\left[\Lambda^2 + M^2 \ln\left(\frac{\Lambda^2}{M^2}\right)\right] + \dots\,, \tag{2.10}$$

where the M^2 can no longer be removed by a choice in regularisation [115]. Higher order loop corrections can make this sensitivity even worse. This sensitivity to high scale physics suggests new physics at the weak scale, an extremely finely tuned theory, *or* no new particles with masses above the elctroweak scale. Supersymmetric WIMPs are of particular interest here as a basic extension to the SM that solves the hierarchy problem. Supersymmetry was first proposed as the natural extension of the Poincaré symmetry to the super Poincaré symmetry. With this comes a copy of the SM which also contributes to the measurable Higgs mass but with the opposite sign to that of the SM, therefore exactly cancelling the problematic quadratic divergence. Simultaneously, most incarnations of supersymmetry provide a weak scale particle that is stable with no charge. This particle is often referred to as the *neutralino* and is the archetypal WIMP candidate.

WIMPs are, by definition, heavy particles that have tree-level interactions with the weak gauge bosons W and Z but do not take part in the strong or electromagnetic forces at leading order. If this type of particle exists, it is produced in the early universe with a relic abundance consistent with the observed DM abundance. This generic production mechanism is coined the *WIMP miracle* and is one of the primary reasons WIMPs have received so much attention as a DM particle candidate.

We can use an illustrative example of thermal freeze out in the early Universe

[116]. Let us assume that at high temperatures, DM is kept in thermal equilibrium with the standard model through the interaction,

$$\chi + \bar{\chi} \leftrightarrow \mathrm{SM} + \mathrm{SM}\,, \tag{2.11}$$

where χ and $\bar{\chi}$ indicates the WIMP and its antiparticle[12]. Whilst thermal equilibrium is maintained in an expanding universe, the comoving number density of DM, $N_\chi = n_\chi/s$[13], is described by the Boltzmann equation,

$$\frac{\mathrm{d}N_\chi}{\mathrm{d}t} = -s\langle\sigma v\rangle\left[N_\chi^2 - \left(N_\chi^{\mathrm{eq}}\right)^2\right]\,, \tag{2.12}$$

where $\langle\sigma v\rangle$ is the thermally averaged annihilation cross section and N_χ^{eq} is the number density of DM in thermal equilibrium. The entropy density s is given by

$$s = \frac{2\pi^2}{45} g_{\star S}(T) T^3\,, \tag{2.13}$$

where $g_{\star S}(T)$ is the effective number of degrees of freedom in entropy. Equation 2.12 describes the effect of two process; the dilution of the number density during expansion and the annihilation and production from both processes in Eq. 2.11. *Thermal freeze out* is defined when the interaction in Eq. 2.11 becomes inefficient. In thermal equilibrium at low temperatures, there would be an exponential suppression of the number density of heavy DM particles since Eq. 2.11 would be much more efficient from left to right. Thermal freeze out is when the annihilation rate also becomes inefficient. This happens when the Universe expands and DM particles becomes increasingly less likely to meet.

Assuming that the relic abundance of DM is much lower than at freeze out and that freeze out occurs during the radiation era we can write the relic abundance as

$$N_\chi^\infty \simeq \frac{x_f}{\lambda}\,, \tag{2.14}$$

where $x = m_\chi/T$ and f indicates the quantity at freeze out. We have also defined

$$\lambda \equiv \frac{2\pi^2}{45} g_{\star S} \frac{m_\chi^3 \langle\sigma v\rangle}{H x^2}\,. \tag{2.15}$$

Using the above result and assuming that the WIMP number density is conserved we

[12] Note that the WIMP can be its own anti-particle.
[13] n_χ is the regular number density

can write the relic density today as

$$\Omega_X h^2 \sim 0.1 \left(\frac{x_f}{10}\right) \left(\frac{10}{g_\star\left(m_\chi\right)}\right)^{1/2} \frac{10^{-8}\,\mathrm{GeV}^{-2}}{\langle\sigma v\rangle} \tag{2.16}$$

where $g_\star$ is the effective number of relativistic degrees of freedom. From here we see that a thermally averaged cross section close to the weak scale would produce the observed DM abundance today. This is known as the *WIMP miracle* [116, 41].

2.3.2 Axions

The QCD lagrangian admits the CP-violating term

$$\mathcal{L} = -\frac{g_s^2\theta}{32\pi^2} G_{\mu\nu}\tilde{G}^{\mu\nu}\,, \tag{2.17}$$

where θ is the coupling term describing the CP-violating phase and $G_{\mu\nu}$ is the gluon field strength tensor (the tilde indicates its dual). This term should induce a neutron electric dipole moment but current observations have shown this to be extremely tightly constrained, $|\theta| < 10^{-10}$ [117, 118]. Since naively $\theta \sim \mathcal{O}(1)$, this is a fine tuning problem. The solution, proposed by Pecci and Quinn, is to promote θ to a field which can dynamically relax to the CP-conserving value below the symmetry breaking scale f_a [112]. The breaking of the Pecci-Quinn symmetry gives rise to a Goldstone boson, called the QCD axion. It is generally thought to have generic couplings to both gluons, quarks, and electromagnetism whilst its mass is connected to the decay constant $m_a \propto 1/f_a$.

Axions generally behave like non-relativistic matter and therefore are a primary DM candidate. In theory they can be produced non-thermally, although the favoured production scenario is known as the misalignment mechanism. Above the symmetry breaking scale, all values of the axion field are allowed since it is not yet needed to solve the strong CP problem. After the Pecci-Quinn symmetry breaking, the field relaxes to its minimum and generates a relic density of axions. If inflation occurs after the symmetry breaking then the Universe has one relic density given by [119, 107],

$$\Omega_a \sim 0.4\theta_i \left(\frac{f_a}{10^{12}\,\mathrm{GeV}}\right)^{1.18}\,, \tag{2.18}$$

where θ_i is the misalignment angle, a $\mathcal{O}(1)$ initial condition.

2.4 Direct Detection of Dark Matter

The Direct Detection (DD) of Dark Matter is one of the three primary methods currently employed to detect WIMP interactions. It was first proposed by Goodman and Witten [120] and subsequently developed by Drukier, Freese and Spergel [121] . The basic idea is to build a detector quiet and large enough to detect the very rare DM interactions with the target material. By measuring these scattered particles and their kinematics one can hopefully detect, and infer the properties of, DM. Here we briefly review the theory of nuclear scattering for use in Chapters 5, 6, and 7. We introduce the non-relativistic effective theory of DM interactions (as proposed in [122, 123]) for use in Chapter 5 and emphasise the spin independent interaction term which we explore further in Chapters 6 and 7. Finally, we briefly review the experimental status of DD searches.

2.4.1 Dark Matter Scattering Theory

Heavy, $\mathcal{O}(1)$ GeV, DM is observationally determined to be non-relativistic within the Galaxy [124]. We will therefore focus on non-relativisitc DM-nucleon elastic interactions. For more detailed derivations see Refs. [125, 126, 127, 128, 122, 123]. Importantly, we focus on DM-nucleus interactions rather than interactions between DM and electrons, although in principle these can also be detected [129, 130]. Figure 2.5 shows an elastic interaction between a DM particle, χ, and nucleus, N. Most DD experiments are cooled to very low temperatures, therefore we assume that the kinetic energy of the nucleus is negligible. The initial and final energies of the system are given by

$$E_{\rm ini} = \frac{\mathbf{p}'^2}{2m_\chi}, \qquad E_{\rm fin} = \frac{\mathbf{p}^2}{2m_\chi} + \frac{\mathbf{q}^2}{2m_N}\,, \tag{2.19}$$

where bold indicates the momentum vectors and m indicates the mass of the DM particle or nucleus. By momentum conservation we find $\mathbf{q} = \mathbf{p}' - \mathbf{p}$, where $\mathbf{q}$ is often referred to as the momentum transfer. Defining $\hat{\mathbf{p}}' \cdot \hat{\mathbf{q}} = \cos\theta$ and equating final and initial energies we find,

$$\frac{\mathbf{q}^2}{2\mu_{\chi N}} = \frac{|\mathbf{p}'||\mathbf{q}|\cos\theta}{m_\chi}\,, \tag{2.20}$$

where $\mu_{\chi N} = m_\chi m_N/(m_\chi + m_N)$ is the reduced mass of the DM-nucleus system. We can now find the maximum energy transfer for a single recoil

$$E_R^{\rm max} = \frac{|\mathbf{q}|^2_{\rm max}}{2m_N} = \frac{2\mu_{\chi N}^2 v^2}{m_N}\,, \tag{2.21}$$

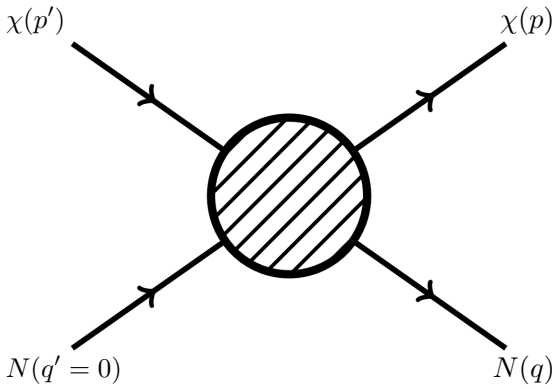

Figure 2.5: **Feynman diagram for Dark Matter Direct Detection**. The vertex depicts a generic elastic interaction between DM χ and a nucleus N. The different forms of this interaction are explored in the text.

where v is the magnitude of the velocity of the incoming DM particle which is typically taken to be $v \sim 3 \times 10^{-3}$ in the lab frame [124]. Plugging in typical numbers for the DM and nuclear masses we find the maximum recoil energy is $\mathcal{O}(10 - 100)\,\text{keV}$.

For a given energy we can define the minimum velocity a DM particle must have to induce that recoil:

$$v_{\text{min}} = \sqrt{\frac{E_R m_N}{2\mu_{\chi N}^2}}\,. \tag{2.22}$$

The maximum velocity of DM is given by the escape velocity of the galactic potential. Together with the DM velocity distribution, $f(\mathbf{v})$, (that encodes the kinematics of the expected interactions) we can write the recoil spectra per unit target mass as

$$\frac{\mathrm{d}R}{\mathrm{d}E_R} = \sum_i \frac{\xi_i \rho_{\text{DM}}}{m_N^i m_\chi} \int_{v_{\text{min}}}^{v_{\text{max}}} \mathrm{d}^3 v f(\mathbf{v}) v \frac{\mathrm{d}\sigma}{\mathrm{d}E_R}\,, \tag{2.23}$$

where ξ is the mass fraction of the nuclide i, ρ_{DM} is the DM density at the Earth's position in the galaxy, and $\mathrm{d}\sigma/\mathrm{d}E_R$ is the DM-nucleus recoil rate.

There are two approaches to calculating $\mathrm{d}\sigma/\mathrm{d}E_R$; the first involves writing down a microscopic DM particle theory that interacts with gluons or quarks which can then be matched on to the corresponding operators involving nucleons from which we can calculate the final recoil rate. The second method, and one we will adopt, is to directly write down all the non-relativistic effective operators and there nuclear responses.

The non-relativistic differential cross section can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_R} = \frac{1}{32\pi}\frac{1}{m_\chi^2 m_N v^2}|\mathcal{M}_{\chi N}|^2\,, \tag{2.24}$$

where $\mathcal{M}_{\chi N}$ is the matrix element that encodes the information of the interaction type and the relevent nuclear physics. The general Lagrangian for interactions with nucleons, $\mathcal{N}$, takes the form

$$\mathcal{L} = \sum_i c_i^{\mathcal{N}}\mathcal{O}_i\,, \tag{2.25}$$

where $\mathcal{O}_i$ represent the different non-relativistic operators and the coefficients c_i are functions of the relativistic theory. The operators themselves are built to be Galilean and rotationally invariant, containing only the relevant degrees of freedom of the system. For a DD experiment the relevant degrees of freedom are; the relative velocity between the DM and nucleus $\sim \mathbf{v}$, the momentum exchange $\mathbf{q}$, the spin of the DM $\mathbf{S}_\chi$, and the spin of the nucleon $\mathbf{S}_{\mathcal{N}}$. Averaging over spins, we can write the matrix element as

$$|\mathcal{M}_{\chi N}|^2 = \frac{m_N^2}{m_{\mathcal{N}}^2}\sum_{i,j=1}^{16}\sum_{\mathcal{N},\mathcal{N}'=p,n} c_i^{\mathcal{N}}c_j^{\mathcal{N}'}F_{ij}^{\mathcal{N},\mathcal{N}'}\,, \tag{2.26}$$

where $F_{ij}^{\mathcal{N},\mathcal{N}'}$ are the nuclear response functions that critically depend on the nuclear structure of the target material. These factors were computed for some of the most commonly used DD experimental materials in [123]. The scattering amplitude can be written in this way purely because of the non-relativistic expansion in powers of the momentum transfer and DM velocity. Additionally, the coefficients of Eq. 2.26 have momentum dependence only through the scalar quantities $\mathbf{v}^2$ and $\mathbf{q}^2$.

The two most well known operators[14] used in the DD community are $\mathcal{O}_1$ and $\mathcal{O}_4$ corresponding to the 'spin-independent' (SI) and 'spin-dependent' (SD) interactions respectively. They are given by,

$$\mathcal{O}_1 = \mathbb{1}\,, \qquad \mathcal{O}_4 = \mathbf{S}_\chi\cdot\mathbf{S}_{\mathcal{N}}\,. \tag{2.27}$$

For the spin-independent case we can write the total recoil rate in its common form,

$$\frac{\mathrm{d}R}{\mathrm{d}E_R} = \sum_i \frac{\xi_i\rho_{\mathrm{DM}}}{2m_\chi\mu_{\chi\mathcal{N}}^2}\sigma_{\mathrm{SI}}\int_{v_{\min}}^{v_{\max}}\mathrm{d}^3\mathbf{v}\frac{f(\mathbf{v})}{v}A_N^2F_{\mathrm{helm}}^2(E_R)\,, \tag{2.28}$$

where $F_{\mathrm{helm}}(E_R)$ is the Helm form factor [131], A_N is the atomic number of the nucleus, and σ_{SI} is the spin independent cross-section[15]. We plot the recoil spectra

[14] These are the most commonly used since they are the lowest order in the non-relativistic expansion.

[15] Note that this is the coupling to protons which is why we get an A_N^2 enhancement to the recoil rate.

for the first eleven operators in Fig. 2.6, excluding $\mathcal{O}_2$ for convenience of plotting and its lack of obvious UV completion due to its v^2 dependence. These spectra, possible extensions to this formalism, and the velocity distribution of DM will be discussed further in Chapter 5.

2.4.2 Experimental Status

Dark Matter scattering events are rare due to their extremely weak coupling to the SM. To detect these events experiments must therefore be large and extremely quiet. The primary background is from cosmic ray interactions. To remove this, experiments are generally built in underground laboratories using the Earth as a shield. Unfortunately, there are radioactive elements within the Earth itself, providing an additional background. Experiments must therefore be shielded against these background sources and be able to detect multiple scatters, the latter being used to veto SM events. Another primary source of nuclear recoils is natural radioactivity of the target material. This must be reduced to exceptionally low levels and constantly monitored [132, 133]. Finally, experiments often aim to distinguish between electromagnetic and nuclear recoils, since the latter is induced by DM interactions.

There are three primary ways to detect the DM-nucleus interaction. Firstly, the emission and detection of scintillating photons from the de-excitation of the recoiling atom. Secondly, electrons can be ionized from the atom which are then drifted and subsequently detected. Finally, the small energy deposition from the interaction can slightly warm the target material which can be detected, normally referred to as phonon detection. We discuss one of the most successful detector types here, two-phase liquid noble experiments, but refer the reader to Ref. [134, 135] for a more complete review.

Two-phase liquid noble experiments have been the most successful experimental design to gain sensitivity to $\mathcal{O}(1-100)\,\text{GeV}$ scale WIMP DM. The most prominent examples are XENON1T [18], LUX [136], and PandaX-II [20] although future experiments such as DarkSide-20k will have competitve sensitivity [137].

In these detectors, a nuclear recoil will produce a scintillation and ionization signal. In Time Projection Chambers the scintillation is directly measured by photon multiplier tubes at the edge of the detector. Ionization electrons are drifted in an applied electric field and detected in the gas chambers at either end of the detector. By summing these two signals, the recoil energy of the interaction can be reconstructed whilst the timing of the two signals can be used to localise the event within the detector. Localisation allows for easy vetoing of potentially spurious events close to the detector wall.

Despite exceptional achievements in reducing backgrounds in Xenon experiments, some still remain. The most prominent of the nuclear recoil backgrounds is that of ra-

diogenic neutrons from the detector materials. In addition, there are coherent elastic neutrino nucleus scatters mainly from solar neutrinos as well as cosmogenic neutrons from secondary particles produced by muon showers outside the detector. For further discussion of backgrounds see Ref. [18] and references therein.

As the name indicates, noble liquids are used in these detectors, the most popular of which is Xenon. Xenon is used due to its sensitivity to both SI and SD interactions. Sensitivity to SD interactions comes from the unpaired neutron in some of the Xenon isotopes. Additionally, the mass of a Xenon nucleus is kinetically favourable for $\mathcal{O}(10-100)$ GeV DM interactions as well as receiving a significant coherent enhancement due to the large nucleon number. Figure 2.7 shows the sensitivity of current experiments to the SI cross section (above) and SD cross section (below). Experiments will continue to gain sensitivity to both of these interactions in the next generation of detectors (currently under construction). We discuss an alternative approach, known as *paleo-detectors*, in Chapters 6 and 7. Finally, the yellow shaded region of Fig. 2.7 shows the *neutrino floor*; the point at which neutrino interactions will become a primary nuclear recoil background that cannot be removed. The neutrino floor has three primary contributions from different neutrino sources: solar neutrinos, atmospheric neutrinos, and the so called *diffuse supernovae background* (DSNB). These neutrinos are of primary interest if they can be detected and characterised. This the topic of Chapter. 7 and Ref. [4].

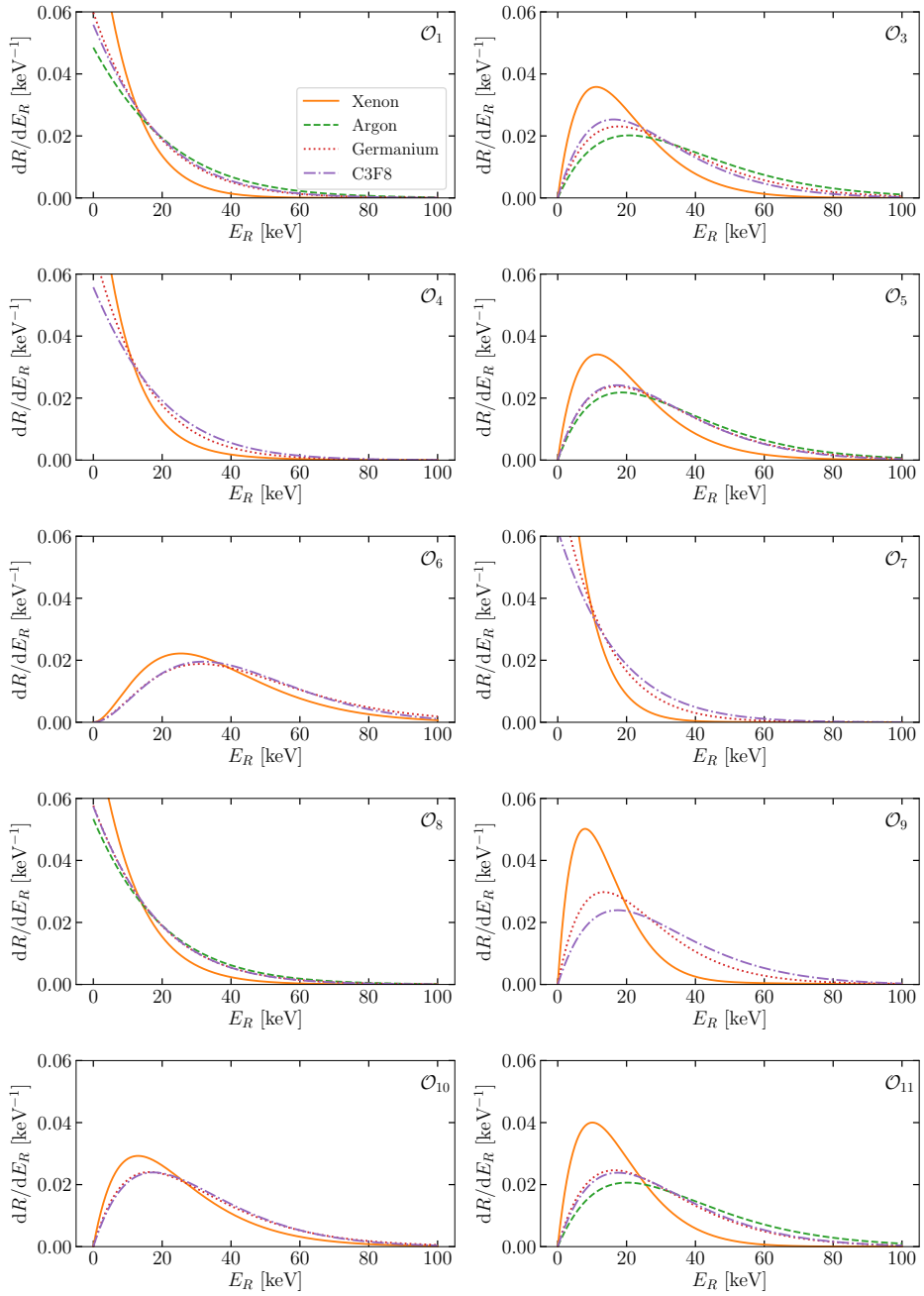

Figure 2.6: **Recoil spectra for 50 GeV DM particle**. The different panels represent the eleven main operators of the NREFT. $\mathcal{O}_2$ is excluded $\mathcal{O}_2$ for ease of plotting and unlikely UV completion.

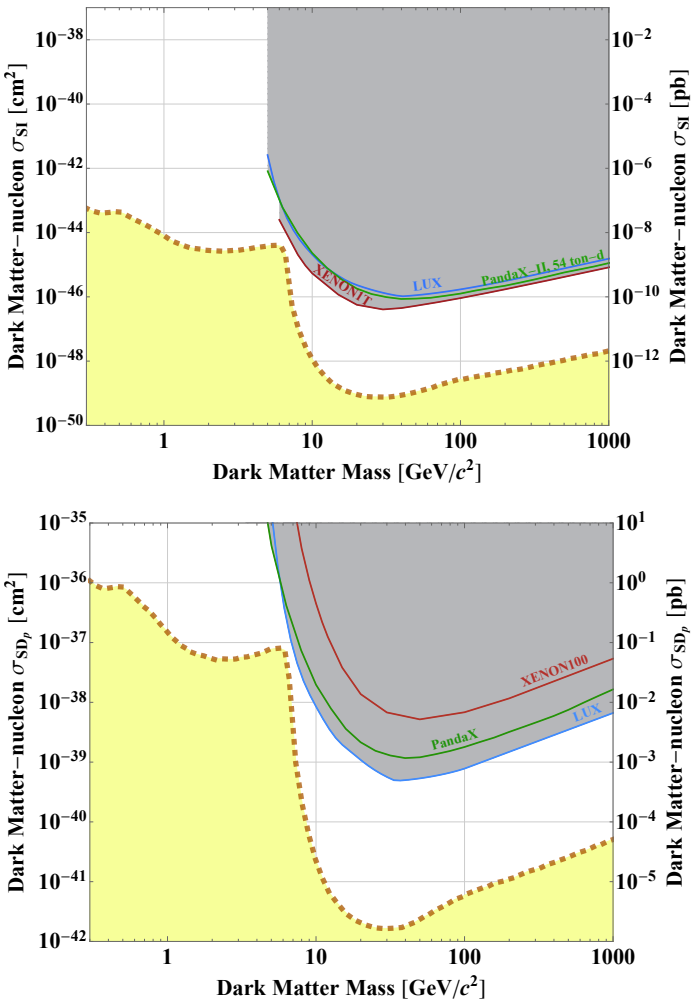

Figure 2.7: **Dark Matter DD Limits on spin-independent (above) and spin-dependent (below) cross sections**. Limits taken from Ref. [138, 23, 139]

3 Statistics

The purpose of this chapter is to lay the foundations of the statistics that will become important throughout the thesis. In particular, we will focus on the procedures used in Dark Matter searches (and throughout this thesis) such as setting exclusion limits, projecting exclusion limits, and calculating the ability of experiments to discriminate between models. Together with particle physicists, these methods are generally employed from a Frequentist perspective on statistics. Bayesian statistics, on the other hand, dominates fields such as Cosmology. We therefore begin by introducing both, but focus on the Frequentist approach, whilst making a connection to Bayesian statistics towards the end. For a concise review of the methods used in collider physics see Ref. [140].

3.1 Introduction

Without a deep understanding of statistics, complex experiments would struggle to find evidence for new physics. Statistics is the mathematical formalism to connect our physics models to experiments and data. Statistical procedures are the necessary step to quantify the information gain from an experiment but also crucial in planning experimental setups and procedures. Unfortunately statistical methods are normally treated as an after thought, a necessary evil. This is an oversight, since data processing and statistical procedures can be time consuming and difficult. Instead, statistical methods should be incorporated into the framework of physics models, therefore playing a central role for both theorists and experimentalists alike.

Physicists use mathematical *models* to describe the world around them. They are predictive devices, most of the time taking some initial data which can then be used to predict, for example, a systems time evolution. These models are the way physicists are taught to view the world. We will focus on the logical structure of the statistical procedures throughout, providing mathematical proofs when they are concise or references for longer derivations. Finally, we will try to keep the language

simple and accessible introducing all the technical terms used. For a review of statistical methods used in particle physics see Ref. [140]. Many of the methods used in particle physics were developed in Ref. [17].

3.2 Technical Concepts

For a rigorous discussion we first need to clarify some important definitions. First of all, our definition of probability is extremely important.

Within statistics, there are two primary camps; the Bayesians and the Frequentists. Although their differences are usually phrased around the methods they use, the primary difference is actually much deeper, stemming from the very definition of probability. A Frequentist views nature as exact, with precise values for particular quantities such as the masses of particles. The definition of probability is instead assigned to the notion of repeated trials. The typical example is flipping a coin, giving two possible results; heads or tails. The Fequentist probability associated to getting a tails $P\left(\text{tails}\right)$ is defined by its relative occurrence in a large number of trials,

$$P\left(\text{tails}\right) = \lim_{n_{\text{flips}} \to \infty} \frac{n_{\text{tails}}}{n_{\text{flips}}}\,, \tag{3.1}$$

where n_{flips} is the number of times the coin is flipped and n_{tails} is the number of times tails occurred in n_{flips} flips. Obviously we cannot actually perform infinite numbers of flips of a coin, but it is the baseline philosophy of a Frequentist that this probability will converge exactly to its true result. In practice, physicists try to replicate this behaviour by doing many experiments, hopefully approaching the a point where the relative frequency is a good approximation of the *true frequency*.

The Bayesian view differs in that probability is simply a quantification of an observers belief. There is no need to discuss the value of something in nature, instead we simply quantify how well we can measure such a quantity. Importantly, the Bayesian definition of probability requires us to use Bayes theorem [141],

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.2}$$

where $P(A|B)$ is the probability of A given B (vice versa for $P(B|A)$), $P(A)$ is the prior belief of A, and $P(B)$ is normally referred to as the evidence. Note that $P(A|B)$ is often referred to as the *posterior probability*. The important difference here is the use of the prior $P(A)$, which requires us to define a subjective belief to the occurrence of an event. Bayesian probability is closer to everyday reasoning since we generally associate some degree of belief to an event occurring based off past experience. This prior has lead to large disagreement between the communities since it can be subjec-

tive, not necessarily in keeping with scientific language of objectivity. In theory, when large volumes of data are available, Frequentists and Bayesains will agree, though this doesn't often occur for rare event searches. It is important to note that Frequentists don't dispute Bayes' theorem, rather they don't believe that the necessary quantities are available.

Finally we note that the physics community has displayed some interest in the field known as *objective Bayesianism*, where one tries to define a prior that is independent of the users belief. This allows Bayes theorem to extract the posterior probability simply from the data. Unfortunately, there is still some disagreement amongst statisticians about how to do this exactly so we will only refer to this towards the end of the chapter. See Ref. [142] for a review.

Probabilities and Likelihoods

The basic definition of probability is anything that follows the *Kolmogorov axioms*, as given in the box below. Note that these are included for completeness and not necessary for the logical flow of the rest of the chapter.

The Kolomogorov axioms are defined on the space of events E with exclusive events E_i. The three axioms are given below;

1. The probability of an event is a real non-negative number i.e. we can assign a numerical value to the probability of an event occurring:

$$P(E_i) \in \mathbb{R}, \quad P(E_i) \geq 0 \quad \forall \quad E_i \in E\,. \tag{3.3}$$

2. Upon a measurement, one the E_i's will be selected with probability one. Said more simply, when a measurement is performed the outcome needs to be specified within the space of possible outcomes:

$$\sum_{E} P(E_i) = 1 \tag{3.4}$$

3. The probability of these exclusive events can be added

$$P(E_1 \text{ or } E_2 \text{ or } \dots) = P(E_1) + P(E_2) + P(\dots) \tag{3.5}$$

Both the Frequentist and Bayesian definitions of probability follow these axioms and we will adopt the Frequentist definition unless otherwise specified. For a review of the different rules of probability that are a natural consequence of the Kolmogorov axioms, see Ref. [141].

Consider an experiment making measurements of the continuous variable x. For example, one might want to measure the time taken for a ball to hit the ground when

dropped from a specific height. Multiple measurements will give different values of x_i, different times, which can be accommodated into a quantity known as the *probability density function* (pdf) $f(x)$ which has the important property,

$$\int f(x)\mathrm{d}x = 1\,. \tag{3.6}$$

Pdf's play an important role since most experimental outcomes will have variation due to some randomness in the system. For example, if we want to measure the number of photons coming from the sun, we perform a series of observations at different times to create an ensemble of measurements. The pdf then reflects the probability of getting the different measurements within this ensemble. In physics we normally have a description for the likelihood that the data will occur given a set of physical parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \dots\}$ which we can write as

$$f(x \,|\, \boldsymbol{\theta})\,. \tag{3.7}$$

Continuing on the example of measuring the Sun, these physical parameters might contain parameters intrinsic to the Sun such as its temperature, radius, and chemical composition as well as extrinsic parameters such as the position of Earth within the solar system. These extrinsic parameters are sometimes referred to as *nuisance parameters*.

Finally, let us define the likelihood function which gives the probability that you will observe a given value of the continuous variable x as a function of the parameters $\boldsymbol{\theta}$. It is numerically equivalent to the pdf but is a function of the parameters, not the random variable. This is an important distinction as the likelihood does not possess the property shown in Eq. 3.6. We want to generalise the likeilhood to be a function of data sets with multiple measurements $\mathcal{D} = \{x_1, x_2, \dots\}$. We will write our likelihood as $\mathcal{L}(\mathcal{D} \,|\, \boldsymbol{\theta})$ but it is important to keep in mind that the dependence is on the parameters[1].

Test Statistics and Hypothesis Testing

A statistic is an object one constructs from the data to summarise the information contained within it. The *mean* of a data set is the archetypal example of a statistic. A *Test Statistic* (TS) is just the word used to describe a statistic constructed for hypothesis testing. Hypothesis testing is just the act of comparing the null hypothesis vs the alternative hypothesis. The main purpose of a test statistic is to differentiate between these hypotheses. Hypothesis testing will take two relatively distinct forms here,

1. The comparison of θ_1 and θ_2 within a parameter space to see which of the two

[1]Note also that the likelihood is often written as $\mathcal{L}(\boldsymbol{\theta}|\mathcal{D})$ to make the dependence more clear.

is preferred or if they can be distinguished with a future experiment.

2. The comparison of a background only (null) hypothesis with the background + signal (alternative) hypothesis. We therefore want to either quantify a discovery of the alternative hypothesis or rule out a region of the signal parameter space by placing an exclusion limit

We will discuss both of these situations below.

3.3 Maximum Likelihood Ratio

We want to construct an object which will maximise the difference between the null and alternative hypothesis. The most commonly used TS is the maximum likelihood ratio (MLR) defined by,

$$\mathrm{TS}(\boldsymbol{\theta} \,|\, \mathcal{D}) = -2 \ln \frac{\max_{\nu} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\nu} \,|\, \mathcal{D})}{\max_{\boldsymbol{\theta},\boldsymbol{\nu}} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\nu} \,|\, \mathcal{D})} \,, \tag{3.8}$$

where $\boldsymbol{\nu} = \{\nu_1, \nu_2, \dots\}$ are our nuisance parameters. The MLR owes its ubiquitous use to the Neyman-Pearson lemma for simple hypothesis testing. The lemma shows that for two models with no parameters, the MLR has the maximum power of discrimination. For a graphical proof of the Neyman-Pearson lemma see Ref. [140].

The behaviour of the MLR as defined in Eq. 3.8 is such that its minima coincides with the best fit values of $\boldsymbol{\theta}$. If the model were to fit the data exactly then this is at exactly zero, although randomness in the data will make sure this cannot happen. The behaviour of the MLR close to the minima will determine how to set our confidence intervals and will be discussed below.

Here it is useful to introduce a more concrete example, in particular a counting experiment. We imagine a simple signal consisting of a Gaussian bump and a smooth power law background component. This kind of situation arises regularly in DM searches such as searching for a gamma-ray bump for indirect detection, or looking for a particle resonance at the LHC. We will use the variable E which could refer to energy but could also be any other variable. Our signal (S) and background (B) are defined number of counts per energy bin and given by

$$S \propto \mathcal{N}(E_0 = 10, \sigma^2 = 2) \,, \quad B \propto E^{-1} \,, \tag{3.9}$$

where $\mathcal{N}(E_0 = 10, \sigma^2 = 2)$ refers to a Gaussian centered on $E_0 = 10$ with variance $\sigma^2 = 2$. The signal and background models can be seen in Fig. 3.1 where we show the same normalisation for the background and two different normalisation's for signal in the left and right panels. The normalisations are our variables here;

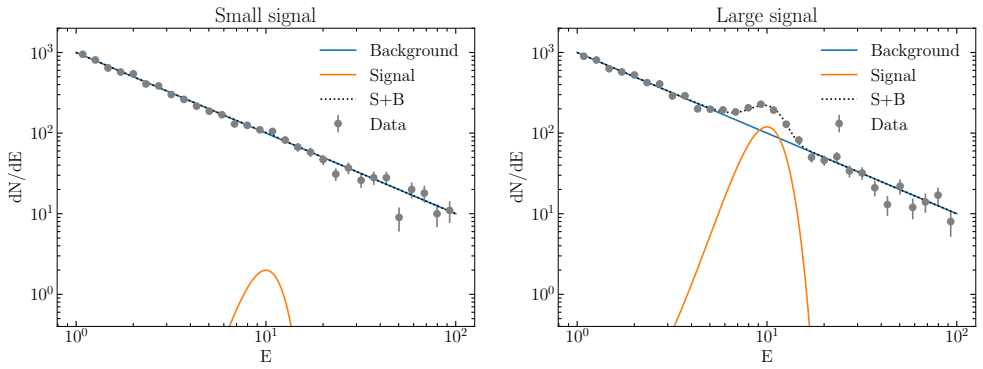

Figure 3.1: **Signal and background Example.** The signal is a Gaussian with fixed width and position. The *left* panel has a signal normalisation of $\theta_S = 10$ where as the *right* has $\theta_S = 600$. The background power law has a slope of -1 with a normalisation of $\theta_B = 1000$. The data (in the 30 log-spaced bins) is generated by sampling from a Poisson distribution with expectation value as the sum of the signal and background.

$\boldsymbol{\theta} = \{\theta_S, \theta_B\}$, where θ_S and θ_B are the normalisation of the signal and background respectively. We set these to be $\theta_S = 10,600$ for the small and large signal cases and $\theta_B = 1000$ throughout. Finally, we have $N = 30$ log spaced bins in energy ($E = \{E_1, E_2, \ldots, E_N\}$) for which we have generated data. These are created by drawing from a Poisson distribution with expectation value given by the combination of the signal and background. Our data set is then made up of a number of events in each bin i, $\mathcal{D} = \{x_1, x_2, \ldots, x_N\}$. To make use of the MLR description we need to define our likelihood. Since we are imagining a counting experiment our likelihood is described by the product of Poisson distributions over all bins,

$$\mathcal{L}(\boldsymbol{\theta} \,|\, \mathcal{D}) = \prod_{i=1}^{N} e^{-S(E_i,\theta_S)+B(E_i,\theta_B)} \frac{(S(E_i,\theta_S) + B(E_i,\theta_B))^{x_i}}{x_i!} \tag{3.10}$$

For numerical simplicity we actually work with the log likelihood given by

$$\ln \mathcal{L}(\boldsymbol{\theta} \,|\, \mathcal{D}) = \sum_{i=1}^{N} -S(E_i,\theta_S) + B(E_i,\theta_B) + (S(E_i,\theta_S) + B(E_i\theta_B,))^{x_i}\,, \tag{3.11}$$

where we dropped the constant term that only depends on the data since it will not be important for the calculation of the likelihood ratio. Imagine now that we are given the data set $\mathcal{D}$ and we want to constrain the normalisation of the signal, our

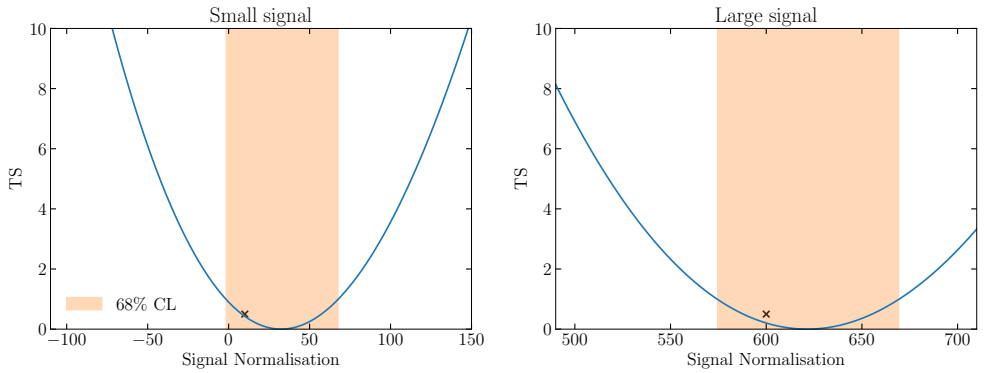

Figure 3.2: **TS distribution for model in Fig. 3.1.** Blue line shows the TS value as a function of the signal normalisation θ_S. The orange band shows the 68% confidence interval and the black cross shows the true value. In both we can see that the randomness ensures the minima of the TS is not precisely at the true parameter value.

test statistic then takes the form,

$$\mathrm{TS}(\theta_S \,|\, \mathcal{D}) = -2 \ln \frac{\max_{\theta_B} \mathcal{L}(\theta_S, \theta_B \,|\, \mathcal{D})}{\max_{\theta_S, \theta_B} \mathcal{L}(\theta_S, \theta_B \,|\, \mathcal{D})} \,. \tag{3.12}$$

The minimum of the TS gives us the best fitting value θ_S for the data set $\mathcal{D}$. The distribution of the TS for both cases is shown in Fig. 3.2. Importantly we can now see that, for a given data set, the minimum of the TS might not exactly match the *true* value of the parameter (marked as black crosses in Fig. 3.2). To capture this behaviour we need to define two new concepts, *confidence intervals* and *coverage*.

Confidence Intervals and Coverage

All data sets contain some element of randomness in them. For example, the number of photons arriving in a gamma ray telescope will sometimes be slightly larger, sometimes slightly smaller, even for a steady state source. This is simply statistical Poisson error and cannot be avoided. It is therefore inevitable that the minima of the TS for any given data set will not precisely reflect the true value of the parameter. Instead we need to define an interval of parameters over which we expect the true parameter to exist within to a given *confidence level*, Z.

A confidence interval approximately represents the interval within which one would be expect to find the true parameter with a confidence of Z. More precisely Z represents the frequency, assuming repeated experiments, that the associated confidence interval will contain the true value. Take the example above, the orange band in Fig. 3.2 shows the 68% confidence interval. This means that if one was to form

100 data sets, constructing a new interval for each, then 68 of them would contain the true parameter. The interval is defined to be

$$\{\,\hat{\theta}\,|\,\mathrm{TS}(\theta) < c\,\} \tag{3.13}$$

where c is a value set such that the confidence interval has the desired behaviour as defined above and $\hat{\theta}$ is the best fit value of θ (the minima of the TS). Finally, the probability that an interval will contain the true value is called the *coverage probability* (coverage for short). If the confidence interval and coverage are the same then the interval is said to be exactly covered. In some cases, exact coverage may not be possible and the confidence interval is said to under or over cover. Under coverage refers to the when the stated confidence level is higher than the true coverage. The opposite scenario is when the confidence interval contains the true parameter more than the stated confidence level, referred to as over coverage. If an interval over covers it is said to be conservative.

Discovery Tests

Before we move on to looking at methods to construct confidence intervals, we briefly mention the most common way to establish whether a discovery has been made or not. In particle physics these questions are normally answered through a hypothesis test where we have the null being a background only hypothesis vs the alternative being background + signal. Assuming that we have a good model of the background, we take data and quantify the probability that the observed data is compatible with the null hypothesis. If the data set is extremely unlikely to have been produced by the background alone, we would reject the null and accept the alternative. This probability is normally called the p-value and for a simple one-bin counting experiment is quantified by the number of excess events over the background. The general definition is given by,

$$p = \Pr\left(\mathrm{TS}(x) \geq \mathrm{TS}\left(x_{\mathrm{obs}}\right) \,|\, H_0\right)\,, \tag{3.14}$$

i.e. the probability that the test statistic $\mathrm{TS}(x)$ is more extreme than the observed $\mathrm{TS}(x_{\mathrm{obs}})$ under the null hypothesis.

Note that this is not a statement about the probability of the alternative being correct. For this we would require Bayesian reasoning, not commonly used amongst particle physicists. For the detection of gravitational waves they use Frequentist statements to quantify the probability of a data set occurring in a background only scenario combined with Bayesian statements about the likelihood of a signal being of astrophysical origin. A discovery is only made if both tests provide evidence for a gravitational wave. We won't discuss this further here but refer to Ref. [143] for more information.

We will now discuss three methods that utilise the MLR to construct confidence

intervals. Importantly, we are now addressing the question of how to either set an upper limit or construct a closed confidence interval, for a non-discovery or discovery respectively.

3.3.1 Neyman construction

The Neyman construction is a method to develop confidence intervals with exact coverage. Intervals constructed in this way are therefore ideal since they have precisely the desired confidence level. In an ideal scenario, one would not decide to set a two sided confidence interval or upper limit based of the data and would have a technique that smoothly transitions between the two. Feldman and Cousins developed a method to do this based around the Neyman construction [144], which we will briefly review here. Imagine an experiment that counts numbers of events, n, with a background and signal contribution b and s respectively. Constructing the confidence interval now has three steps, shown in Fig. 3.3:

1. For each value of s one defines an acceptance region of allowed values of n as described by $P(n \,|\, s)$
$$P\left(n \in [n_1, n_2] \,|\, s\right) = \alpha\,, \tag{3.15}$$
where α is the confidence level.
2. Construct a *belt* of acceptance regions of n as a function of s.
3. Upon measuring n_0, the confidence interval is set by the intersecting values of s along the belt at n_0.

For more details on the exact construction of the acceptance region and how to perform computations in more complicated scenarios see Ref. [144].

3.3.2 Large Sample Limit

The large sample limit is generally defined by having a large number of background events. In this scenario we are lucky that the distribution of the MLR is well described by a chi-squared distribution χ^2_d with degrees of freedom specified by the difference in the number of free parameters between the numerator and denominator of the MLR [145]. For example, Eq. 3.12 has one free parameter, θ_S, and is therefore described by $\chi^2_{d=1}$. We could modify the TS to instead constrain the background and signal normalisation's simultaneously,

$$\mathrm{TS}(\theta_s, \theta_b) = -2 \ln \frac{\mathcal{L}(\theta_s, \theta_b \,|\, \mathcal{D})}{\max_{\theta_s, \theta_b} \mathcal{L}(\theta_s, \theta_b \,|\, \mathcal{D})} \sim \chi^2_{d=2}\,. \tag{3.16}$$

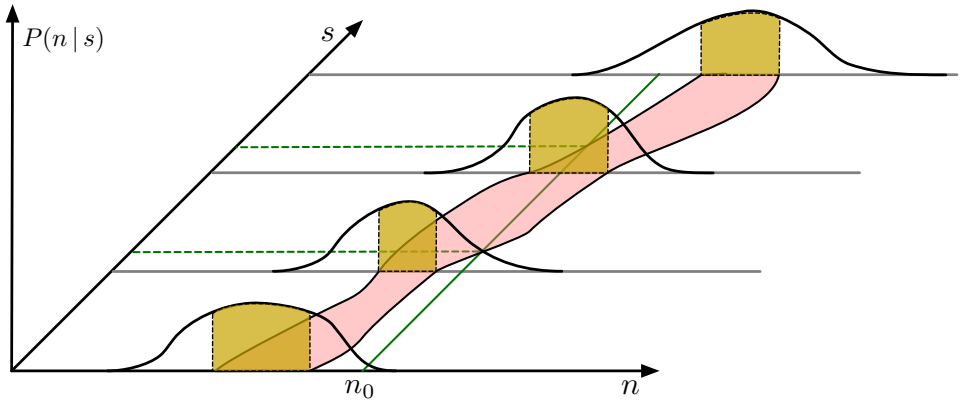

Figure 3.3: **Neyman construction as developed by Feldman and Cousins.** For each value of the signal s one calculates an *acceptence region* from the distribution $P\left(n \mid s\right)$ (for details on computing the acceptence region see Ref. [144]. The acceptence regions are shown in yellow. Next, construct a belt of acceptence regions, as shown in red. Finally, the confidence interval, on s, upon measuring n_0 is given by the intersection of the vertical line through the belt, as shown in green.

Knowing the distribution of the TS drastically reduces the computational cost of calculating confidence intervals. Instead of having to computationally *map* the TS distribution through monte-carlo (MC) simulations, we can instead find threshold values which define the boundary of the confidence interval. The only computation is then to find the values of θ that correspond to these threshold values. Imagine we want to compute a 90% CL interval, the threshold value is given by 2.706. This is calculated from the inverse cumulative distribution function (known as the percentage point function), $\mathrm{PPF}(d, \alpha)$. For the example above we have $\mathrm{PPF}(1, 0.9) = 2.706$, as shown in Fig. 3.4.

If we did not make a discovery, we need to instead set an upper limit. Our TS will no longer be symmetric as in Fig. 3.2 and Wilks theorem does not apply exactly. Instead we need to modify our TS to,

$$\mathrm{TS}(\theta_s) = \begin{cases} -2\ln \frac{\max_{\theta_B} \mathcal{L}(\theta_S, \theta_B \mid \mathcal{D})}{\max_{\theta_S, \theta_B} \mathcal{L}(\theta_S, \theta_B \mid \mathcal{D})} & \theta_s \geq \hat{\theta}_s\,, \\ 0 & \theta_s < \hat{\theta}_s\,, \end{cases} \tag{3.17}$$

where $\hat{\theta}_s$ is the best fit value. This also produces a change to the TS distribution which now goes as,

$$\mathrm{TS} \sim \frac{1}{2}\delta(0) + \frac{1}{2}\chi^2_{d=2}\,, \tag{3.18}$$

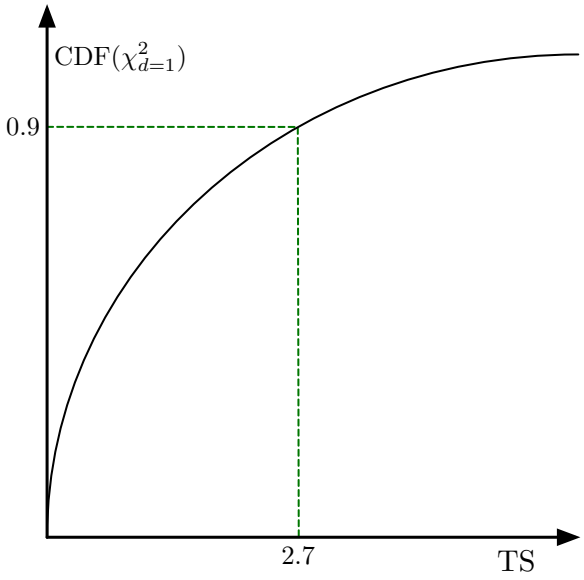

Figure 3.4: **How Wilks theorem is used to calculate threshold values.** Wilks theorem shows that the TS is distribution like a chi-squared distribution. The cumulative distribution function of the χ^2_d distribution can be used to find the necessary threshold values for confidence intervals. Here we show an example for $\chi^2_{d=1}$. The inverse of the cumulative distribution function is called the percentage point function, $\text{PPF}(1, 0.9) = 2.706$.

where d is specified in the same way as above. From here we simply need to solve for the threshold value as we did before but with a modified cumulative distribution function. The exclusion limit is then the value of θ_s at the threshold value, with all values of θ_s above *excluded*.

Proof of Wilks' Theorem

Here we simply sketch the proof of Wilks' theorem and refer to Ref. [146] for details.

1. First we imagine that we observed only the maximum likelihood estimators $\hat{\theta}_j\left(x_1,\ldots,x_N\right)$, allowing us to replace the likelihood of the observed data with a new pdf as so.

$$P\left(x_1,\ldots,x_N|\theta_1,\ldots,\theta_n\right) \longrightarrow \quad P\left(\hat{\theta}_1,\ldots,\hat{\theta}_n|\theta_1,\ldots,\theta_n\right). \tag{3.19}$$

2. If we have enough data to constrain all the parameters sufficiently well, we can assume that the pdf is normally distributed

$$P\left(\hat{\theta}_1,\ldots,\hat{\theta}_n|\theta_1,\ldots,\theta_n\right) \propto e^{-\frac{1}{2}\sum_{ij}\left(\hat{\theta}_i-\theta_i\right)\Sigma^{-1}_{ij}\left(\hat{\theta}_j-\theta_j\right)}, \tag{3.20}$$

3. We can now generically apply the rotation and rescaling,

$$z_k = \sum_{j,i=1}^{n} R_{kj}U_{ji}\theta_i\,, \tag{3.21}$$

such that

$$P\left(\hat{z}_1,\ldots,\hat{z}_n|z_1,\ldots,z_n\right) \propto e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\hat{z}_i-z_i\right)^2}, \tag{3.22}$$

where R_{kj} is a rotation matrix and U_{ji} is used to rescale[a].

4. We now examine the new MLR with the new variables,

$$-2\ln\frac{\max_{z'_{k+1},\ldots,z'_n} P\left(\hat{z}_1,\ldots,\hat{z}_n|z_1,\ldots,z_k,z'_{k+1},\ldots,z'_n\right)}{\max_{z'_1,\ldots,z'_n} P\left(\hat{z}_1,\ldots,\hat{z}_n|z'_1,\ldots,z'_n\right)} = \sum_{i=1}^{k}\left(\hat{z}_i-z_i\right)^2, \tag{3.23}$$

which is a chi-squared distribution with k degrees of freedom.

[a] It is not necessary to know the exact form of there matrices for the proof here, only that they can be applied.

3.3.3 Small Sample Limit

Now let us discuss the most problematic scenario, when the number of background and signal events is low. Here, Wilks theorem breaks down since there is not enough data to constrain the parameters well. Instead one must resort to MC realisations of the TS to map its distribution; an extremely computationally intensive task. To set an exclusion limit the process is as follows:

1. Generate a data set from the model, assuming the background only hypothesis.
2. Find the minimum value of the TS for the alternative hypothesis, as specified in Eq. 3.17. Each realisation of the data will give you a different minimum value.
3. Repeat as many times as is necessary to accurately produce a cumulative distribution function for finding threshold values.
4. Follow steps one and two as above but instead using the real data. Use the threshold values calculated in step three to calculate the exclusion limit on the model parameter.

Due to the computational overhead of this procedure, only a small subset of physics models get tested rigorously. Alleviating this computational burden is the topic of Chapter 4.

3.3.4 Forecasting

Throughout this section we have imagined that data has been taken and we want to assess how well the data constrains physical models. Another common task is to instead examine the ability of future experimental setups to constrain the parameter space of different models. The methods used here are extremely similar to those mentioned above with some small differences. For example, when setting an upper limit, instead of using the experimental data to calculate the limit (as described above) one would simulate many data sets for the background only hypothesis calculating upper limits for each and then report the median as the *expected exclusion limit*. This exact procedure and how to perform these calculations more efficiently will be discussed in Chapter 4.

In addition to calculating the expected sensitivity of an experiment to new physics, one might want to quantify its ability to constrain parameters given a detection. There are two commonly used approaches here, both involve using *benchmark scenarios*. Imagine we have a physics model and assume it has the true parameter θ_s^t, this is known as a benchmark point. Using the description of the signal with this parameter we can perform the same exercise as above but instead simulate data for the signal + background hypothesis. Once we have the distribution of the TS we can construct the *expected two sided confidence interval*.

If we instead want to calculate whether two different parameters, θ_{s1}^t and θ_{s2}^t, will be *distinguishable*, we must calculate two sided confidence intervals for θ_{s1}^t. We can then see if θ_{s2}^t is within the interval for θ_{s1}^t, a measurement of their distinguishability. More concretely, one would construct the TS for a hypothesis test between θ_{s1}^t vs θ_{s2}^t[2].

[2]Note that for almost all cases of interest this is a symmetric to the inverted hypothesis θ_{s2}^t vs θ_{s1}^t.

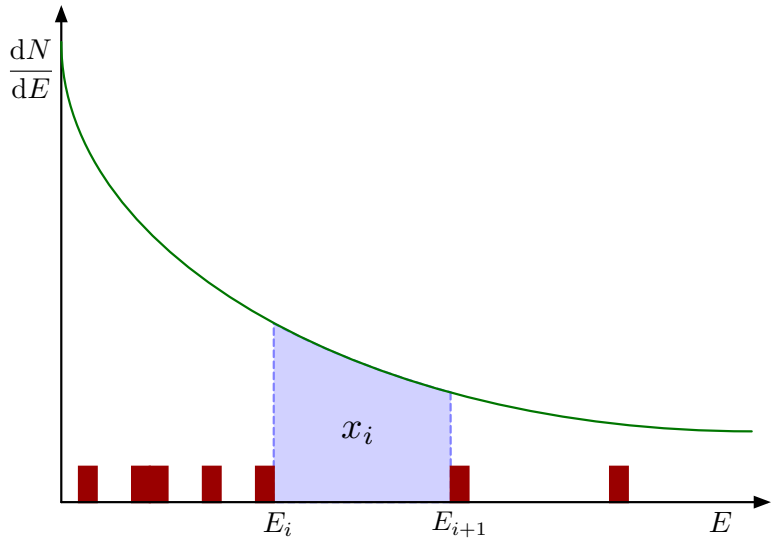

Figure 3.5: **Maximum gap method.** This method was developed by Yellin [147] to be able to set limits even when an unknown background source produces data points. By finding the maximum gap, x_i, and assuming all the data is from a DM signal, one can exclude parameters that are too large to have produced a gap this large.

Again this is a computationally intensive task, even for small numbers of benchmark scenarios. For instance, if one wants to compare n benchmark points, the number of computations is $n \times N$ where N is the number of data samples required to map the TS accurately i.e. the MC. In Chapter 4 we develop an efficient method to perform this procedure, reducing the required computations to just n and essentially allowing for *benchmark free forecasts*.

3.4 Maximum Gap Method

The maximum gap method was developed by Yellin [147] as a way to place an exclusion limit without knowing the background model. This has played a particularly important role in DD DM experiments when some unknown background component contributes to a data set and cannot be subtracted.

Imagine we have N observed events in a data set, with energies $\mathbf{E} = \{E_1, \ldots, E_N\}$. We can treat all these as observed signal events and calculate which parameters would have been too large to have produced this signal. Imagine that we have a model to describe the signal, such as the 'spin-independent' scattering recoil spectrum described in § 2.4, $\mathrm{d}N/\mathrm{d}E$. Between the energy threshold and maximum energy, there exists

$N+1$ of gaps where no events were detected. We now want to calculate the maximum gap size, $x_{\max}$, where the gap size is defined as,

$$x_i = \int_{E_i}^{E_{i+1}} \mathrm{d}E \frac{\mathrm{d}N}{\mathrm{d}E} \,. \tag{3.24}$$

See Fig. 3.5 for a diagram of the constructed maximum gap. We can now compute the probability that the maximum gap is smaller than the observed one as,

$$P\left(x_{\max} < x \,|\, \mu\right) = \sum_{k=0}^{m} \frac{(kx-\mu)^k e^{-kx}}{k!} \left(1 + \frac{k}{\mu - kx}\right) \,, \tag{3.25}$$

where μ is the expectation value of the total number of events and m is the greatest integer $\leq \mu/x$. For a proof of this formula see Ref. [147]. To find the exclusion limit, we set Eq. 3.25 to the desired confidence level and solve for μ. This can then be converted to the relevant model parameter such as the spin independent proton scatting cross section σ_{SI}.

Note that this method is conservative by construction. By considering all the events as signal, the coverage will not match the stated confidence and the limit will be *weaker* as a result. Although this is the lesser of two evils when it comes to over or under coverage, it can leave parameter space *alive* when in fact it has been ruled out. Additional experiments may therefore be required to probe the extra parameter space.

The maximum gap method also has a generalisation to the optimum interval method where one considers gaps with one or more events. This is in contrast to the maximum gap method where we only considered gaps with zero events. See Ref. [147] for more information on the optimum interval method.

3.5 Connection to Bayesian Statistics

Finally, we make a connection to Bayesian methods. This is of particular interest for discovery tests where the commonly used p-value criterion doesn't necessarily give the full picture. In the end, a physicist would like to know how likely a model is to be true given the data set. A small p-value merely quantifies the probability of obtaining a particular data set under the null hypothesis. Instead we can report the likelihood of the null hypothesis to be true. For a one-bin Poisson process this is given by,

$$P\left(H_0 \,|\, N_{\mathrm{obs}}\right) = \frac{P\left(N_{\mathrm{obs}} \,|\, H_0\right) P\left(H_0\right)}{P\left(N_{\mathrm{obs}} \,|\, H_0\right) P\left(H_0\right) + P\left(N_{\mathrm{obs}} \,|\, H_1\right) P\left(H_1\right)} \,, \tag{3.26}$$

where H_0 represents the null hypothesis, H_1 the alternative, and N_{obs} is the observed number of events. From here we can calculate the minimum probability of the null hypothesis over a wide class of priors $P\left(H_0 \,|\, N_{\text{obs}}\right)_{\text{min}}$ which is a direct measure of the probability of the null hypothesis. For a concrete example see Ref. [9].

4 Fisher Forecasting Methods for Astroparticle Physics

This chapter is based on work from Refs. [1, 6]

4.1 Introduction

Progress in astroparticle physics and dark matter (DM) searches is driven by the comparison of theoretical models with experimental data. During this process, estimating the sensitivity of existing or future experiments for the detection of astrophysical or new physics signals is a ubiquitous task, usually requiring the calculation of the expected exclusion and discovery limits. It is done both by phenomenologists who are interested in observational prospects of their theoretical models, as well as experimentalists who are interested in optimizing experimental design. As such, efficient and informative forecasting plays a central role in shaping the development of the field. It is rather common in the (astro-)particle physics community that forecasting is done in the framework of Frequentist statistics. Acceptance of the signal+background hypothesis, H_1, and rejection of the background-only hypothesis, H_0, requires that some test statistic (TS) exceeds a predefined threshold, which depends both on the aspired significance level of the detection as well as the probability distribution function (PDF) of the TS under H_0 [148, 143]. The Frequentist method (in contrast to Bayesian techniques) has the advantage of known false positive rates for hypothesis testing, and known coverage for upper limits. These features can be especially desirable when prior knowledge about the (non-)realization of a theoretical models in Nature is poor.

One of the most commonly used TSs (used for parameter regression, calculation of confidence intervals, and the goodness-of-fit) is the *chi-squared*. Its application is limited to binned data with errors that are approximately normal distributed, which is often realized in the large-sample limit. Asymptotic formulae for its statistical properties are available [17]. In the small sample limit, and generally for more complicated likelihood functions, the maximum likelihood ratio (MLR) [17] is a very common TS.

Again, in the large sample limit, asymptotic distributions are available [146, 149]. In the small-sample limit, however, its statistical properties have to be inferred from Monte Carlo (MC) simulations.

When estimating experimental sensitivities, one general question that arises is '*What is the maximum information that can be in principle extracted from a given observation?*'. Information gain here corresponds to the reduction of the uncertainty associated to the model parameters of interest. The famous Cramér-Rao bound [e.g. 143] provides a general way to derive, for a given experimental setup, a lower bound on the achievable variance of any unbiased model parameter estimator, which holds for *any* statistical method employed to analyse the data. This bound corresponds to an upper limit on the achievable information gain. The Cramér-Rao bound is based on the *Fisher information matrix*, which quantifies how 'sharply peaked' the likelihood function describing the observational data is around its maximum value.

Forecasting instrumental sensitivities with Fisher information is rather common, *e.g.*, in cosmology. It rests on the assumption that estimators which saturate the Cramér-Rao bound are available, and that these estimators approximately follow a multi-variate normal distribution. In the large-sample limit, this is indeed often (but not always) the case. The important limitations of this approach were pointed out many times [*e.g.*, 150, 151], as well as proposals to extend the framework to account for, *e.g.*, non-Gaussian effects [152] (for a recent review see Ref. [153]).

The Fisher information matrix has an impressive range of attractive properties that – if used wisely – can significantly ease the life of anybody interested in performing forecasting, from simple tasks to problems with many experiments, targets and a high-dimensional parameter space. It can be quickly calculated, it is additive, allows for efficient handling of nuisance parameters, and it is at the root of the powerful concept of information geometry [*e.g.*, 154] (and, as we will show in this work, *information flux*). Fisher forecasting is largely unused in the astroparticle physics and DM communities however there are some exceptions, *e.g.* [150, 155, 156, 157, 158, 154, 159].

In this paper, we present an overview of how Fisher information can be used in the context of astroparticle physics and DM searches. Throughout this paper, we focus our attention on Poisson likelihoods and *additive component models*, where the shape (in for instance energy or positional space) of the model components are fixed and the normalization coefficients are the only regression parameters. In addition, we will study the impact of additional *external constraints* on the model parameters. Such constraints can be used to account for various model or instrumental uncertainties. Focusing on these scenarios allows for an in-depth discussion of the specific capabilities and limitations of Fisher forecasting, while still covering many interesting use cases.

We introduce various (to the best of our knowledge) new prescriptions for the efficient estimate of the expected exclusion and discovery limits that are valid in the

small- and large-sample regimes based on the Fisher information matrix. We compare the accuracy of these prescriptions with results from the MLR method, and for a few simple cases with results from the full Neyman belt construction. We demonstrate how to incorporate the effect of correlated background systematics. Lastly, we introduce the new concept of Fisher *information flux*. It generalizes the commonly used signal-to-noise ratio (SNR), while fully accounting for background and other uncertainties. We illustrate the power of this new concept in a few simple examples.

Some of the techniques discussed in this paper are extensively used in other fields of science. Part of the work is inspired by the discussion of information geometry in Ref. [154], and by the notion of effective backgrounds in Refs. [160, 158]. We mention explicitly when and why we deviate from notations introduced in these works if we deem this to be necessary. Some of the methods discussed here have already been applied by some of the present authors in Refs. [156, 157].

The paper is organized as follows: In § 4.2, we introduce and define the Fisher information matrix, additive component models, and equivalent signal and background counts. In § 4.3, we present prescriptions to derive approximate expected exclusion limits and discovery sensitivities from the Fisher information matrix, and study the validity of the results with MC techniques. In § 4.4 we show in a few examples how systematic background uncertainties can be incorporated in the sensitivity estimates. In § 4.5 we introduce the notion of Fisher information flux, and various connected concepts. In § 4.7 we finally conclude.

In Appendix 4.8.1 we discuss relevant properties of the Poisson likelihood function and the associated Fisher information. In Appendix 4.8.2, we discuss conventional methods for the calculation of the discovery reach and expected upper limits, that we use for comparison with our techniques. Finally, in Appendix 4.8.3 we collect some more technical derivations for results in the paper.

4.2 Fisher Information of the Poisson likelihood function

4.2.1 The Fisher information matrix

For any sufficiently regular likelihood function $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$, with n model parameters $\boldsymbol{\theta} \in \mathbb{R}^n$ and data $\mathcal{D}$, the Fisher information matrix is a $n \times n$ matrix that can be defined as

$$\mathcal{I}_{ij}(\boldsymbol{\theta}) \equiv \left\langle \left(\frac{\partial \ln \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})}{\partial \theta_i} \right) \left(\frac{\partial \ln \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})}{\partial \theta_j} \right) \right\rangle_{\mathcal{D}(\boldsymbol{\theta})} = - \left\langle \frac{\partial^2 \ln \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right\rangle_{\mathcal{D}(\boldsymbol{\theta})} , \quad (4.1)$$

where the average is taken over multiple realizations of a model with parameters $\boldsymbol{\theta}$. The second equality holds given some weak regularity conditions [141]. The Fisher information matrix quantifies the *maximum precision* at which the model parameters can be inferred from the data.[1] This follows from the Cramér-Rao bound (CRB) [161, 162], which states that for *any* set of unbiased estimators of the model parameters, $\hat{\boldsymbol{\theta}}$, the inverse of the Fisher information matrix provides a lower limit on its variance. The CRB generalises to the multivariate case to give [148],

$$\mathrm{cov}\left[\hat{\theta}_i, \hat{\theta}_j\right] \equiv \langle(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)\rangle_{\mathcal{D}(\boldsymbol{\theta})} \geq (\mathcal{I}(\hat{\boldsymbol{\theta}})^{-1})_{ij} \ . \tag{4.2}$$

Estimators that saturate the bound exactly are called 'minimum variance'. An estimator is called 'asymptotically efficient' when the bound is saturated in the large sample limit. Note that this bound only holds for *unbiased* estimators, and becomes stronger in the presence of a constant bias [148].

A widely used estimator is the maximum likelihood estimator (MLE). The MLE is *exact*, meaning that although it is in general biased it becomes unbiased in the large sample limit. Furthermore, the MLE is asymptotically efficient. Although the Fisher information matrix as an estimate for the variance of MLEs becomes exact only in the large-sample limit, we will see below that it remains a powerful tool in the small-sample regime.

Searches for new physics in both astrophysics and DM detection experiments are often, at their core, counting experiments. The number of events recorded in a detector is described by the Poisson distribution. We assume here for simplicity that events can be described by one variable, for example photon energy E (the generalization to multiple variables is straightforward and will be used below). Then, the Poisson log-likelihood can be written as (details can be found in Appendix 4.8.1)

$$\ln \mathcal{L}_{\mathrm{pois}}(\mathcal{D}|\boldsymbol{\theta}) = \int dE \ \left[\mathcal{C}(E) \ln \Phi(E|\boldsymbol{\theta}) - \Phi(E|\boldsymbol{\theta})\right] \ , \tag{4.3}$$

where the integration is done over the energy range of interest, and we dropped terms that do not depend on $\boldsymbol{\theta}$ since they do not affect the rest of the discussion. Here, $\Phi(E|\boldsymbol{\theta})$ is the model counts spectrum (with units $1/E$). Furthermore, we defined the 'unbinned' count spectrum

$$\mathcal{C}(E) = \sum_k \delta_D(E - E_k) \ , \tag{4.4}$$

where E_k is the energy of event k, and $\delta_D(E)$ is the Dirac delta function. $\mathcal{C}(E)$

[1] Here, we refer to the maximum precision in terms of a Frequentist approach. In the Bayesian case it is possible to gain further insight with the addition of informative priors and/or a hierarchical model structure.

can be interpreted as a counts histogram with zero bin size. It represents a specific realization of a Poisson process with mean $\Phi(E|\boldsymbol{\theta})$, and has the useful property that $\langle \mathcal{C}(E) \rangle_{\mathcal{D}(\boldsymbol{\theta})} = \Phi(E|\boldsymbol{\theta})$.

We note that the above definition resembles a *Poisson point process*, defined here on an interval on the real line that is given by the energy range of interest [163, 164]. In this context, $\Phi(E|\boldsymbol{\theta})$ would be referred to as 'intensity measure', but we will continue to use here the expressions 'model counts spectrum' or, in some 2-dim examples below, 'model counts map'.

It is straightforward to show that the Fisher information matrix of the above Poisson likelihood is given by

$$\mathcal{I}_{ij}^{\text{pois}}(\boldsymbol{\theta}) = \int dE \, \frac{\partial \Phi(E|\boldsymbol{\theta})}{\partial \theta_i} \frac{1}{\Phi(E|\boldsymbol{\theta})} \frac{\partial \Phi(E|\boldsymbol{\theta})}{\partial \theta_j} \, . \tag{4.5}$$

This form is fully general and holds for any parametric dependence of $\Phi(E|\boldsymbol{\theta})$. The inverse of the Fisher information matrix (irrespective of whether it is derived from Poisson or other likelihood functions) approximates the expected covariance matrix as shown in Eq. 4.2. As mentioned above, this relation holds exactly in the large-sample limit only.

For the diagonal elements of the inverse of the Fisher matrix, we will sometimes use of the notation

$$\sigma_i^2(\boldsymbol{\theta}) \equiv (\mathcal{I}(\boldsymbol{\theta})^{-1})_{ii} \, . \tag{4.6}$$

Furthermore, we note that in general the full Fisher matrix can have, beside the Poisson part that we discussed above, parts that introduce additional constraints on the model parameters (see § 4.4 below). In that case, we can write

$$\mathcal{I}_{ij}(\boldsymbol{\theta}) = \mathcal{I}_{ij}^{\text{pois}}(\boldsymbol{\theta}) + \mathcal{I}_{ij}^{\text{syst}} \, , \tag{4.7}$$

where we assumed that the systematics term is (approximately) independent of the model parameters.

4.2.2 The profiled Fisher information matrix

Typically, one is only interested in a few 'parameters of interest' (PoI), say $\theta_1, \ldots, \theta_k$, while the remaining $n-k$ model parameters are nuisance parameters that parametrize background, instrumental or signal uncertainties. The canonical method to deal with MLEs of nuisance parameters in a Frequentist approach is to maximise the likelihood function $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta})$ with respect to the parameters $\theta_{k+1}, \ldots, \theta_n$, which gives rise to a 'profile likelihood' function that only depends on the k PoI. This method leaves one with a description of the remaining parameter space whilst accounting for the effects

of the $n-k$ nuisance parameters.

To perform the analogous procedure for the Fisher information matrix, we write $\mathcal{I}_{ij}(\boldsymbol{\theta})$ in block form,

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_A & \mathcal{I}_C^T \\ \mathcal{I}_C & \mathcal{I}_B \end{pmatrix} , \tag{4.8}$$

where $\mathcal{I}_A$ is a $k \times k$ matrix that corresponds to the PoI, $\mathcal{I}_B$ is a $(n-k)\times(n-k)$ matrix that corresponds to the nuisance parameters, and $\mathcal{I}_C$ is the mixing between both. We then define the *profiled Fisher information matrix*, where the nuisance parameters are removed in such a way that the (co-)variance of the PoIs remains unchanged. It is given by[2]

$$\widetilde{\mathcal{I}}_A = \mathcal{I}_A - \mathcal{I}_C^T \mathcal{I}_B^{-1} \mathcal{I}_C \ . \tag{4.9}$$

This expression is general and does not depend on the details of the problem.

4.2.3 Additive component models

To simplify the discussion we will, as mentioned above, assume that the model counts spectrum $\Phi(E|\boldsymbol{\theta})$ consists of a number of additive components with free normalization θ_i but fixed shape $\Psi_i(E)$,

$$\Phi(E|\boldsymbol{\theta}) = \sum_{i=1}^{n} \theta_i \Psi_i(E) \ , \tag{4.10}$$

where the sum is taken over the n model components. The model counts spectra can be in many cases calculated as

$$\Psi_i(E) = \mathcal{E}(E)\, I_i(E) \ , \tag{4.11}$$

where $\mathcal{E}(E)$ denotes the exposure (usually effective area or volume times effective observation time) as function of energy, and $I_i(E)$ the differential flux of component i. We furthermore will occasionally use the total flux $I(\boldsymbol{\theta}|\Omega) \equiv \sum_i \theta_i I_i(\Omega)$. The number of expected counts per component is given by

$$\lambda_i(\theta_i) \equiv \theta_i \int dE\, \Psi_i(E) \ . \tag{4.12}$$

For the additive component model, the Fisher information matrix acquires the simple form

$$\mathcal{I}_{ij}^{\text{pois}}(\boldsymbol{\theta}) = \int dE \frac{\Psi_i(E)\Psi_j(E)}{\Phi(E|\boldsymbol{\theta})} \ . \tag{4.13}$$

[2]This result can be readily understood by acknowledging that the upper left $k \times k$ part of the inverse of Eq. 4.8 is given by $(\mathcal{I}_A - \mathcal{I}_C^T \mathcal{I}_B^{-1} \mathcal{I}_C)^{-1}$.

Note that the model parameters enter here only through the model predictions in the denominator.

4.2.4 Equivalent number of signal and background events

In § 4.3 we will present methods for calculating approximate upper limits and discovery reaches which rely on the definitions of the number of signal and background counts. We refer to the method as the Equivalent Counts Method (ECM) since it is a generalisation of the single binned case where it is clear that the number of signal and background events provides information about (a) the expected significance of the signal, (b) the signal-to-background ratio (SBR) and (c) the sample size and hence the relevance of the discreteness or skewness of the Poisson distribution. The ECM definitions capture the same information but for more general cases i.e. situations with large numbers of bins or even the unbinned case as is presented here. In general, not all signal events $\lambda_i(\theta_i)$ are statistically relevant, since some might overlap with regions containing strong backgrounds, whilst other signal events might reside in regions that are almost background free. A reasonable definition of equivalent signal and background events should account for this effect.

We propose definitions here for the number of equivalent signal and background counts that are defined *solely* in terms of the Fisher information matrix. They provide information about the expected signal significance, the SBR and the effective sample size in rather general situations. These definitions are used below for the calculation of expected exclusion and discovery limits.

We define the equivalent signal, $s_i(\boldsymbol{\theta})$, and background, $b_i(\boldsymbol{\theta})$, events for any component i implicitly in terms of the SNR,

$$\mathrm{SNR}_i(\boldsymbol{\theta}) = \frac{s_i^2(\boldsymbol{\theta})}{s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})} \,, \tag{4.14}$$

and in terms of the SBR,

$$\mathrm{SBR}_i(\boldsymbol{\theta}) = \frac{s_i(\boldsymbol{\theta})}{b_i(\boldsymbol{\theta})} \,. \tag{4.15}$$

The SNR for model component i is given by the corresponding diagonal term of the inverse of the Fisher information matrix, times the factor θ_i^2, which we can write as

$$\mathrm{SNR}_i(\boldsymbol{\theta}) \equiv \frac{\theta_i^2}{\sigma_i^2(\boldsymbol{\theta})} \,. \tag{4.16}$$

If systematic (non-Poisson) contributions to the $\mathcal{I}_{ii}(\boldsymbol{\theta})$ as well as mixing with other components are negligible, $\mathrm{SNR}_i(\boldsymbol{\theta})$ is essentially given by the corresponding diagonal component of Eq. 4.13. As expected, this is simply the standard SNR of $\theta_i \Psi_i(E)$

over a fixed background $\Phi(E|\boldsymbol{\theta})$. The definition of the SBR in terms of the Fisher information matrix is less obvious. We find that the following expression serves the purpose,

$$\mathrm{SBR}_i(\boldsymbol{\theta}) \equiv \frac{\sigma_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta}_0)} - 1 \,. \tag{4.17}$$

Here, we used the definition $\boldsymbol{\theta}_0 \equiv (\theta_1, \ldots, \theta_{i-1}, 0, \theta_{i+1}, \ldots, \theta_n)^T$, *i.e.* it is $\boldsymbol{\theta}$ with the normalization of the signal component set to zero. This choice will be further justified below.

From the definitions in Eqs. 4.16 and 4.17, the equivalent number of signal and background events can be directly obtained using Eqs. 4.14 and 4.15. They are given by

$$s_i(\boldsymbol{\theta}) = \frac{\theta_i^2}{\sigma_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta}_0)} \,, \tag{4.18}$$

and

$$b_i(\boldsymbol{\theta}) = \frac{\theta_i^2 \sigma_i^2(\boldsymbol{\theta}_0)}{(\sigma_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta}_0))^2} \,. \tag{4.19}$$

Discussion. Although these definitions are relatively abstract, they have a number of useful properties. These become particularly clear when mixing between the components is negligible, *i.e.* $\sigma_i^2(\boldsymbol{\theta}) \simeq 1/\mathcal{I}_{ii}^{\mathrm{pois}}(\boldsymbol{\theta})$. If component shapes are very similar, such a situation can be enforced by introducing constraint terms for all non-signal components, in the way that we will discuss below in § 4.4.1. If *component mixing is negligible*, we find:

(i) If all $\Psi_i(E)$ are constant, which corresponds to a single-bin scenario, the number of equivalent signal events just equals the total number of signal events, $s_i(\boldsymbol{\theta}) = \lambda_i(\boldsymbol{\theta})$, and the equivalent background events equals the sum of the events from all other components, $b_i(\boldsymbol{\theta}) = \sum_{j \neq i} \lambda_j(\boldsymbol{\theta})$.

(ii) In the large-signal limit, here defined as $\theta_i \Psi_i(E) \gg \sum_{j \neq i} \theta_j \Psi_j(E)$, the number of equivalent signal events of component i equals the total number of events in the signal component, $s_i(\boldsymbol{\theta}) \simeq \lambda_i(\theta_i)$. In other cases, the equivalent number of signal events is in general smaller, $s_i(\boldsymbol{\theta}) \leq \lambda_i(\theta_i)$.

(iii) The definition of the SBR for model component i in Eq. 4.17 becomes clearer when writing it in terms of the additive components functions $\Psi_i(E)$ and their sum $\Phi(E)$. In the small-signal regime, $\theta_i \Psi_i(E) \ll \sum_{j \neq i} \theta_j \Psi_j(E)$, this yields

$$\mathrm{SBR}_i(\boldsymbol{\theta}) \simeq \left(\int dE \, \frac{\Psi_i(E)^2}{\Phi(E|\boldsymbol{\theta}_0)} \right)^{-1} \int dE \, \frac{\Psi_i(E)^2}{\Phi(E|\boldsymbol{\theta}_0)} \, \frac{\theta_i \Psi_i(E)}{\Phi(E|\boldsymbol{\theta}_0)} \,. \tag{4.20}$$

Here, the ratio $\theta_i \Psi_i(E)/\Phi(E)$ can be interpreted as the SBR of component i as

function of E, and $\Psi_i^2(E)/\Phi(E)$ is proportional to the SNR as function of E. Thus, Eq. 4.20 effectively averages the SBR over regions of E where the signal component $\Psi_i(E)$ is most significant w.r.t. the background $\Phi(E)$. In other words, regions in E where the background $\Phi(E)$ is intense enough that the signal $\Psi_i(E)$ is swamped do not contribute to Eq. 4.20, even if the number of signal photons in that region is high.

(iv) On the other hand, in the large-signal limit (as defined above), the SBR of component i becomes

$$\mathrm{SBR}_i(\boldsymbol{\theta}) \simeq \left(\int dE\, \Psi_i(E) \right)^{-1} \int dE\, \Psi_i(E)\, \frac{\theta_i \Psi_i(E)}{\Phi(E|\boldsymbol{\theta}_0)} \,. \tag{4.21}$$

The weighting by the significance of the signal is now replaced by a weighting over the signal strength.[3]

In the *general case with mixing* between the components, the above simple analytic expressions do not hold anymore. If we concentrate on Fisher information matrices of the from Eq. 4.7, where the non-Poisson part is independent of the model parameters, one can however still show that $\mathrm{SBR}_i(\boldsymbol{\theta}) \geq 0$, which implies the important property that $b_i(\boldsymbol{\theta}) \geq 0$ and $s_i(\boldsymbol{\theta}) \geq 0$ in all cases. We confirmed this numerically by calculating SBR_i for a large number of randomly generated models.

Illustration. Finally, we illustrate the definitions introduced in this subsection with a simple example. Consider the following model

$$\Phi(E) = \theta_1 \left(\frac{\alpha}{1+\alpha} N_1(E) + \frac{1}{1+\alpha} N_3(E) \right) + \theta_2 e^{-E} \,, \tag{4.22}$$

where α is a shape parameter of the signal component, $N_1(E)$ and $N_3(E)$ are normal distributions with variance 0.5 that are centered on the indexed values, and we set the signal normalization to $\theta_1 = 1$ and the background normalization to $\theta_2 = 8$. Example spectra are shown in the upper panel of Fig. 4.1, for various values of α. We neglect mixing between the components by assuming that the background is fixed via external constraints.

In the lower panel of Fig. 4.1 we show the corresponding equivalent signal and background counts of component $i = 1$, as well as the corresponding SBR s_1/b_1. For $\alpha \ll 1$, the signal is dominated by the high energy peak, which is in a region of low background. Indeed, we find approximately $b_1 \simeq 0.33$ and $s_1 \simeq 0.89$. On the other

[3] Eq. 4.21 is underlying the definition of the 'effective background', $b_{\mathrm{eff}} = n_s^2/\mathrm{TS}$, used in Refs. [160, 158], if we identify $n_s = \lambda_i$ and $\mathrm{TS} = \mathrm{SNR}_i$ (see Eq. 4.16). We find that this definition can significantly underestimate the SBR in cases where most of the signal events are located in regions where they are statistically not significant. This can happen for instance in low-energy tails of a steeply falling astrophysical spectra. Our definitions based on Eq. 4.17 do not exhibit this problem.

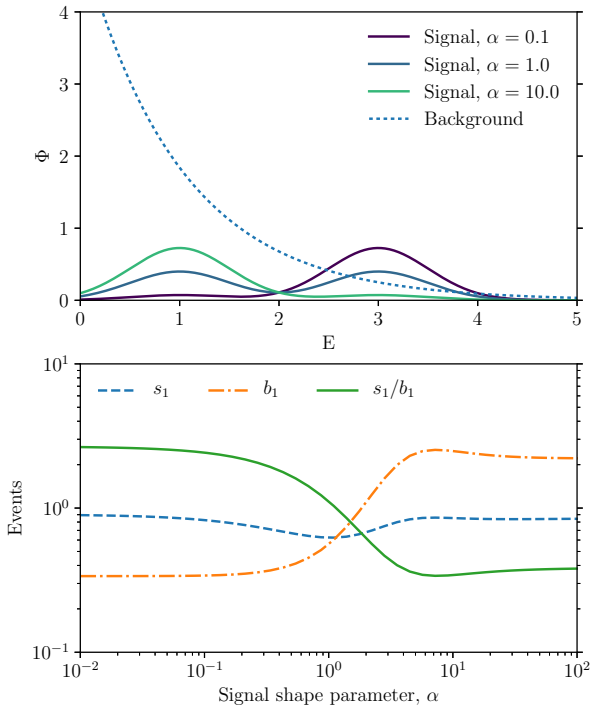

Figure 4.1: **Example signal and background (upper) and corresponding equivalent signal and background (lower).** *Upper panel:* Example spectra of Eq. 4.22 for various values of α. *Lower panel:* Illustration of equivalent signal counts s_1, background counts b_1, and SBR s_1/b_1 as function of the signal shape parameter α.

hand, for $\alpha \gg 1$, the signal is present at lower energies, where the background is larger. Indeed, we find here $b_1 \simeq 2.2$.

In the transition region, the equivalent number of signal events s_1, which otherwise remains close to one, drops somewhat. This is expected, since at $\alpha \sim 1$ half of the signal is in a region of large backgrounds, whereas the other half is in a region with low backgrounds. Since the low-background component dominates the signal-to-noise ratio in that case, it is also mostly this component that contributes to the equivalent signal number counts. The effect becomes more pronounced if the background below $E < 2$ is further increased, or above $E > 2$ further reduced.

4.3 Expected exclusion limits and discovery reach

In this section, we introduce prescriptions for deriving expected exclusion and discovery limits from the Fisher information matrix. These prescriptions build on the equivalent signal and background counts that we defined above. We refer to the corresponding prescriptions as the Equivalent Counts Method (ECM). For simplicity, we take the model component of interest to be $i = 1$. We put particular emphasis on the case of very small or a vanishing number of background events, and validate the approach for a few examples using MC simulations. Caveats are discussed in the last subsection.

4.3.1 Expected exclusion limits

Equivalent counts method. Projected approximate upper limits on the parameter θ_1, assuming that the true value is $\theta_1 = 0$, can be derived from the Fisher information matrix by solving the equation

$$s_1(\boldsymbol{\theta}^U) = Z(\alpha) \cdot \sqrt{s_1(\boldsymbol{\theta}^U) + b_1(\boldsymbol{\theta}^U)} \,, \tag{4.23}$$

for θ_1^U, while keeping the remaining $n-1$ parameters fixed to their respective values. Here, we defined $\boldsymbol{\theta}^U = (\theta_1^U, \theta_2, \ldots, \theta_n)^T$. Furthermore, s_i and b_i refer to the equivalent signal and background counts that we introduced above in § 4.2.4. Finally, $Z(\alpha)$ is connected to the desired confidence level of the limit, $100(1-\alpha)\%$ CL, and α is the significance level. It is derived from the inverse of the standard normal cumulative distribution function, denoted $F_{\mathcal{N}}$, as

$$Z(\alpha) = F_{\mathcal{N}}^{-1}(1-\alpha) \,. \tag{4.24}$$

In the case of, say, a 95% CL upper limits, we have $\alpha = 0.05$ and hence $Z = 1.64$ [17]. It is convenient to rewrite Eq. 4.23 as

$$\theta_1^U = Z(\alpha) \cdot \sigma_1 \left(\boldsymbol{\theta}^U\right) \,, \tag{4.25}$$

which is in practice much easier to evaluate than Eq. 4.23.

In the *background-limited regime*, $\theta_1^U \Psi_1(E) \ll \Phi(E|\boldsymbol{\theta})$, Eq. 4.25 implies that the exclusion limit satisfies

$$\theta_1^U \simeq Z(\alpha) \cdot \sigma_1(\boldsymbol{\theta}) \,, \tag{4.26}$$

where $\boldsymbol{\theta}_0 = (0, \theta_2, \ldots, \theta_n)^T$. This is exactly what is expected if Gaussian background noise with variance σ_1^2 dominates. In the *signal-limited regime*, here defined

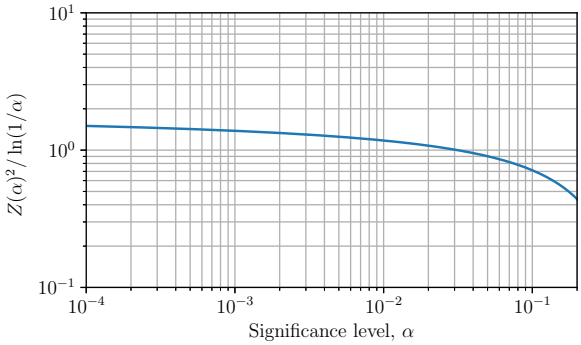

Figure 4.2: **Ratio between expected exclusion limits from the ECM and the exact Neyman belt construction.** We assume a single bin with vanishing expected background, as function of the significance level of the limit α. See Eq. 4.27 and text for details.

as $\theta_1^U \Psi_1(E) \gg \sum_{i\geq 2} \theta_i \Psi_i(E)$, Eq. 4.25 implies on the other hand

$$\lambda_1(\theta_1^U) \simeq Z(\alpha)^2 \,. \tag{4.27}$$

The upper limit is here independent of the signal model spectrum, $\Psi_1(E)$. Eq. 4.27 should be compared with $\lambda_1 = \ln 1/\alpha$, which is the proper upper limit on the mean of a Poisson process when zero events are observed and the expected background is negligibly small. We show in Fig. 4.2 that the fractional difference between $Z^2(\alpha)$ and $\ln(1/\alpha)$ is indeed small for typical significance level values used to set upper limits in the literature. Namely, for $\alpha = 10^{-3}$–10^{-1} the deviation is smaller than 38%.

We note that the general method to estimate expected upper limits in Eq. 4.23 works very well even in the presence of parameter mixing and background systematics as discussed in § 4.4, as long as the associated changes of the background flux remain 'sufficiently small' (say, below a few tens of percent) in the signal region. A quantitative discussion can be found in subsection 4.3.3 below.

Comparison with exact methods. The most general exact method for deriving upper limits with the correct coverage, and actually any sort of confidence intervals, is based on the Neyman Belt construction [165] (an instructive overview can be found in Ref. [144]). In practice, often the more specific MLR method is used to construct confidence intervals of any sort. In the small-sample limit, MC simulations are required to establish the statistics of the MLR and construct intervals with the desired coverage (this is what we do here in all cases). Details about the construction of confidence intervals using both methods can be found in Appendix 4.8.2.

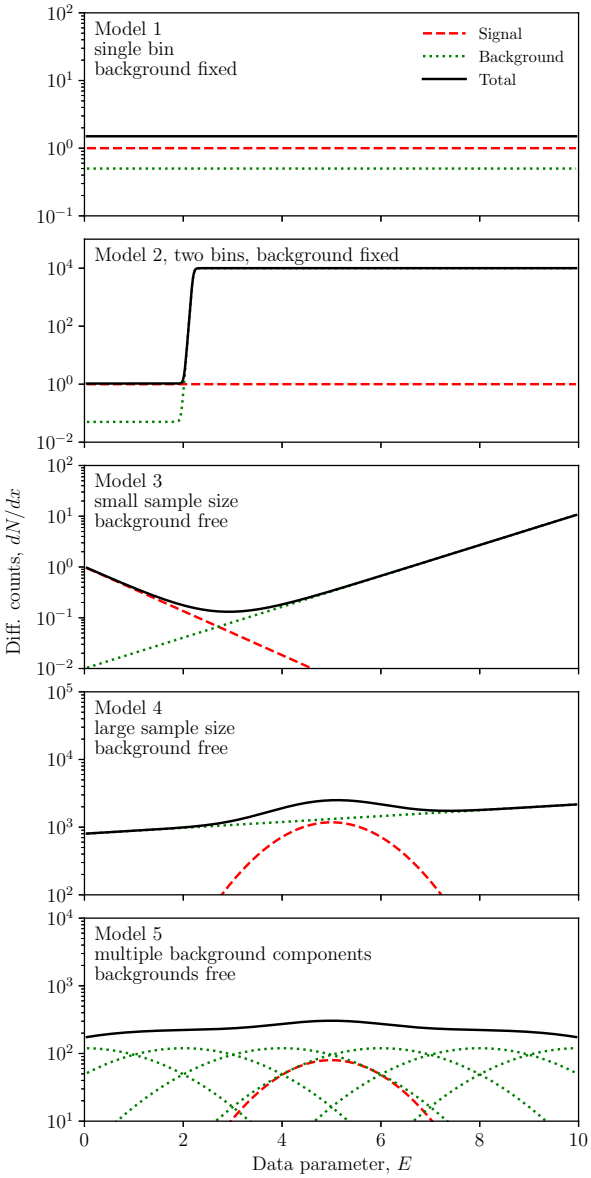

Figure 4.3: **Five example models that we use to test the EC method against the full MLR method.** The signal component refers to $\Psi_1(E)$, the background component to $\Psi_2(E)$. We also indicate whether background components are treated as fixed or free in the analysis. Model 5 has six background components, $\Psi_{2-7}(E)$, instead of one.

As specific examples, we consider the signal and background functions shown in Fig. 4.3. Model 1 is a simple single-bin example, whereas model 2 is a basic two-bin example with a strong difference in the expected background counts in both bins. In both cases, we assume that the background normalization is known and fixed. Model 3 provides an example in the small-sample regime, and model 4 in the large-sample regime. The background normalization is here assumed to be determined by a fit to the data and hence free. Model 5 is a scenario with multiple background components that are to some degree (but not completely) degenerate with the signal.

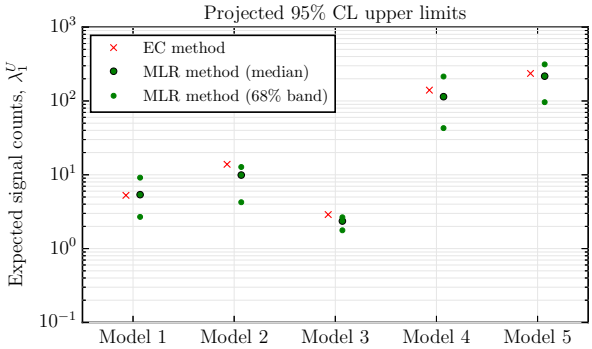

Figure 4.4: **Expected exclusion limits for the five models in Fig. 4.3, derived using the ECM, Eq. 4.25, compared to the median limit from the full coverage-corrected MLR method.** For comparison, we also show the 68% containment regions of the MLR limits. These regions contain, for multiple realizations of background-only data, 68% of the corresponding upper limits.

In Fig. 4.4, we show the expected 95% CL exclusion limits that we obtain from our EC method, Eq. 4.25, for the five example models. We compare these limits with the median limits obtained from the MLR method. In three of the five cases (model 1, 4 & 5) the agreement is remarkably good. To emphasize this, we also indicate the extent of the 68% containment band of the MLR limits, which contains 68% of the upper limits when multiple realizations of the data are considered. They are significantly wider than the difference between the limits from the EC and the MLR methods.

It is instructive to discuss the single-bin model 1 in more detail. In Fig. 4.5, we show the projected 95% CL upper limits as derived from (1) the EC method, (2) the MLR method and (3) the full Neyman belt construction. The upper limits are shown as a function of the number of equivalent background counts. In the large-sample limit, all methods give consistent results. In the limit of vanishing background counts, the EC method yields slightly stronger projected limits than the Neyman belt construction (consistent with Fig. 4.2 and the above discussion). In the intermediate

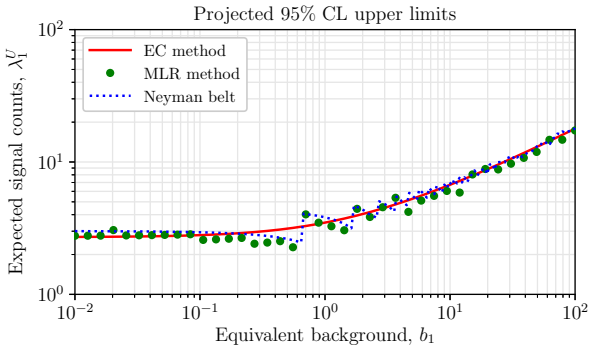

Figure 4.5: **Various projected upper limits for model 1 from Fig. 4.3, as a function of the total equivalent background counts b_1.** Limits from the EC method (red solid line). The median limit based on a full coverage-corrected MLR method (green dots). The median limit from the Neyman Belt construction (blue dotted line).

regime, the Neyman belt construction shows a step-like structure, which is related to the discreteness of the Poisson likelihood. This is not visible in the EC results, but is a small effect almost everywhere. The MLR method yields results consistent with the Neyman belt construction. Deviations are due to MC noise.

Finally, for models 2 and 3, we observe relevant differences between the methods. The EC limits are here weaker ('more conservative') than the limits from the MLR method by up to $\sim 40\%$. We found that this is a rather general behaviour in the Poisson regime, which we observed for a large range of non-trivial scenarios. However, as discussed above, these large deviations are not observed in the single-bin case model 1.

To shed further light on the difference between the single-bin and multi-bin scenarios in the Poisson regime, we show in Fig. 4.6 for model 2 the expected exclusion limits derived from the EC and the MLR methods, as function of the normalization of the background component, θ_2. Three regimes can be clearly discriminated. For $\theta_2 \ll 10^{-5}$, the number of expected background counts over the entire range of x is negligible. This case is equivalent to a simple single-bin scenario, where the EC and MLR methods agree well (see also Fig. 4.5 for model 1). On the other hand, for values of $\theta_2 \gg 10^2$, we enter the Gaussian regime where the number of background counts is large over the entire range of x. Again, EC and MLR results agree well.

However, in the intermediate range, $\theta_2 \sim 10^{-3}$–10^{-1}, we find a plateau where the MLR and EC methods yield different results. In the plateau region, the number of background events is large at $2 < x < 10$, but negligible at $0 < x < 2$. This effectively reduces the number of statistically relevant signal events by a factor of

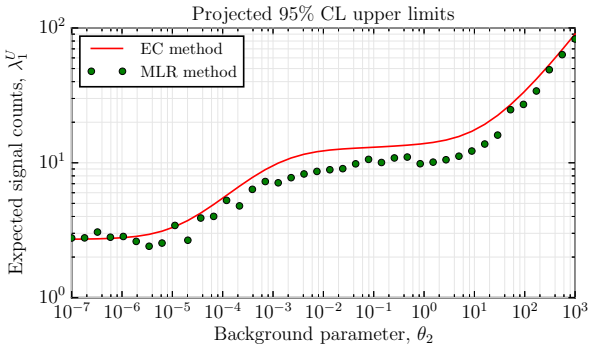

Figure 4.6: **Projected upper limits for model 2 from Fig. 4.3, using the EC method and the coverage-corrected MLR method.** See text for a detailed discussion.

five, and should consequently weaken the expected exclusion limits by the same factor w.r.t. the $\theta_2 \ll 10^{-5}$ case. This happens indeed for the limits derived using the EC method, but not for the limits from the MLR method. We trace this behaviour back to the fact that in the plateau region, the events from $2 < x < 10$ introduce a noise in the otherwise Poissonian likelihood from $0 < x < 2$. This effectively removes the discreetness of the Poisson distribution. As a consequence, the MLR limits in the plateau region can cover exactly, while the MLR limits in the $\theta_2 \ll 10^{-5}$ actually over-cover.

Scenarios like model 3 have a low- and a high-background regime, and hence behave similar to model 2 in the plateau region, leading to discrepancies between the MLR and the EC results (see Fig. 4.4). We never found the effect to exceed 40%.

4.3.2 Expected discovery reach

Equivalent counts method. The discovery limit for θ_1, *i.e.* the value of θ_1 that leads in 50% of the cases to a rejection of the $\theta_1 = 0$ hypothesis with a significance level α, can be approximately obtained by numerically solving the following equation for θ_1^D:

$$\Big(s_1(\boldsymbol{\theta}^D) + b_1(\boldsymbol{\theta}^D)\Big) \ln\left(\frac{s_1(\boldsymbol{\theta}^D) + b_1(\boldsymbol{\theta}^D)}{b_1(\boldsymbol{\theta}^D)}\right) - s_1(\boldsymbol{\theta}^D) = \frac{Z(\alpha)^2}{2} \,. \tag{4.28}$$

Here, we use the notation $\boldsymbol{\theta}^D = (\theta_1^D, \theta_2, \ldots, \theta_n)^T$, and s_1 and b_1 refer to the equivalent signal and background counts discussed at the end of § 4.2.4. Heuristically, Eq. 4.28 is motivated by the analytic structure of profile likelihood ratios, see for instance

discussion in Ref. [17]. However, its main motivation comes from the fact that it leads to approximately correct results both in the signal- and background- limited regimes.

It is useful to consider limiting cases. In the *background-limited case*, $b_1(\boldsymbol{\theta}^D) \gg s_1(\boldsymbol{\theta}^D)$, Eq. 4.28 implies that

$$\theta_1^D \simeq Z(\alpha) \cdot \sigma_1(\boldsymbol{\theta}) \,, \tag{4.29}$$

with $\boldsymbol{\theta}_0 = (0, \theta_2, \ldots, \theta_n)^T$. This is exactly what we expect for Gaussian background noise with variance σ_1^2. On the other hand, in the *signal-limited case*, $b_1(\boldsymbol{\theta}^D) \ll s_1(\boldsymbol{\theta}^D)$, one can show that the solution to Eq. 4.28 satisfies the following equation (details can be found in Appendix 4.8.3):

$$\frac{b_1(\boldsymbol{\theta}^D)^{s_1(\boldsymbol{\theta}^D)}}{\Gamma\left(s_1(\boldsymbol{\theta}^D)+1\right)} = \alpha \cdot \sqrt{\frac{Z^2(\alpha)}{s_1(\boldsymbol{\theta}^D)}} \,. \tag{4.30}$$

Here, $\Gamma(\cdot)$ is the gamma function. This equation is very similar to the exact expression derived from the Poisson distribution using Asimov data [17] in the low-background limit. To see this, note that, given some background $b_1 \ll 1$ and zero signal, the probability to detect λ or more photons is approximately $(b_1)^\lambda/\Gamma(\lambda+1)$. If λ were the discovery reach corresponding to significance level α, this expression should equal α. The difference between this exact and the above approximate expressions is hence the square-root on the right-hand side of Eq. 4.30. In practice, this turns out to be a small effect, as we will see below.

Comparison with exact methods. In Fig. 4.7 we compare the expected discovery limits derived using the EC method, Eq. 4.28, with the ones from the full MLR method, for the five example models in Fig. 4.3 (details can be found in Appendix 4.8.2). We find that for all cases the EC results remain very close to the MLR results. Deviations are usually much less than 1σ, and largest for model 3 where the equivalent background is lowest. This good agreement is quite remarkable, given that some of the models are deeply in the Poisson regime.

To further investigate the limitations of our EC method in the low-background regime, we compare in Fig. 4.8 for model 1 the 3σ and 5σ expected discovery limits derived from the EC and the MLR methods. The approximate discovery limits closely resemble the exact results down to equivalent backgrounds of one. For smaller equivalent backgrounds, the discreteness of the Poisson distribution starts to dominate the MLR method, which is not seen in the EC results. However, depending on the significance level, the agreement remains reasonably good even down to and below $b_1 \sim 10^{-3}$. We note that, in very extreme cases, the EC method could lead to

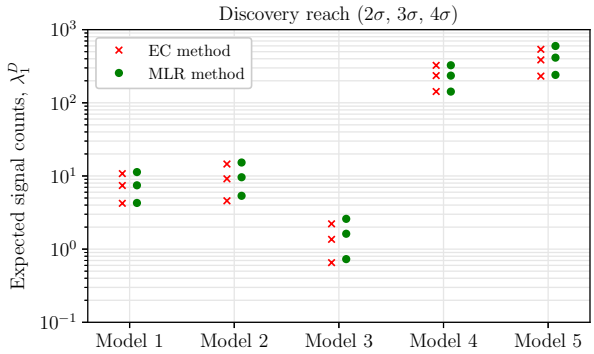

Figure 4.7: **Expected discovery limits derived from the EC method, Eq. 4.28, compared wit full MLR results, for the five models in Fig. 4.3**. Groups of three symbols show from bottom to top 2σ, 3σ and 4σ discovery limits (meaning that 50% of the measurements would lead to a detection with at least the indicated significance). Both methods yield consistent results, see detailed discussion in the text.

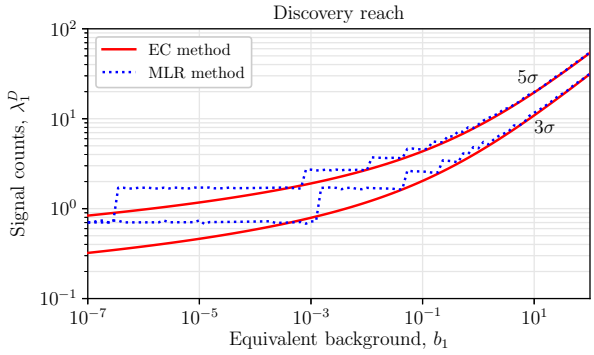

Figure 4.8: **Comparison of 3σ (bottom) and 5σ (top) expected discovery limits based on the EC method, Eq. 4.28, and on the MLR method**. The discreteness of the Poisson distribution becomes apparent for $b_1 \lesssim 1$, where the pure EC methods over-predicts the sensitivity by the indicated amount. See text for a detailed discussion.

expected discovery limits that are much smaller than one. This can be prevented by the additional ad hoc requirement that the equivalent number of signal counts, $s_1(\boldsymbol{\theta})$, should be at least one.

4.3.3 Limitations of the Fisher approach

There are two important situations where the prescriptions for deriving expected exclusion limits, Eq. 4.25, and expected discovery limits, Eq. 4.28, will break down and give potentially wrong results.[4] Here we address both in detail.

Skewness of Poisson likelihood. The Fisher information matrix as defined in Eq. 4.1 encodes complete knowledge about the likelihood function, provided that the likelihood behaves like a multivariate normal distribution. This is equivalent to requiring that higher order derivatives in the expansion of the log-likelihood are negligible. Often, higher derivative terms will only cause deformations to Gaussian contours, but in the extreme case in which parameters are infinitely degenerate, the Fisher formalism will provide not only quantitatively but also qualitatively wrong results [150]. Non-Gaussianity is a common challenge in cosmology [166], but also relevant for the simple additive component model with the Poisson likelihood considered here.

As we saw above, using our EC method, one can obtain reasonable expected exclusion and discovery limits even in the deep Poisson regime, despite the fact that the likelihood functions are clearly non-Gaussian in that case. This is partially due to some lucky numerical coincides that happen to make our proposed prescriptions reasonably accurate. However, these mechanisms only apply to the signal-component of interest. The likelihood functions corresponding to the background (any non-signal) components should obey the usual Gaussianity constraints to ensure that the EC method can be applied.

In Appendix 4.8.1, we study the behaviour of the Poisson likelihood function up to third order in the model parameters. Assuming that no strong degeneracies exist between parameters, one can derive a simple condition on the equivalent number of background events that should hold for all background components. It reads

$$b_i \gtrsim \frac{4t}{9f^2} \,, \tag{4.31}$$

where b_i is the number of *equivalent background* counts of component i, t the value of $\Delta(2\ln\mathcal{L})$ at the boundary of the confidence region of interest, and f the tolerable fractional uncertainty of t at that boundary. If, for instance, we want to have 2σ intervals (for one dimension this corresponds to $t = 4$) with a fractional significance error of less than 20% (this corresponds to $f \simeq 40\%$, since t depends quadratically on the significance in standard deviations), this implies that the equivalent background for the components i should exceed $b_i \gtrsim 11$. In cases where parameter degeneracies might play a role, we recommend to use the full expressions provided in Appendix 4.8.1

[4] This is the case for the additive component models considered in the present work. For more general Poisson problems, *e.g.* with shape uncertainties, the number of potential problems is larger.

instead.

Parameter degeneracies. Using the Fisher formalism implicitly assumes that model parameters are unbound and only constrained by the likelihood function. This means that there is nothing that prevents some of the parameters θ_i to become negative. Consider, as a simple example, a signal, $\Psi_i(E)$, that is exactly degenerate with a background component $k \neq i$, $\Psi_k(E) \propto \Psi_i(E)$. In that case, the Poisson part of the Fisher information matrix becomes singular, and the error of the signal component, σ_i, diverges. An arbitrarily large signal i can be compensated by an equally large negative background component k. Since this behaviour is usually unphysical (unless for instance absorption effects are part of the model), it must be prevented when performing Fisher forecasting.

A sufficient condition to exclude the problems with negative components reads

$$s \cdot \sigma_k(\boldsymbol{\theta}) < \theta_k \quad \text{for all} \quad k \neq i \,, \tag{4.32}$$

where s is the significance of interest in standard deviations. If errors are sufficiently small, the parameter θ_k will not cross zero. As will be discussed below in section 4.4, if the projected data alone is not sufficient to break the degeneracy between different background components or the background and the signal, it is possible (and necessary) to include additional constraints on parameters such that Eq. 4.32 is fulfilled. If this is not possible, the EC method cannot be directly applied.

4.4 Modeling of instrumental and background systematics

We give here a few instructive examples of how to model background uncertainties within the Fisher information framework.

4.4.1 Basic parameter systematics

In many cases of practical importance, additional information about nuisance parameters is available, which must be included in the sensitivity projections to obtain realistic results. Within a Bayesian approach, these additional constraints would be included as priors on the nuisance parameters. Within the Frequentist treatment, which we focus on here, a common approach is to include additional parameter constraints as effective likelihoods, as described in the following. These additional constraints can, for instance, come from 'sideband measurements' in signal-free regions

of the data space.

If we assume that the constraints on parameter θ_i are Gaussian, with standard deviation ξ_i, the associated combined likelihood function can be written as

$$\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}) = \mathcal{L}(\mathcal{D}|\boldsymbol{\theta})_{\text{pois}} \times \prod_i \mathcal{N}(\theta_i^A|\mu = \theta_i, \sigma^2 = \xi_i^2) \; . \tag{4.33}$$

Here, $\mathcal{N}$ refers to the PDF of a normal distribution with variance ξ_i^2 and Asimov value θ_i^A. It accounts for potential sideband measurements and similar external constraints. In this spirit, θ_i^A is taken to equal the mean value, $\theta_i^A = \theta_i$, when the average $\langle\cdot\rangle_{\mathcal{D}(\boldsymbol{\theta})}$ is applied in Eq. 4.1.

The resulting total Fisher information matrix can be split in a Poisson and a systematics part, $\mathcal{I}_{ij} = \mathcal{I}_{ij}^{\text{pois}} + \mathcal{I}_{ij}^{\text{syst}}$, where the latter is here diagonal and given by

$$\mathcal{I}_{ij}^{\text{syst}} = \delta_{ij} \frac{1}{\xi_i^2} \; . \tag{4.34}$$

The generalization to correlated systematic errors reads $\mathcal{I}_{ij}^{\text{syst}} = (\Sigma_{\text{syst}}^{-1})_{ij}$. We will show in a few examples how this is used in practice.

4.4.2 Example 1: Background systematics degenerate with the signal

We start with a simple example where we assume that some component of the background systematic is perfectly degenerate with the signal. More specifically, we consider a three-component model, where $\Psi_1(E)$ is the signal, $\Psi_2(E)$ the nominal background, and the component $\Psi_3(E) = \Psi_1(E)$ accounts for small positive or negative perturbations of this background. Note that this implies that $\mathcal{I}_{1i}^{\text{pois}} = \mathcal{I}_{3i}^{\text{pois}}$ for $i = 1, 2, 3$, which means that the Poisson part of the Fisher matrix is singular. We set the background normalization to $\theta_2 = 1$, and the mean background perturbation to zero, $\theta_3 = 0$. For the systematics part of the Fisher matrix, Eq. 4.34, we assume that the background perturbation θ_3 is externally constrained with a variance of $\xi_3^2 > 0$, and the background normalization θ_2 and the signal normalization are unconstrained, $\xi_1^2, \xi_2^2 \to \infty$.

In that case, one can show, by calculating the profiled Fisher information for component $i = 1$, that the variance of the signal component is given by (details can be found in Appendix 4.8.4)

$$\sigma_1^2(\boldsymbol{\theta}) = (\sigma_1^{\text{pois}})^2(\boldsymbol{\theta}) + \xi_3^2 \; . \tag{4.35}$$

This means that, as one might have expected, statistical and systematic errors are

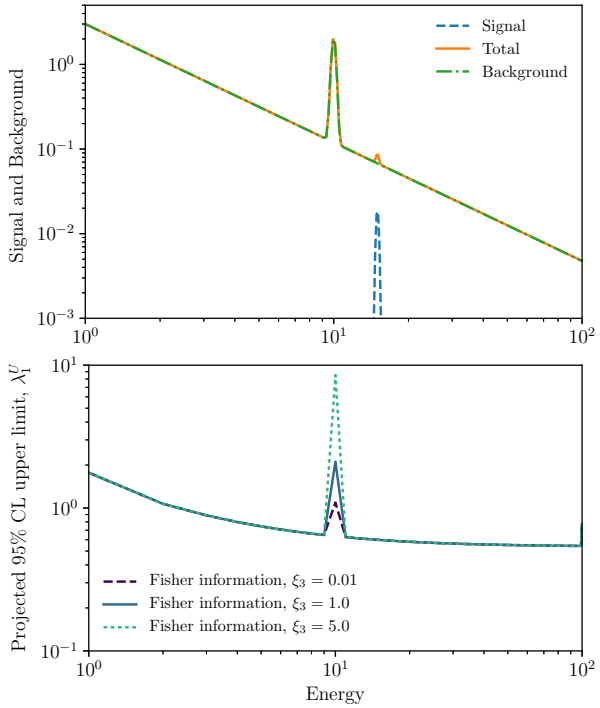

Figure 4.9: **Effect of background systematics on upper limits.** *Upper panel:* Example of a fixed power law background with one instrumental lines at $E = 10$. We show an example of a signal line with identical width to the instrumental lines (here at $E = 15$). *Lower panel:* We calculate expected exclusion limit using the EC method, as a function of the signal position. The limits becomes more constraining at higher energies, due to the increased SBR. However, there is loss of constraining power when the signal is degenerate with the instrumental line, depending on the uncertainty of the instrumental line, ξ_3.

added in quadrature. We remark that constraints on the signal parameter have the opposite effect, $\sigma_1^{-2}(\boldsymbol{\theta}) = (\sigma_1^{\rm pois})^{-2}(\boldsymbol{\theta}) + \xi_1^{-2}$, and *decrease* the overall variance.

To illustrate Eq. 4.35, we consider expected exclusion limits on line-like signals on top of a power-law background plus an 'instrumental line'. We treat the power-law background as fixed, and assume that the instrumental line has identical width to the signal line. However, the normalization of the instrumental line is constrained by $\xi_3 = 1.0$. The fluxes of the background and signal components are illustrated in the upper panel of Fig. 4.9. Specifically we consider a Gaussian located at $E = 10$, with width $\sigma_E^2 = 0.2$. The power law has a slope of 1.4 and normalisation of three, $\theta_2 = 3.0$. The normalisation of the instrumental line is kept fixed with $\theta_3 = 1.0$.

The projected limits are shown in the lower panel of Fig. 4.9. When the signal becomes degenerate the with instrumental lines, the projected limits weaken, according to Eq. 4.35. We also recover the expected behaviour that when the normalization of the instrumental line is unconstrained (see § 4.3.3) , the projected limit calculated by the Fisher information at the position of the instrumental line diverges.

4.4.3 Example 2: Background systematics described by correlation function

A more general approach towards the modeling of background systematics is to write the total flux as

$$\Phi(E) = \theta_1 \Psi_1(E) + (1 + \delta(E))\Psi_2(E) \ , \tag{4.36}$$

where again Ψ_1 denotes the signal, Ψ_2 the nominal background (we fix $\theta_2 = 1$ throughout), and $\delta(E)$ parameterizes small deviations from the background model. In general, systematic uncertainties in the background will be correlated as a function of energy, and one can define the covariance function

$$\langle \delta(E)\delta(E') \rangle = \Sigma_\delta(E, E') \ , \tag{4.37}$$

and we assume $\langle \delta(E) \rangle = 0$. More specifically, $\delta(E)$ can be thought of as *Gaussian random field* with mean zero, whose behaviour is completely determined by the covariance function. It incorporates information both about the variance and correlation of the systematic.

For any practical calculations, one needs to discretize the field $\delta(E)$. One simple way of doing that is to write $\delta(E)$ as a step function,

$$\delta(E) = \sum_{i=1}^{N} \xi_i \chi_{\Delta E_i}(E) \ , \tag{4.38}$$

with the selector function

$$\chi_{\Delta E_i} = \begin{cases} 1 & \text{if } E \in \Delta E_i \\ 0 & \text{if } E \notin \Delta E_i \end{cases} \ , \tag{4.39}$$

where the (very small) energy bins ΔE_i cover the entire energy range of interest, and E_i will denote the corresponding mean of each energy bin. Furthermore, ξ_i are constrained by a multivariate normal distribution, with a covariance matrix defined by $\Sigma_\delta(E_i, E_j)$. The free parameters in the present example are then $\boldsymbol{\theta} = (\theta_1, \xi_1, \ldots, \xi_N)$.

One can now show that the resulting profiled Fisher information for the signal (assuming a sufficiently fine binning which captures all relevant structure of $\Psi_1(E)$),

is given by (details can be found in Appendix 4.8.4)

$$\widetilde{\mathcal{I}}_{11} = \sum_{ij} \frac{\Psi_1}{\Psi_2}(E_i) D_{ij}^{-1} \frac{\Psi_1}{\Psi_2}(E_j) \,, \tag{4.40}$$

where we defined the combined covariance matrix

$$D_{ij} \equiv \frac{\delta_{ij}\Phi(E_i)}{\Delta E_i [\Psi_2(E_i)]^2} + \Sigma_\delta(E_i, E_j) \,, \tag{4.41}$$

which includes both the effects of Poisson noise and background systematics. This expression becomes more simple in the regime where $\Psi_1 \ll \Psi_2$ since then $\Phi(E) = \Psi_2(E)$. In the limit of no systematics, $\Sigma_\delta \to 0$, we recover the standard expression for the Fisher information of the signal component, Eq. 4.13. On the other hand, in the large sample limit, where Ψ_2 becomes large and hence the first term in Eq. 4.41 small, only the covariance matrix Σ_δ matters and determines the limiting accuracy at which Ψ_1 can be measured. Note that results are independent of the bin size as long as it is sufficiently small to fully resolve the discriminating aspects of the different model components.

This approach of estimating the effect of background uncertainties on expected experimental sensitivities has been already used by some of the present authors in Ref. [157]. We will provide another example below in § 4.5, in context of the Fisher information flux.

4.5 Strategy optimization

Experimental design, or the planning of astronomical observational campaigns, often make use of the SNR of some signals of interest (for the simple additive component models, Eq. 4.10 that we discussed above, this corresponds to $\propto \Psi_i(E)/\sqrt{\Phi(E)}$). One of the goals is to maximize the exposure of energy and/or spatial regions that provide the largest SNR for some component i, which then leads to the tightest constraints on the model parameter θ_i.

We will show here that the above SNR is the simplest realization of the Fisher *information flux*, which we newly introduce here. However, the latter is much more general and can also naturally account for the non-local and saturation effects of background and instrumental systematics. We will illustrate this in an example that makes use of the treatment of correlated background systematics that we discussed in the previous section.

4.5.1 Fisher information flux

In this subsection we will look at the model spectrum Φ as function of the sky coordinate, $\Phi(\Omega|\boldsymbol{\theta})$, such that $I_i(\Omega)$ is the intensity of the signal or background component,[5] and $\mathcal{E}(\Omega)$ is the exposure of the instrument towards Ω, see Eq. 4.11.

With this, we can define the differential *Fisher information flux* that corresponds to parameter pair (i, j) as functional derivative w.r.t. the exposure at position Ω. It is given by

$$\frac{d\mathcal{F}_{ij}}{d\Omega}(\boldsymbol{\theta}, \mathcal{E}) \equiv \frac{\delta \mathcal{I}_{ij}(\boldsymbol{\theta}, \mathcal{E})}{\delta \mathcal{E}(\Omega)} \,, \tag{4.42}$$

where we made explicit that the Fisher information is in general a function of the exposure map $\mathcal{E}(\Omega)$. If we consider the Poisson part of the Fisher information alone, we find

$$\frac{\delta \mathcal{I}_{ij}^{\mathrm{pois}}(\boldsymbol{\theta}, \mathcal{E})}{\delta \mathcal{E}(\Omega)} = \frac{1}{\theta_i \theta_j} \frac{I_i(\Omega) I_j(\Omega)}{I(\Omega|\boldsymbol{\theta})} \,. \tag{4.43}$$

The diagonal part of the Fisher information flux corresponds here to the SNR of component i, and the non-diagonal parts provide information about the degeneracy of the components pairs (i, j).

We emphasize that the right-hand side of Eq. 4.43 does *not* depend on the exposure anymore, and is hence constant during the course of the measurement. Similarly, external constraints like in Eq. 4.34 do not depend on the exposure. In these cases, the full Fisher information flux in Eq. 4.42 would equal the Fisher information flux of the Poisson likelihood, Eq. 4.43. This simple situation changes drastically when considering the *effective information flux* for a subset of the model parameters, as we will see in the next subsection.

The information gain about the parameter pair (i, j) that is obtained by increasing the exposure towards direction Ω by the infinitesimal amount $\delta\mathcal{E}(\Omega)$ is given by

$$\delta \mathcal{I}_{ij}(\boldsymbol{\theta}, \mathcal{E}) = \int d\Omega \, \delta\mathcal{E}(\Omega) \frac{d\mathcal{F}_{ij}}{d\Omega}(\boldsymbol{\theta}, \mathcal{E}) \,. \tag{4.44}$$

As a simple application, let us assume that the change in exposure per time is given by

$$\frac{d\mathcal{E}(\Omega)}{dt} = A_{\mathrm{eff}} \frac{dT_{\mathrm{obs}}(\Omega)}{dt} \,, \tag{4.45}$$

where we factored out the effective area, A_{eff}, and $T_{\mathrm{obs}}(\Omega)$ is the accumulated observation time in direction Ω. Then, the information gain per unit time is given

[5] To keep the notation simple, we ignore here the effects of the instrument point-spread function or energy dispersion. They can be added straightforwardly.

by

$$\frac{d\mathcal{I}_{ij}(\boldsymbol{\theta}, \mathcal{E})}{dt} = A_{\text{eff}} \int d\Omega \, \frac{dT_{\text{obs}}(\Omega)}{dt} \frac{d\mathcal{F}_{ij}}{d\Omega}(\boldsymbol{\theta}, \mathcal{E}) \, . \tag{4.46}$$

Integrating this over time would again give the total information obtained in the observation. Note that these equations remain valid also for the effective information flux that we discuss in the next subsection.

4.5.2 Effective information flux

In order to quantify information gain about PoIs in the presence of background and other uncertainties, the above concept of Fisher information flux needs to be extended to the profiled Fisher information that we introduced in Eq. 4.9. This can be done straightforwardly by applying the functional derivative with respect to $\mathcal{E}(\Omega)$, Eq. 4.42, to the profiled Fisher information.

The full expression for the *effective information flux* for the PoI $(\theta_1, \ldots, \theta_k)$ reads

$$\frac{d\widetilde{\mathcal{F}}_A}{d\Omega} = \frac{d\mathcal{F}_A}{d\Omega} - \frac{d\mathcal{F}_C^T}{d\Omega} \mathcal{I}_B^{-1} \mathcal{I}_C + \mathcal{I}_C^T \mathcal{I}_B^{-1} \frac{d\mathcal{F}_B}{d\Omega} \mathcal{I}_B^{-1} \mathcal{I}_C - \mathcal{I}_C^T \mathcal{I}_B^{-1} \frac{d\mathcal{F}_C}{d\Omega} \, , \tag{4.47}$$

where as above A refers to the $k \times k$ part of the Fisher information matrix that corresponds to the PoI, B refers to the $(n-k) \times (n-k)$ part for the nuisance parameters, and C to the mixing between nuisance parameters and the PoI. This expression appears lengthy, but straightforward to evaluate analytically or numerically if the Fisher information matrix and Fisher information flux are already known.

The effective information flux has a number of surprising and useful properties. First, it is in general not constant in time (in contrast to the plain information flux in Eq. 4.42). In the examples considered here, this is due to *saturation effects*, which are related to the observation reaching the systematic limited regime. Second, it is *non-local*, in the sense that it (for instance) depends on the past (non-)observation of sidebands that could help to characterize the backgrounds in the signal region. Technically, the non-locality is due to the fact that the full Fisher information matrix appears in Eq. 4.47, which includes integrals of the signal and background intensities over Ω. We will illustrate these two aspects in the following final example.

Saturation effects. We demonstrate the saturation effects of the effective information flux with a non-trivial example from § 4.4.3. There, we discussed how to treat correlated background systematics around a fixed background in the Fisher formalism, see Eq. 4.36.

For definiteness, we consider a signal component that consists of two Gaussian

peaks,

$$I_1(E) = 0.05 \times \mathcal{N}(E|\mu = 2, \sigma^2 = 0.01) + \mathcal{N}(E|\mu = 6, \sigma^2 = 4) \,, \tag{4.48}$$

one of which is narrow and the other wide. The background is taken to be flat, $I_2(E) = 1$. The corresponding model count spectra $\Psi_{1,2}(E)$ are obtained by multiplication with the exposure $\mathcal{E}(E)$, as in Eq. 4.11. Lastly, we define the covariance function of the background deviations as,

$$\Sigma_\delta(E, E') = 0.01 \times \mathcal{N}(E|\mu = E', \sigma^2 = 1) + 0.01 \times \mathcal{N}(E|\mu = E', \sigma^2 = 2) \,. \tag{4.49}$$

For simplicity, we focus on the case $\theta_1 = 1$. A sum of two Gaussians was chosen to demonstrate the techniques ability to account for multiple correlation lengths.

The profiled Fisher information for the signal component $\Psi_1(E)$ is given by Eq. 4.40 above. The corresponding effective differential information flux at energy E_k can be obtained by differentiating this expression w.r.t. exposure in energy bin ΔE_k (remember that we have bins in energy for practical purposes). The resulting *effective information flux* of the signal reads

$$\frac{d\widetilde{\mathcal{F}}_{11}}{dE}(E_k) = \sum_{ij} \frac{I_1}{I_2}(E_i) D^{-1}_{ik} \frac{I(E_k)}{\Delta E_k^2 \mathcal{E}(E_k)^2 I_2^2(E_k)} D^{-1}_{kj} \frac{I_1}{I_2}(E_j) \,. \tag{4.50}$$

Here, D_{ij} refers to the combined covariance matrix defined in Eq. 4.41.

It is instructive to consider two limiting cases. If background uncertainties are negligible w.r.t. Poisson noise, the first term in the right-hand side of Eq. 4.41 dominates, and we obtain the pure Poisson information flux

$$\frac{d\widetilde{\mathcal{F}}_{11}}{dE} = \frac{I_1(E)^2}{I(E)} \,. \tag{4.51}$$

On the other hand, for a sufficiently large exposure, the second term in the right-hand side of Eq. 4.41 can dominate, and D_{ij} becomes constant in time. In that case, the effective information flux scales like $\propto \mathcal{E}^{-2}$, leading to a finite total measured information even for very large integration times.[6]

In Fig. 4.10 we show the effective information flux from Eq. 4.50, for different values of the past observation time $T_{\rm obs}$ (remember that $\mathcal{E}(E) = A_{\rm eff} T_{\rm obs}$, and we

[6] This argument does not hold if the vector $x_i \equiv I_1/I_2(E_i)$ has components with zero eigenvalues w.r.t. the matrix $M_{ij} \equiv \Sigma_\delta(E_i, E_j)$. Components with zero eigenvalues correspond to characteristics of the signal that are completely uncorrelated with the modeled background uncertainties, and give rise to non-zero contributions to the effective information flux even after very large integration times. This often undesired behaviour can be removed by adding a small diagonal contribution to the background uncertainty M, which also improves the numerical stability of the matrix inversions.

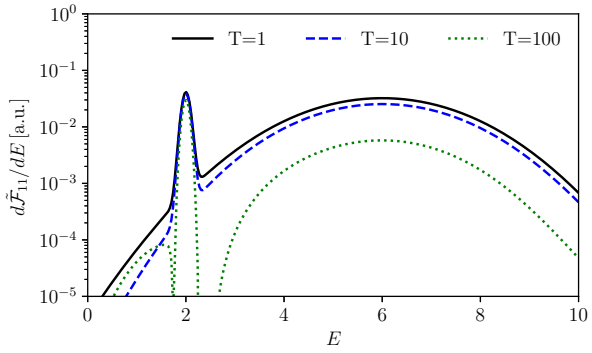

Figure 4.10: **Example of the saturation effects of the effective information flux.** We show here the effective information flux for a signal with narrow and a broad spectral component, on top of a background with correlated uncertainties, after different observation times $T_{\rm obs}$. See Eq. 4.50 and text for further details.

set $A_{\rm eff} = 1$). For small observation times, the flux essentially corresponds to the pure Poisson contribution in Eq. 4.51. However, for larger observation times, the information flux from the broad peak around $E = 6$ becomes increasingly suppressed. This is due to the fact that this broad feature is significantly degenerate with the modeled background uncertainties. On the other hand, the flux from the narrow signal component around $E = 2$, which has a width that is smaller than the correlation scale of the modeled background systematics, remains practically constant.

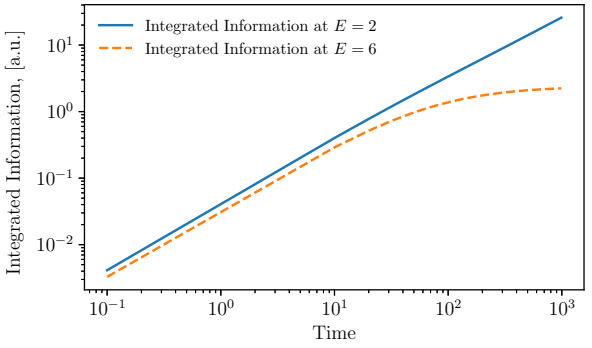

Figure 4.11: **The cumulative information flux at two reference energies.** The corresponding scenario is shown in Fig. 4.10, but integrated over the observation time, as function of the total observation time $T_{\rm obs}$.

To further illustrate the saturation effects when measuring over a long time, we show in Fig. 4.11 the cumulative information obtained by integrating the effective

information flux from Eq. 4.50 at the peaks of the two features, $E = 2$ and $E = 6$. It is clear that at $T = 10^2$ the information obtained from observing the broad peak becomes saturated, while the sharper peak continues to provide information.

Non-locality. We demonstrate the non-locality of the effective information flux with a simple two-component example. With non-locality, we mean that the information flux at E depends in general on the past observation history of $E' \neq E$. This makes sense, since usually a comparable exposure of different observational regions is required to break degeneracies between various model components.

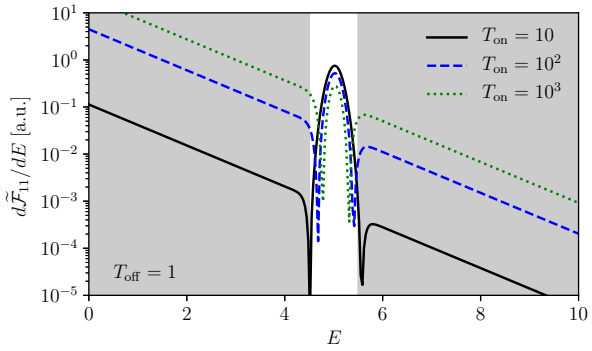

Figure 4.12: **Example for the non-locality of the effective information flux.** We show the effective information flux for a narrow signal on top an exponential background with free normalization. After an initial observation of the sidebands for $T_{\rm off} = 1$, observations are assumed to only take place in the narrow range $E = 4.5$–5.5, and we show the effective information flux after different signal observation times $T_{\rm on}$. See text for details.

In our simple two-component example, the signal flux, $I_1(E)$, is given by a normal distribution, centered at $E = 5$, with a width of $\sigma = 0.2$. The background flux is given by $I_2(E) = 5\exp(-(E-5))$. We assume that after an initial observation of the sidebands with $T_{\rm off} = 1$, observations only take place in the narrow range $E = 4.5$–5.5. We show the resulting effective information flux in Fig. 4.12. At early observation times $T_{\rm on} = 10$, the information flux is completely dominated by observations of the signal region, since due to the initial sideband observations the background in the signal region is already reasonably constrained. However, with growing observation time, the non-observations of the sidebands, becomes increasingly problematic. Consequently, the information flux of the sidebands grows continuously.

4.6 Euclideanised Signals

The model discrimination power of an experiment is often quoted as the significance level z (in terms of standard deviations) at which two benchmark signals $\mathbf{S}_1$ and $\mathbf{S}_2$ can be distinguished. Here, z can be estimated from the expected profile log-likelihood ratio,

$$\mathrm{TS} = -2 \ln \frac{\mathcal{L}_\mathrm{p}(\mathcal{D}_\mathcal{A}(\mathbf{S}_2)|\mathbf{S}_1)}{\mathcal{L}_\mathrm{p}(\mathcal{D}_\mathcal{A}(\mathbf{S}_2)|\mathbf{S}_2)} \,, \tag{4.52}$$

where $\mathcal{D}_\mathcal{A}$ denotes again Asimov data. Using the Fisher approximation, this TS value can be approximated by

$$\mathrm{TS} \simeq \sum_{ij} \Delta S_i D_{ij}^{-1} \Delta S_j \,, \tag{4.53}$$

where $\Delta\mathbf{S} \equiv \mathbf{S}_1 - \mathbf{S}_2$ is the signal difference, and the total (statistic plus systematic) covariance matrix is given by

$$D_{ij} \equiv K_{ij} + \delta_{ij} \frac{S_{2,i} + B_i}{E_i} \,, \tag{4.54}$$

where B_i is the ith background component and E_i is the exposure. The interpretation of the TS value in units of standard deviations is context dependent. We will give here one very typical example. Suppose $\mathbf{S}_1$ and $\mathbf{S}_2$ are part of a parametric signal model, $\mathbf{S}_1 = \mathbf{S}(\boldsymbol{\theta}_1)$ and $\mathbf{S}_2 = \mathbf{S}(\boldsymbol{\theta}_2)$, and that $\boldsymbol{\theta}$ has k relevant components. Furthermore, suppose now that $\boldsymbol{\theta}_1$ are the parameters of the simple null hypothesis that we want to test, and the composite alternative hypothesis that we want to discriminate against is that $\boldsymbol{\theta}$ can aquire any value. The above TS value corresponds then approximately to the median TS value that we would measure in repeated experiments if the true model parameters were $\boldsymbol{\theta}_2$ (we use as data the Asimov data set corresponding to $\boldsymbol{\theta}_2$, in which case the maximum-likelihood estimator of $\boldsymbol{\theta}$ in the alternative hypothesis would be simply $\boldsymbol{\theta}_2$). Since the alternative and null hypothesis are nested, and differ in their degrees of freedom by k, we can assume that the TS value is approximately χ^2_k distributed, with k degrees of freedom [146, 17]. The corresponding threshold value for a $(1-\alpha)$CL contour is then given by

$$Y^2_k(\alpha) = F^{-1}_{\chi^2_k}(1-\alpha) \,. \tag{4.55}$$

For instance, in the case of two parameters, $k = 2$, the 68.7% CL or 95.3% CL contours correspond to $Y^2_{k=2} = 2.32$ and $Y^2_{k=2} = 6.12$ respectively.

Calculating the pair-wise TS-values for N different signals would require N^2 ma-

trix inversions, which is prohibitive if the number of points is large (say, $N \sim 10^6$, which is not a large number in the context of global scans). The entire point of the 'Euclideanized signals' is to enable the rapid calculation of the expected TS value in Eq. 4.53 for a *very* large number of signal combinations (e.g., $\sim 10^{16}$ in the case of a Bayesian scan with 100 million points). This would be possible if the most efficient clustering algorithms, which can handle billions of points, can be used. These, however, happen to work in Euclidean space. The goal is hence to map signals $\mathbf{S}$ onto vectors $\mathbf{x}$ such that the TS value is approximately given by the L2-norm in Euclidean space, $\mathrm{TS} \simeq \|\mathbf{x}_1 - \mathbf{x}_2\|^2$.

We start by approximating the TS-value in Eq. 4.56 by using the Fisher information matrix evaluated at the mean signal $\bar{\mathbf{S}} = \frac{1}{2}(\mathbf{S}_1 + \mathbf{S}_2)$,

$$\mathrm{TS} \simeq \Delta\mathbf{S}^T D_{\bar{\mathbf{S}}}^{-1} \Delta\mathbf{S} \,, \tag{4.56}$$

where $\Delta\mathbf{S} \equiv \mathbf{S}_1 - \mathbf{S}_2$ denotes the signal difference, and $D_{\bar{\mathbf{S}}}$ corresponds to Eq. 4.54 with $S_i \to \bar{S}_i$. If we define now the vector

$$x_i' \equiv \sum_j (D^{-1/2})_{ij} S_j E_j \,, \tag{4.57}$$

it is straightfoward to show that it satisfies

$$\mathrm{TS} \simeq \|\mathbf{x}_1' - \mathbf{x}_2'\|^2 \quad \text{if} \quad D_{\mathbf{S}_1} \approx D_{\mathbf{S}_2} \,. \tag{4.58}$$

However, in the case where the shot noise of the signal has a significant impact on the total background uncertainties, the above relationship will break down. This becomes actually a large effect in the strong-signal limit. In fact, to second order in ΔS_i^2, we find that

$$\|\mathbf{x}_1' - \mathbf{x}_2'\|^2 \approx \frac{1}{4} \sum_i \frac{\Delta S_i^2}{\bar{S}_i} \quad \text{if} \quad \bar{S}_i \gg B_i + E_i K_{ii} \,. \tag{4.59}$$

Hence, we would in this limit underestimate the discrimination power of an instrument systematically by a factor of two. In order to avoid this problem, we multiply $\mathbf{x}'$ with some *fudge factor* that equals one for negligible signals, and equals two when the signal dominates. We will show below that this proceedure leads to satisfactory results. The above discussion motivates the definition of the 'Euclideanized signal',

$$x_i \equiv \left(\sum_j (D^{-1/2})_{ij} S_j E_j \right) \left(1 + \frac{R \cdot S_i}{R \cdot S_i + B_i + K_{ii} E_i} \right) \,, \tag{4.60}$$

where we take the weight $R = 0.1$, which we found to lead to the best results. It has the proprety that the TS-value corresponding to two signals can now be written as

Euclidean distance between their euclideanized signal vectors, as shown in Eq. 4.56.

The mapping depends on the specific noise level of the background as well as the specified background uncertainties. It is defined in the appendix, Eq. 4.60. Here, we just list a few limiting cases. In the background-limited regime where systematics are neglected, we have $x_i \simeq S_i \cdot \sqrt{E_i/B_i}$. In the systematics limited regime, the expression looks like $x_i = \sum_j (K^{-1/2})_{ij} S_j$. In both cases, the signal enters just linearly, since it does not contribute to the background noise. This is however different in the signal limited regime, which requires some extra care. In this limit, we obtain $x_i \simeq 2\sqrt{E_i S_i}$, where the 2 is a fudge factor that compensates for the fact that only the square-root of the signal appears. However, one can show that the latter implies to lowest order in $\Delta S_i/\sqrt{\bar{S}_i}$ that TS $\approx \sum_i \Delta S_i^2/\bar{S}_i$, with $\bar{\mathbf{S}} \equiv \frac{1}{2}(\mathbf{S}_1 + \mathbf{S}_2)$.

In order to test the accuracy of the above proceedure, we randomly generated a large number of models with $n_\mathrm{b} = 10$, using randomized signals $\mathbf{S}$ and randomized covariance matrices K. Without loss of generality, the background and the exposure are kept flat. In Fig. 4.14, we compare the TS values obtained from the profile log-likelihood ratio in Eq. 4.53 with the one obtained from the euclideanized signals, for varying degrees of exposure and magnitude of the covariance matrix.

4.6.1 Validation of approximation methods

First, we motivate heuristically the analytical form of the fudge factor that we used in Eq. 4.60. To this end, we consider the approximate relation

$$\chi^2_G \equiv \frac{(s_1 - s_2)^2}{b + \bar{s}} \approx \chi^2_E \equiv (x_1 - x_2)^2 \,, \tag{4.61}$$

where we defined the mean signal $\bar{s} \equiv \frac{1}{2}(s_1 + s_2)$, and the Euclideanized signal

$$x_i = \frac{s_i}{\sqrt{s_i + b}} \cdot \left(1 + \frac{R \cdot s_i}{R \cdot s_i + b}\right) \tag{4.62}$$

As above, we set the rescaling parameter $R = 0.1$, for which we find the best performance.

The degree to which the approximation in Eq. 4.61 is shown in Fig. 4.13, as function of the signals s_1 and s_2. Only the region that corresponds to a signal difference of $< 5\sigma$ is shown. In this region, the approximation provides a relative agreement to within 16%. The approximation works somewhat worse if the difference between the signals becomes larger, which is however not of much relevance for the intendet applications of our method.

In order to validate the Euclideanized signal method, and to estimate the ex-

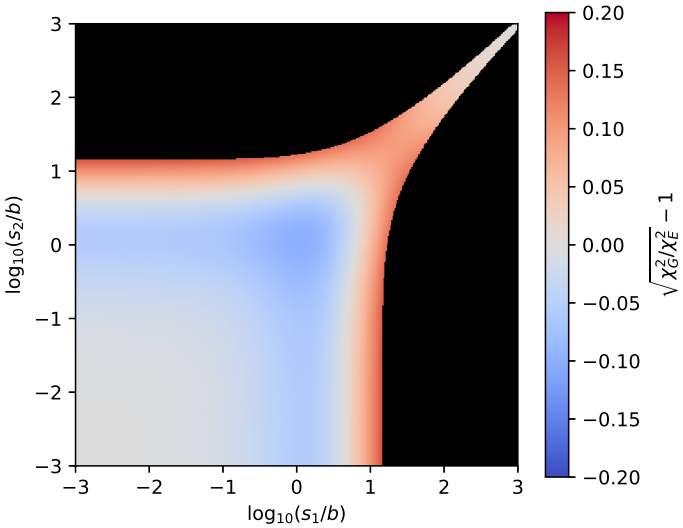

Figure 4.13: **Ratio between $\sqrt{\chi^2_G}$ and $\sqrt{\chi^2_E}$ in Eq. 4.61, as function of s_1/b and s_2/b.** The unmasked area corresponds to $\chi^2_G < 25$, which excludes signal differences larger than $\sim 5\sigma$. Within the unmasked region, the differences between $\sqrt{\chi^2_G}$ and $\sqrt{\chi^2_E}$ are smaller than 16%.

pected approximation errors, we consider a large number of random models. For the purpose of illustration, we kept the models simple. They consists of 3 bins (we find similar results also for a much larger number of bins), with a background set to $\mathbf{B}^T = (1, 1, 1)$, a signal set to $\mathbf{S} = \theta\mathbf{R}$, where $\mathbf{R}$ is a vector of random numbers in the range $[0, 1]$, and $K = k^2 L^T \cdot L$ where L is a random 3×3 matrix with entries in the range $[-1, 1]$. We assume a flat exposure given by E. We consider three benchmark scenarios. First, a *signal limited* case, with $E = 10^{-2}$, $k = 0$, $\theta = 10^{3.5}$. Second, a *systematics limited* case, with $E = 10^6$, $k = 1$, $\theta = 1$. Third, a *background limited* case that is still close to the Poissonian regime, with $E = 10^2$, $k = 0$, $\theta = 1$.

In Fig. 4.14 we compare the TS-value derived via the profile log-likelihood, Eq. 4.52, with the TS-value derived from the Euclideanized signal method, Eq. 4.58. We find that the deviations are largest in the signal-limited case, and up to $\pm 20\%$ on $\sqrt{TS}$. In the systematics limited regime, the deviations are (as expected) much smaller. Interestingly, in the background limited case, we find for the given benchmark point that there is a $\sim 10\%$ bias towards smaller TS-values. This bias disappears if we either increase the exposure (and hence the problem becomes more Gaussian), or if we decrease the exposure and the problem becomes signal dominated. We conclude that *the Euclideanized signal method provides fast estimates for* $\sqrt{\mathrm{TS}}$ *that are correct to within* 20%.

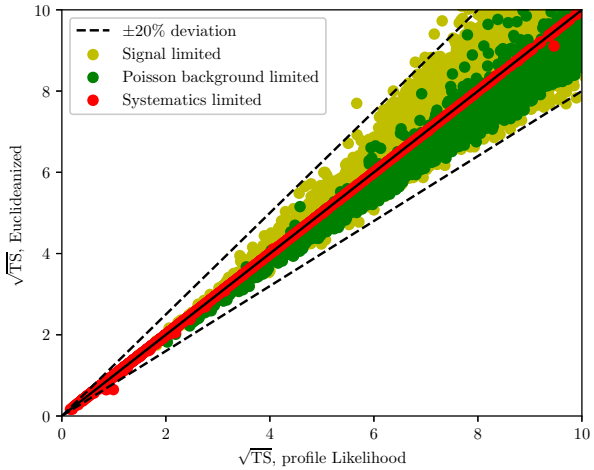

Figure 4.14: **Comparison between exact $\sqrt{\text{TS}}$ value derived from the log profile-likelihood, and the approximate $\sqrt{\text{TS}}$ value derived from the Euclideanized signal method.** The different colors correspond to the signal, the systematics and the (Poisson) background limited regimes (see text for details). As indicated by the dashed lines, the devitations are not larger than 20% for all random models.

4.7 Conclusions

We introduced new methods and concepts for efficient and informative forecasting in astroparticle physics and DM searches, based on the powerful Fisher matrix formalism and unbinned Poisson likelihoods. The Fisher information matrix, Eq. 4.1, is a way of quantifying the maximum information that an observation carries about a set of model parameters. It is at the core of many more advanced statistical methods, and heavily used in other branches of science.

We introduced compact expressions for the approximate derivation of expected exclusion and discovery limits, Eqs. 4.25 and 4.28 (*equivalent counts method*). The equivalent counts method is based on new definitions for the equivalent number of signal and background events, Eqs. 4.18 and 4.19. These are solely based on the Fisher information matrix. The equivalent counts method leads to surprisingly accurate results, even deeply in the Poisson regime, as we showed by comparison with the exact Neyman belt construction and maximum likelihood ratio techniques. In this work, we focused on additive component models and Poisson processes, and neglected the effects of shape uncertainties.

We furthermore showed in two examples how systematic uncertainties can be ef-

ficiently accounted for within the Fisher formalism. In the first example, we assumed that a component of the background is completely degenerate with the shape of the signal, Eq. 4.35. In the second example, we modeled background uncertainties using the idea of Gaussian random fields with arbitrary correlation functions. The resulting expressions, Eq. 4.40, can be efficiently solved numerically by matrix inversion, without profiling or marginalizing over the potentially thousands of nuisance parameters.

Finally, we introduced the new concept of *information flux*, which we obtained from the Fisher information matrix by applying a functional derivative w.r.t. the instrument exposure, Eq. 4.42. For unbinned Poisson likelihoods, it is equivalent to the commonly used signal-to-noise ratio. However, when including the effect of nuisance parameters, the resulting *effective information flux* accounts automatically for the non-local properties and saturation effects of background and instrumental uncertainties. We illustrated these effects in two examples in Figs. 4.10 and 4.12.

Our motivation for this work was to provide a both solid and efficient statistical framework for the systematic study of optimal search strategies for a large range of dark matter models in indirect and other searches which will be the subject of future publications. We furthermore plan to expand the discussion towards models with shape uncertainties, model discrimination and Fisher information geometry.

In summary, we showed how to make the powerful Fisher matrix formalism useful for typical problems in astroparticle physics and DM searches. The equivalent counts method for calculating expected exclusion and discovery limits is applicable in a large range of diverse situations, ranging from direct DM searches to the detectability of extended gamma-ray sources. The effective Fisher information flux is a flexible tool for search strategy optimization, and we expect it to be particularly interesting when confronted with a large number of potential targets, like in indirect searches for DM or a large number of analysis channels in particle colliders. In this work, we just scratched the surface of what can be done with the Fisher formalism, and expect fruitful further theoretical developments of the formalism in the future.

4.8 Appendix

4.8.1 Poisson likelihood properties

We discuss here the general Poisson likelihood function that we use in the main part of the paper, its higher-order derivatives, expectation values and skewness.

Generalized Poisson likelihood

We concentrate in this section on univariate data (the generalization to the multivariate case is straightforward). A typical example are photon energy spectra, measured over some fixed energy range, $E^- \dots E^+$. The data is then fully described by an unordered list of the energies of the N measured photons,

$$\mathcal{D} \equiv \{E_1, E_2, \dots, E_N\} \; . \tag{4.63}$$

The corresponding PDF is

$$P(\mathcal{D}|\boldsymbol{\theta}) = \frac{e^{-\mu(\boldsymbol{\theta})}}{N!} \prod_{i=1}^{N} \Phi(E_i|\boldsymbol{\theta}) \; , \tag{4.64}$$

where $\Phi(E|\boldsymbol{\theta})$ is the model counts spectrum of the photons, and the expected total number of events can be calculated as

$$\mu(\boldsymbol{\theta}) \equiv \int_{E^-}^{E^+} dE \; \Phi(E|\boldsymbol{\theta}) \; . \tag{4.65}$$

The PDF in Eq. 4.64 is correctly normalized to one, which can be checked integrating over photon energies and summing over N.

The unbinned count map $\mathcal{C}(E)$, introduced in Eq. 4.4, carries exactly the same information as the unordered photon list $\mathcal{D}$. It turns out to be useful to rewrite the Poisson likelihood as function of the unbinned count map $\mathcal{C}(E)$. A formal expression that is completely equivalent to Eq. 4.64 is

$$P(\mathcal{C}|\boldsymbol{\theta}) = \exp\left(\int dE \, [\mathcal{C} \ln \Phi - \Phi] - \Gamma(N+1)\right) \; , \tag{4.66}$$

where we used the gamma function instead of the factorial, $N! = \Gamma(N+1)$, and the total number of measured events is given by

$$N \equiv \int_{E^-}^{E^+} dE \, \mathcal{C}(E) \; . \tag{4.67}$$

The key advantage is here that the domain on which $P(\mathcal{C}|\boldsymbol{\theta})$ is defined can be immediately extended to arbitrary functions $\mathcal{C}(E)$, which can be continuous and/or feature non-integer total measured events numbers. This will become useful below.

The likelihood function corresponding to the Poisson process is given by

$$\mathcal{L}(\mathcal{C}|\boldsymbol{\theta}) \equiv A(\mathcal{C}) P(\mathcal{C}|\boldsymbol{\theta}) \; , \tag{4.68}$$

where $A(\mathcal{C})$ is an arbitrary positive function of the data that does not affect the discussion. The logarithm of Eq. 4.68, together with the choice $A(\mathcal{C}) = \exp(-\Gamma(N+1))$, leads Eq. 4.3 in the main text.

Expectation values and 'Asimov data'

Writing the Poisson log-likelihood in the form of Eq. 4.3 has the advantage that it is a linear function of the counts map $\mathcal{C}$. This means that averages over projected data are trivial. Remember that the unbinned count map averaged over many realizations of model $\boldsymbol{\theta}$ just equals the expected count map, $\langle \mathcal{C} \rangle_{\mathcal{D}(\boldsymbol{\theta})} = \Phi(\boldsymbol{\theta})$. Hence, for linear functions of $\mathcal{C}$, averaging over data is equivalent to substituting the unbinned counts map by the 'Asimov' data set [17], $\mathcal{C} \to \Phi(\boldsymbol{\theta})$.

Maximizing $\ln \mathcal{L}$ with respect to $\boldsymbol{\theta}$ requires that $\partial \ln \mathcal{L}/\partial\theta_k = 0$, which is equivalent to

$$\int_{E^-}^{E^+} dE \left(\mathcal{C}(E) \frac{\Psi_k(E)}{\Phi(E|\boldsymbol{\theta})} - \Psi_k(E) \right) = 0 \,, \tag{4.69}$$

for all components k. We adopted here the additive component model defined in Eq. 4.10. Replacing $\mathcal{C}$ by Asimov data for model $\boldsymbol{\theta}$ yields zero, as expected.

For the additive component model, averages over higher order derivatives of the log-likelihood take a simple form. The n-th order derivative (for $n \geq 2$) reads

$$\left\langle \frac{\partial^n(-\ln \mathcal{L})}{\partial\theta_{k_1} \dots \partial\theta_{k_n}} \right\rangle_{\mathcal{D}(\boldsymbol{\theta})} = (-1)^n (n-1)! \int_{E^-}^{E^+} dE \, \frac{\Psi_{k_1}(E) \dots \Psi_{k_n}(E)}{\Phi^{n-1}(E|\boldsymbol{\theta})} \,. \tag{4.70}$$

For $n = 2$, we recover Eq. 4.13 in the main text. The expression for $n = 3$ is important for the discussion of the non-Gaussianity effects in the next subsection.

The Gaussian regime

We start by Taylor expanding the log-likelihood around the model parameter $\boldsymbol{\theta}$,

$$\ln \mathcal{L} = \text{const} + \sum_i \Delta\theta_i \frac{\partial \ln \mathcal{L}}{\partial\theta_i} + \frac{1}{2!} \sum_{ij} \Delta\theta_i \Delta\theta_j \frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j} + \frac{1}{3!} \sum_{ijk} \Delta\theta_i \Delta\theta_j \Delta\theta_k \frac{\partial^3 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j \partial\theta_k} + \dots \,, \tag{4.71}$$

where $\Delta\boldsymbol{\theta}$ is the deviation from the expansion point. If we average the $\ln \mathcal{L}$ now over model realizations $\mathcal{D}(\boldsymbol{\theta})$, and assume Poisson likelihoods, this expansion can be written as

$$\langle -\ln \mathcal{L}(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \rangle_{\mathcal{D}(\boldsymbol{\theta})} = \text{const} + \frac{1}{2} \sum_{ij} \Delta\theta_i \Delta\theta_j \left(\mathcal{I}_{ij} + \frac{2}{3} \sum_k \Delta\theta_k \frac{\partial \mathcal{I}_{ij}}{\partial\theta_k} \right) + \dots \,. \tag{4.72}$$

Note that the additional factor two in front of $\partial\mathcal{I}_{ij}/\partial\theta_k$ comes from the fact that the derivative is here also affecting the parameters of the Asimov data. The precise factor depends on the actual likelihood function and its dependence on the data.

In order to understand the impact of higher-order terms in Eq. 4.72, it is convenient to think about how they affect the significance at the boundaries of confidence intervals. Naively, the boundary of a confidence interval that extends to a threshold value t in a certain direction would correspond to the ellipsoid constructed by values of $\Delta\theta_i$ that satisfy the equation

$$\sum_{ij} \Delta\theta_i \Delta\theta_j \mathcal{I}_{ij} = t \ . \tag{4.73}$$

When taking into account third-order terms from Eq. 4.72, the *actually* realized threshold value at point $\Delta\theta_i$ on the ellipsoid changes by

$$\Delta t = \frac{2}{3} \sum_{ijk} \Delta\theta_i \Delta\theta_j \Delta\theta_k \frac{\partial\mathcal{I}_{ij}}{\partial\theta_k} \ . \tag{4.74}$$

If $\Delta t \ll t$ at all points of the ellipsoid, higher-order (more precisely third order) terms can be indeed ignored and do not affect the result.

In practice, it is usually sufficient to concentrate on the principal axes of the ellipsoid, which correspond to the eigenvectors of $\mathcal{I}_{ij}$. For simplicity, we will here further assume that there are no strong degeneracies between the components. In that case, the principal axes approximately align with the individual parameters θ_i. We can then, for every direction i, require that

$$\frac{\Delta t}{t} = \frac{2}{3} \frac{\Delta\theta_i}{\mathcal{I}_{ii}} \frac{\partial\mathcal{I}_{ii}}{\partial\theta_i} < f \ , \tag{4.75}$$

which means that the fractional change in the threshold value t should be smaller than f. For the Poisson likelihood that we assumed already above in Eq. 4.72, we can use $\Delta\theta_i (\mathcal{I}_{ii})^{-1} \partial\mathcal{I}_{ii}/\partial\theta_i = \Delta\theta_i \sqrt{\mathcal{I}_{ii}}/\sqrt{b_i} = \sqrt{t/b_i}$, where we used the definition of the equivalent background in Eq. 4.19. This implies the condition

$$b_i > \frac{4t}{9f^2} \ , \tag{4.76}$$

which is identical to Eq. 4.31 in the main text and further discussed there.

We emphasize that this is a quite naive estimate, and does not take into account the possible effects of parameter degeneracies, deviations of the log-likelihood ratio from a chi-square distribution, etc. But it gives a useful heuristics for when the Fisher formalism is safe to use, and when it should be used with care. As we have

seen in § 4.3, if the appropriate prescriptions are used, reasonable results for expected exclusion and discovery limits on *one* signal component of interest can be obtained even for vanishing equivalent backgrounds, as long as the other components are well constrained and behaved.

4.8.2 Expected exclusion and discovery limits

We describe here very briefly some of the exact methods that are used to calculate expected exclusion and discovery limits. These are in the main text compared against the results from our EC method.

Neyman belt construction

For details about the Neyman belt construction, we refer to Ref. [144]. We repeat here just the technical result. A conventional one-sided upper limit on the number of expected signal events, s, given k observed events and b expected background events, is given by the s^U that satisfies the equation

$$\sum_{k' \leq k} P(k'|s^U + b) = \alpha \,, \tag{4.77}$$

where α is the significance level of the limit, and $P(c|\mu)$ the PDF of the Poisson distribution with c observed and μ expected counts. Since the connection between k and s^U is monotonic, the median expected exclusion limit can be obtained by considering, for a given expected background b, the median expected count number k. Since k can only acquire discrete values, this introduces jumps in the projected upper limits as function of b, which are clearly visible in Fig. 4.5.

The minimum number of events k that leads to a signal detection with a significance level of (at least) α, above an expected background of b events, is given by the smallest k that satisfies the inequality

$$\sum_{k' \geq k} P(k'|b) \leq \alpha \,, \tag{4.78}$$

which we call here k_{th}. The expected discovery limit of the signal for the given sensitivity level is now given by the smallest s that corresponds to a median count of k_{th}.

Maximum likelihood ratio method

For general likelihood functions with n free parameters, we can define the MLR test statistic (see Ref. [17] for details)

$$TS_{\mathcal{D}}(\theta_1) = -2 \ln \frac{\max_{\theta'_2,\ldots,\theta'_n \geq 0} \mathcal{L}(\mathcal{D}|\theta_1, \theta'_2, \ldots, \theta'_n)}{\max_{\theta'_1,\ldots,\theta'_n \geq 0} \mathcal{L}(\mathcal{D}|\theta'_1, \ldots, \theta'_n)} \,. \tag{4.79}$$

We fix here only one parameter, θ_1, and maximize w.r.t. the remaining ones. We define the modified test statistic

$$q_{\mathcal{D}}(\theta_1) = \begin{cases} TS_{\mathcal{D}}(\theta_1) & \text{if} \quad \theta_1 > \hat{\theta}_1(\mathcal{D}) \\ 0 & \text{otherwise} \end{cases} \,, \tag{4.80}$$

where $\hat{\theta}_1 \geq 0$ is the maximum likelihood estimator given data $\mathcal{D}$. For a given data set, an one-sided confidence interval that corresponds to the desired upper limits is given by

$$C = \{\theta_1 \geq 0 \mid q_{\mathcal{D}}(\theta_1) \leq t(\theta_1)\} \,, \tag{4.81}$$

where the threshold $t(\theta_1)$ is in general a function of θ_1 and depends on the aspired significance level α. The threshold must be set such that C has correct coverage properties. This means that C should cover the true value of θ_1 in $1-\alpha$ of the cases. In the large-sample limit, asymptotic formulae for the statistical distribution of $q_{\mathcal{D}}(\theta_1)$ are available [17]. However, in the small-sample regime, MC simulations are required to derive appropriate threshold values for $t(\theta_1)$. Note that C can be the empty set in some cases, which corresponds to downward fluctuations of the background. Although there are numerous ways to deal with this situations [167, 144], this is not problematic for the purposes of the present work. Note that, due to the discreteness of Poisson processes, $t(\theta_1)$ is in general not a continuous function of θ_1 or the other background parameters.

In the main text, we usually show median limits obtained from a large set of data realizations with $\theta_1 = 0$, using a $t(\theta_1)$ that is derived from MC simulations.

Expected discovery limits are derived in a similar way. We first find the threshold value $t_{\rm th}$ that corresponds to a test of the hypothesis $\theta_1 = 0$ with the significance level of α,

$$P(TS_{\mathcal{D}}(\theta_1 = 0) \geq t_{\rm th} | \theta_1 = 0) \leq \alpha \,. \tag{4.82}$$

The discovery limit is then given by the smallest value of θ_1 that leads to a detection in at least 50% of the cases, namely we search for the θ_1^D that satisfies

$$P(TS_{\mathcal{D}}(\theta_1 = 0) \geq t_{\rm th} | \theta_1 = \theta_1^D) \geq 0.5 \,. \tag{4.83}$$

Note that t_{th} and hence θ_1^D are not necessarily smooth functions of the background parameters $\theta_2, \ldots, \theta_n$, if the discreteness of the Poisson likelihood plays a role.

4.8.3 Technical calculations

We present here some more details about derivations of equations related to discovery limits as well as the treatment of background systematics, used in the main part of the paper.

Expected discovery limits

Given $b \ll 1$ expected background events, one can derive an approximate discovery limit for the number of required signal events s and statistical significance α, by solving

$$P(s+b|b) = \frac{e^{-b} b^{s+b}}{\Gamma(s+b+1)} = \alpha \,, \tag{4.84}$$

for s. Here, $P(s+b|b)$ is the continuum version of the Poisson probability mass function, with $s+b$ observed events while b are expected.

We compare this expression with Eq. 4.28 (we use the symbols s and b for simplicity). In the limit $b \ll s$ it can be written as

$$s \ln\left(\frac{s}{b}\right) - s = \frac{Z^2}{2} \,. \tag{4.85}$$

Now, one can consider the first two terms of the expansion of α in $1/Z$,

$$\ln \frac{1}{\alpha} \simeq \frac{Z^2}{2} + \frac{1}{2} \ln 2\pi Z^2 \,, \tag{4.86}$$

which can be substituted into the right-hand side of Eq. 4.85. The large-s approximation to the log of the gamma function, Stirling's formula, reads (we use $b \ll s$)

$$\ln \Gamma(s+1) \approx s \ln s - s + \frac{1}{2} \ln 2\pi s \,, \tag{4.87}$$

which can be substituted in the left-hand side of Eq. 4.85. One can then rearrange the terms such that they read

$$\frac{b^s}{\Gamma(s+1)} = \alpha \cdot \sqrt{\frac{Z^2}{s}} \,. \tag{4.88}$$

This has exactly the form shown in Eq. 4.30 (remember that $\lambda_i \simeq s_i$ in the low-background limit).

4.8.4 Examples with systematic errors

We will show here in some detail how to arrive at the results in Eqs. 4.35 and 4.40.

In the example leading to Eq. 4.35, we have a three component system with a signal, a background, and some third component which parametrizes variations in the background that are completely degenerate with the signal. The full Fisher matrix of the system is given by

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_{11}^{\rm pois} & \mathcal{I}_{12}^{\rm pois} & \mathcal{I}_{11}^{\rm pois} \\ \mathcal{I}_{12}^{\rm pois} & \mathcal{I}_{22}^{\rm pois} & \mathcal{I}_{12}^{\rm pois} \\ \mathcal{I}_{11}^{\rm pois} & \mathcal{I}_{12}^{\rm pois} & \mathcal{I}_{11}^{\rm pois} + \frac{1}{\xi_3^2} \end{pmatrix} , \tag{4.89}$$

where we already used the various symmetry properties of the Fisher information matrix elements as well as the fact that $\Psi_1 = \Psi_3$. We are interested in the profiled 1×1 Fisher matrix where the parameter of interest is the signal parameter θ_1, and we have removed θ_2 and θ_3. This profiled Fisher matrix is given by

$$\widetilde{\mathcal{I}} = \mathcal{I}_{11}^{\rm pois} - \begin{pmatrix} \mathcal{I}_{12}^{\rm pois} \\ \mathcal{I}_{11}^{\rm pois} \end{pmatrix}^T \begin{pmatrix} \mathcal{I}_{22}^{\rm pois} & \mathcal{I}_{12}^{\rm pois} \\ \mathcal{I}_{12}^{\rm pois} & \mathcal{I}_{11}^{\rm pois} + \frac{1}{\xi_3^2} \end{pmatrix}^{-1} \begin{pmatrix} \mathcal{I}_{12}^{\rm pois} \\ \mathcal{I}_{11}^{\rm pois} \end{pmatrix} . \tag{4.90}$$

It is straightforward to invert the 2×2 matrix analytically, and one can show that the inverse of the profiled Fisher matrix can be written in the simple form

$$\frac{1}{\widetilde{\mathcal{I}}} = \xi_3^2 + \frac{1}{\mathcal{I}_{11}^{\rm pois} - (\mathcal{I}_{22}^{\rm pois})^{-1}(\mathcal{I}_{12}^{\rm pois})^2} . \tag{4.91}$$

If we now identify $\sigma_1^2(\boldsymbol{\theta}) = \widetilde{\mathcal{I}}^{-1}$, and $(\sigma_1^{\rm pois})^2(\boldsymbol{\theta}) = (\mathcal{I}_{11}^{\rm pois} - (\mathcal{I}_{22}^{\rm pois})^{-1}(\mathcal{I}_{12}^{\rm pois})^2)^{-1}$, we arrive at Eq. 4.35. Note that the latter is just the profiled Fisher information that we would have obtained in absence of the third component, or equivalently in the limit $\xi_3 \to \infty$.

The derivation of Eq. 4.40 follows a similar pattern, but is technically slightly more involved. Again, we are interested in the profiled 1×1 Fisher information for the signal component. It is here useful to associate the index $i = 0$ with the signal component, and the indices $i = 1, \ldots, N$ with the discrete energies from Eq. 4.38. If we think about the underlying full Fisher matrix of the system in the block form shown in Eq. 4.8, then the components A (associated with $i = 0$), B and C are given by

$$A = \sum_{i=1}^{N} \Delta E_i \frac{\Psi_1(E_i)^2}{\Phi(E_i)} , \tag{4.92}$$

where we discretized the integral,

$$C_i = \Delta E_i \frac{\Psi_1(E_i)\Psi_2(E_i)}{\Phi(E_i)} = \Delta E_i \Psi_1(E_i) \,, \tag{4.93}$$

where $i = 1, \ldots, N$ and we used in the second step that $\Phi(E) = \Psi_2(E)$, and

$$B = \delta_{ij} \Delta E_i \Phi(E_i) + \Sigma_\delta^{-1} \,, \tag{4.94}$$

where we included the inverse of the covariance matrix for ξ_i. Then, the profiled Fisher information can be written as

$$\begin{aligned}
\widetilde{\mathcal{I}}_{11} &= \sum_{i=1}^{N} \Delta E_i \frac{\Psi_1(E_i)^2}{\Phi(E_i)} - \sum_{i,j=1}^{N} \Delta E_i \Delta E_j \Psi_1(E_i)\Psi_1(E_j) \left[\mathrm{diag}(\Delta E_i \Psi_2(E_i)) + \mathbf{\Sigma}_\delta^{-1}\right]^{-1}_{ij} \\
&= \sum_{i=1}^{N} \Delta E_i \frac{\Psi_1(E_i)^2}{\Phi(E_i)} - \sum_{i,j=1}^{N} \sqrt{\frac{\Delta E_i \Delta E_j}{\Phi(E_i)\Phi(E_j)}} \Psi_1(E_i)\Psi_1(E_j) \left[\frac{\mathbf{\Sigma}'_\delta}{1 + \mathbf{\Sigma}'_\delta}\right]_{ij} \\
&= \sum_{i,j=1}^{N} \sqrt{\frac{\Delta E_i \Delta E_j}{\Phi(E_i)\Phi(E_j)}} \Psi_1(E_i)\Psi_1(E_j) \left[\frac{1}{1 + \mathbf{\Sigma}'_\delta}\right]_{ij} \,.
\end{aligned} \tag{4.95}$$

Here, we used the definitions $(\mathbf{\Sigma}_\delta)_{ij} = \Sigma_\delta(E_i, E_j)$ and

$$(\mathbf{\Sigma}'_\delta)_{ij} = \sqrt{\Delta E_i \Delta E_j \Phi(E_i)\Phi(E_j)}\, \Sigma_\delta(E_i, E_j) \,. \tag{4.96}$$

In the first step, we rearrange some factors of ΔE_i and $\Phi_1(E_i)$, and use the general matrix relation $(1+M^{-1})^{-1} = M/(1+M)$. In the second step, we include the matrix identity in the form $\delta_{ij} = [(1 + \mathbf{\Sigma}'_\delta)/(1 + \mathbf{\Sigma}'_\delta)]_{ij}$ in the first summation which helps to further collapse the whole expression. The last line is after some more rewriting equivalent to Eq. 4.40.

5 Benchmark-free Forecasting for Direct Searches

This chapter is based on work from Ref. [2]

5.1 Introduction

Searching for the elusive Dark Matter (DM) particle has been the preoccupation of physicists for many years [168, 169]. Over the past decade, two-phase scintillator direct detection experiments [170, 171] have found much success with the LUX [136], XENON [172] and PANDA-X [20] collaborations providing the most stringent constraints on DM particles in the GeV − TeV mass range to date. Such experiments will continue to improve in sensitivity for many years to come. In the case of a detection, it should be possible to study the astro- and particle-physics properties of DM using a variety of detectors and detection methods (see [173, 174, 175, 176, 177] and many others), but the precise parameter regions in which these properties can actually be measured is hard to quantify.

Exploring the prospects for discriminating between different DM-nucleon interactions usually relies on comparing a number of benchmark models [178, 179, 180, 181, 182, 183, 184]. However, the pair-wise comparison of different benchmark points in the model parameter space (DM couplings or masses) is time-consuming, does not scale well with the number of benchmark points, and is in particular problematic in high-dimensional parameter spaces. In direct detection, such a high-dimensional parameter space appears in the framework of non-relativistic effective field theory (NREFT) [122, 123, 185, 186, 187], in which the space of DM-nucleon interactions may have more than 30 dimensions [188, 189, 190]. With current techniques, it is hence difficult to study model degeneracies and the degeneracy-breaking power of future instruments in a reliable and exhaustive way. For such tasks, dedicated techniques are required.

In this *Letter*, we introduce a new framework for studying the signal discrimination

power of future detectors in a fundamentally benchmark-free way. The key questions we aim to address are: How many *observationally distinct* signals does a given model predict for a set of future experiments? How many of these signals are compatible with specific subsets of the signal model? In which regions of parameter space is signal discrimination and parameter reconstruction possible?

We first summarize the basics of the non-relativistic effective field theory approach to direct deteciton, as well as the basics of our approach. We then discuss the dark matter models and experiments we consider in the current work. Finally, we show our results and conclude with a short discussion about possible future directions and applications[1].

5.2 Information Geometry

Consider a New-Physics model $\mathcal{M}$ with some d-dim model parameter space, $\vec{\theta} \in \Omega_{\mathcal{M}} \in \mathbb{R}^d$, and a combination of future experiments X that are described by some likelihood function $\mathcal{L}_X(\mathcal{D}|\vec{\theta})$, where $\mathcal{D}$ is data. We expect that two model parameter points $\vec{\theta}$, $\vec{\theta}'$ can be discriminated by experiments X if the parameter point $\vec{\theta}'$ is inconsistent with Asimov data [17] $\mathcal{D} = \bar{\mathcal{D}}(\vec{\theta})$. More concretely, distinctiveness requires that the log-likelihood ratio

$$\mathrm{TS}(\vec{\theta}')_{\bar{\mathcal{D}}(\vec{\theta})} \equiv -2 \ln \frac{\mathcal{L}(\bar{\mathcal{D}}(\vec{\theta})|\vec{\theta}')}{\max\limits_{\vec{\theta}''} \mathcal{L}(\bar{\mathcal{D}}(\vec{\theta})|\vec{\theta}'')} \simeq \|\vec{x}(\vec{\theta}) - \vec{x}(\vec{\theta}')\|^2 \,, \tag{5.1}$$

exceeds a threshold value $r_\alpha(\mathcal{M})^2$. The threshold value depends on the chosen statistical significance, which we set here to $\alpha = 0.045$ (2σ), as well as the sampling distribution of $\mathrm{TS}(\vec{\theta}')$. In the large-sample limit and under certain regularity conditions, the sampling distribution follows a χ^2_k distribution with $k = d$ degrees of freedom [145, 17].

The last part of Eq. 5.1 is an approximation based on the 'euclideanized signal' method [6], an embedding $\vec{\theta} \mapsto \vec{x}(\vec{\theta}) \in \mathbb{R}^n$ into some, usually higher-dimensional, space with unit Fisher information matrix (n usually equals the total number of data bins). This approximation maps statistical distinctiveness onto euclidean distances, and works to within 20% if the number of counts is order one, see [6] for a discussion and caveats. Confidence regions in the model parameter space correspond then to hyperspheres of radius $r_\alpha(\mathcal{M})$ in the euclideanized signal space.

Often one is interested in sub-models $\mathcal{S}$ that are nested inside model $\mathcal{M}$, and

[1] Code associated with the paper available at https://github.com/tedwards2412/benchmark_free_forecasting/.

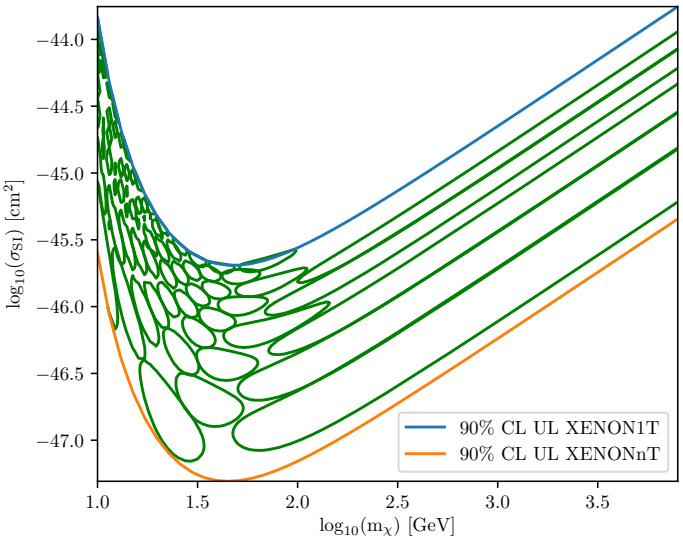

Figure 5.1: **Limit on the SI cross-section vs. DM mass, assuming operator $\mathcal{O}_1$, for a XENONnT detector.** Contours show 68% confidence contours in $d = 2$ dimensions. The radius of these contours in the Euclidean space is therefore $r_\alpha(\mathcal{M}) = 1.52$. The number of discriminable signals in the blue+blue/orange region of the middle panel of Fig. 5.2 can be approximated by counting the number of closed green contours.

which are obtained by restricting $\mathcal{M}$ to a d'-dim subregion $\Omega_\mathcal{S} \subset \Omega_\mathcal{M}$. A parameter point $\vec{\theta}$ of $\mathcal{M}$ leads to a signal that is distinct from any signal in submodel $\mathcal{S}$ if

$$-2\ln\frac{\max_{\vec{\theta}'\in\Omega_\mathcal{S}}\mathcal{L}(\bar{\mathcal{D}}(\vec{\theta})|\vec{\theta}')}{\max_{\vec{\theta}'}\mathcal{L}(\bar{\mathcal{D}}(\vec{\theta})|\vec{\theta}')} \simeq \min_{\vec{\theta}'\in\Omega_\mathcal{S}}\|\vec{x}(\vec{\theta})-\vec{x}(\vec{\theta}')\|^2\,, \tag{5.2}$$

exceeds a certain threshold value $r_\alpha(\mathcal{S},\mathcal{M})^2$. Here 'distinct' means that the composite null hypothesis $\mathcal{S}$ can be rejected for data $\bar{\mathcal{D}}(\vec{\theta})$. The sampling distribution of Eq. 5.2 follows in general a $\chi^2_{k=d-d'}$ distribution. In the euclideanized signal space $\vec{x}$, parts of model $\mathcal{M}$ cannot be discriminated from model $\mathcal{S}$ that lie within a 'shell' of thickness $r_\alpha(\mathcal{S},\mathcal{M})$ around the signal manifold of $\mathcal{S}$.

Finally, nuisance parameters can be accounted for by replacing the likelihood function in Eq. 5.1 with a profile likelihood, $\mathcal{L}(\mathcal{D}|\vec{\theta}) = \max_{\vec{\eta}}\mathcal{L}(\mathcal{D}|\vec{\theta},\vec{\eta})\mathcal{L}_\eta(\vec{\eta})$, where the last factor can incorporate additional constraints on the nuisance parameters from data external to X.

5.3 Distinct signals

To quantify the signal discrimination power of a set of future experiments X in the context of model $\mathcal{M}$ we may define the figure of merit

$$\nu^{\alpha}_{\mathcal{M},X}(\Omega_{\mathcal{M}}) \simeq \begin{array}{l}\text{Total number of signals from model } \mathcal{M} \\ \text{discriminable by experiments } X.\end{array}$$

More specifically, $\nu^{\alpha}_{\mathcal{M},X}$ equals the maximum number of points that can populate the parameter space of $\mathcal{M}$ while remaining mutually distinct according to Eq. 5.1. Any such set of points provides a complete sample of the phenomenological features of model $\mathcal{M}$. Loosely speaking, the points correspond to a set of non-overlapping confidence contours as shown in Fig. 5.1. Furthermore, when considering a sub-model $\mathcal{S}$ nested in $\mathcal{M}$, we can define

$$\nu^{\alpha}_{\mathcal{M},X}(\Omega_{\mathcal{S}}) = \begin{array}{l}\text{Number of signals from model } \mathcal{M} \text{ dis-} \\ \text{criminable by experiments } X, \text{ and con-} \\ \text{sistent with model } \mathcal{S}.\end{array} \tag{5.3}$$

With these definitions, the phenomenological distinctiveness of various regions in the parameter space of model $\mathcal{M}$ can be visualized using standard Venn diagrams [191]. The technical definition for the measure $\nu^{\alpha}_{\mathcal{M},X}(\cdot)$, which is used in the subsequent examples, is given in Appendix A of the supplementary material.

5.4 DM-nucleon interactions

While direct detection is typically analysed in terms of the standard spin-dependent (SD) and spin-independent (SI) interactions [192], the range of possible signals is much broader. The framework of non-relativistic effective field theory (NREFT) [122, 123, 185, 186, 187] aims to classify possible elastic DM-nucleon interactions and thus possible signals in DM-nucleus scattering experiments. NREFT is realised as a power series in the DM-nucleus relative velocity $\vec{v}$ and the nuclear recoil momentum $\vec{q}$, valid for non-relativistic, short-range interactions. The resulting operators (labelled $\mathcal{O}_1$, $\mathcal{O}_3$, $\mathcal{O}_4$, ...) give rise to a range of novel energy spectra [193, 179, 181, 184], directional signals [194, 195] and annual modulation signatures [196, 197]. We focus here on the three operators $\mathcal{O}_1$, $\mathcal{O}_4$ and $\mathcal{O}_{11}$ because they allow us to explore a diverse range of signals with only a small number of operators[2]. Operator $\mathcal{O}_1 = 1_\chi 1_N$ couples to nucleon number while the operator $\mathcal{O}_4 = \vec{S}_\chi \cdot \vec{S}_N$ couples to nuclear spin, allowing us to explore the complementarity between nuclei of different size and spin [200].

[2] We neglect the effects of operator mixing [198, 199] which require us to specify the structure of the dark sector.

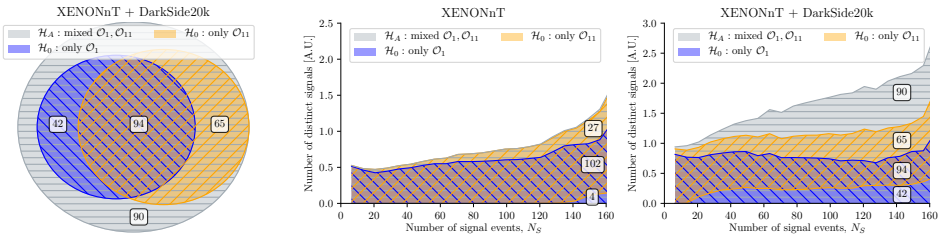

Figure 5.2: **Signal-discrimination power of combinations of direct detection experiments, summarized in *infometric Venn diagrams.*** *Central/Right panel:* The full area corresponds to the *number of distinct detectable signals* within the alternative hypothesis $\mathcal{H}_A$, unfolded as function of number of XENONnT signal events. The subsets indicate the fraction of signals consistent with various null hypotheses $\mathcal{H}_0$. Overlapping subsets correspond to signals consistent with multiple null hypotheses simultaneously. The numbers correspond to $\nu_{\mathcal{M},X}(\Omega_{\mathcal{M}})$ in each region. In the right panel, the overlapping (blue+orange) region corresponds to the model parameters between the purple-dot-dashed and orange contours in Fig. 5.3. The non-overlapping (blue-only) $\mathcal{O}_1$ region corresponds to parameters between the purple-dashed and blue contours in Fig. 5.3. *Left panel:* Standard Venn diagram summed over number of signal events.

Operator $\mathcal{O}_{11} = i\vec{q} \cdot \vec{S}_\chi / m_N$ may arise as the leading-order interaction in certain scalar-mediated models [187]. Similar to $\mathcal{O}_1$, it also couples to nucleon number and receives a coherent enhancement to the rate, but has a characteristic peaked recoil spectrum owing to an extra $\mathrm{d}\sigma/\mathrm{d}E_R \propto E_R$ scaling of the cross section [181]. This allows us to explore how well different recoil spectra can be discriminated in future experiments.

Unfortunately, NREFT cannot encompass all possible signals. In particular, in its original formulation [123] it cannot describe interactions through light mediators. In this case, the typical momentum transfer is larger than the mediator mass and an expansion in q is no longer appropriate[3]. The scenario in which this mediator is the Standard Model photon has been studied extensively [202, 126, 203]. Here, we consider millicharged DM [204] which has long-ranged, coherently-enhanced interactions with nuclei, with a differential cross section scaling as E_R^{-2} [205, 126, 125]. Alternatively, DM may have non-zero electric and magnetic moments [206, 207], particularly if it takes the form of a composite state, such as a Dark Baryon [208, 209]. In the context of model discrimination, most interesting for us will be the magnetic dipole, $(\mu_\chi/2)\,\overline{\chi}\sigma^{\mu\nu}\chi\,F_{\mu\nu}$, which leads to both long-range and short-range contributions to the rate, arising from charge-dipole and dipole-dipole interactions respectively [210, 211, 203].

[3]Note, however, that because the DM is still non-relativistic, the effects of light mediators can be incorporated into the NREFT by including the appropriate propagator [126, 201].

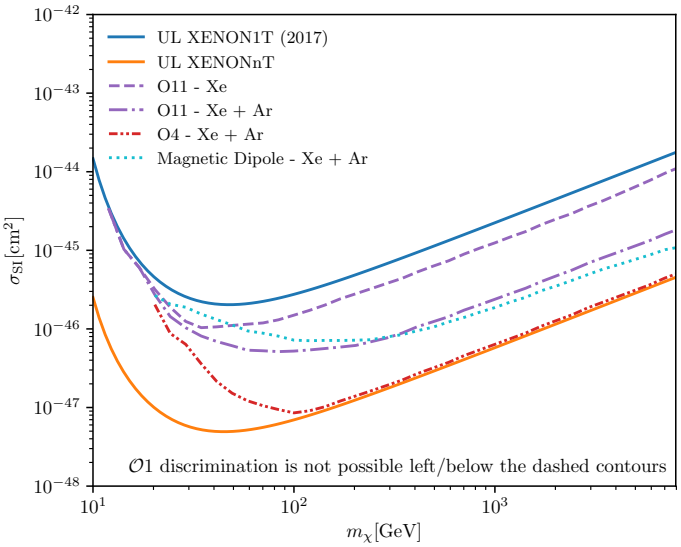

Figure 5.3: **Discriminability of DM interactions.** To the left/below each broken line, it is not possible to discriminate an $\mathcal{O}_1$-signal (with the indicated cross-section and DM mass) from the corresponding best-fit $\mathcal{O}_4$, $\mathcal{O}_{11}$ or magnetic dipole signal. Above/right of each broken line, such a discrimination is possible with at least 2σ significance. Solid lines display 90% CL limits for XENON1T-2017 and XENONnT. All lines include DM halo uncertainties.

The five DM-nucleon interaction models we have outlined above encompass a range of phenomenologically-driven as well as more theoretically-motivated models, leading to a wide range of direct detection signals. We calculate the signal spectra in each case using the publicly-available code WIMpy [212], implementing expressions from [186] and [126][4]. The required nuclear response functions are taken from the mathematica package provided in [186], supplemented by those calculated in [213]. We assume iso-spin conserving ($c_p = c_n$) NREFT interactions and that the particle producing a signature makes up 100% of the local DM density (which we fix to $\rho_\chi = 0.3$ GeV cm^{-3} [72, 214]). We incorporate standard Gaussian halo uncertainties from [214]; details can be found in Appendix B of the supplementary material.

5.5 Direct dark matter searches

We implement two toy detectors, designed to resemble the expected advancement in direct DM searches over the upcoming 5-10 years.

[4]Note that the operator normalisations in [126] and [186] differ.

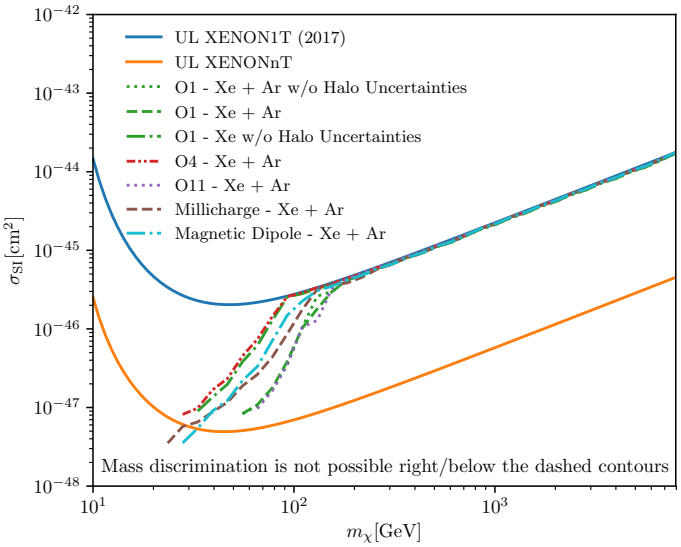

Figure 5.4: **Discriminability of the DM mass.** To the right of each broken line, the DM mass is unbounded from above (at 2σ). Lines for other operators are mapped onto $\sigma_{\rm SI}$ by converting to an effective cross section and rescaled to match $\mathcal{O}_1$ at high masses. For $\mathcal{O}_1$, we also show constraints for the Xenon-only case and when halo uncertainties are neglected. Solid lines display 90% CL limits for XENON1T-2017 and XENONnT.

We implement a Xenon detector in light of the stringent constraints on $\mathcal{O}(10)$ GeV DM set by XENON1T [172], with numerous Xenon-based experiments on the horizon [215, 216, 217, 218]. We model this detector on the future *XENONnT* [217] experiment. In addition we implement a detector containing a target material with no nuclear spin, namely Argon, modeling this detector on *Darkside20k* [137]. In this way we maximize discriminability of spin-dependent operators[5]. Our detector implementations and background assumptions are briefly described in Appendix B of the supplementary material.

5.6 Results

In Fig. 5.1, we show the expected 68% CL reconstruction regions for a set of mutually distinct parameter points, for our *XENONnT*-like detector. The confidence regions are constructed by querying spheres with radius $r_\alpha(\mathcal{M}) = 1.52$ in the euclideanized

[5]Liquid noble detectors typically do not have sensitivity to DM particles lighter than a few GeV, so we restrict our attention to $m_\chi > 10$ GeV in the current work.

signal space. The number of these regions corresponds approximately to the figure of merit in Eq. 5.3.

In the central panel of Fig. 5.2, we illustrate the power of *XENONnT* to discriminate between operators in the 3-dimensional model space $\mathcal{M}$ of mass, $\mathcal{O}_1$, and $\mathcal{O}_{11}$. With increasing numbers of events, the number of discriminable signals increases, though the majority of signals are compatible with both $\mathcal{O}_1$ and $\mathcal{O}_{11}$. In the right panel of Fig. 5.2, we include also information from *DarkSide20k*. The addition of an Argon detector not only increases the number of discriminable signals (from 133 to 291) but also enlarges the region of parameter space where $\mathcal{O}_1$ and $\mathcal{O}_{11}$ can be discriminated from each other. The left panel of Fig. 5.2 corresponds to the same scenario as the right panel, instead summed over the number of *XENONnT* signal events. 160 signal events approximately corresponds the expected number of events in XENON-nT if the true model were at the current sensitivity. We note that the Venn diagrams we have introduced here significantly increase in complexity when comparing a large number of models at once. However, we emphasize that the number of discriminable regions, Eq. 5.3, is completely general and remains a useful measure for assessing model discriminability.

Figure 5.3 shows the regions of the parameter space of spin-independent ($\mathcal{O}_1$) DM in which discrimination from $\mathcal{O}_4$, $\mathcal{O}_{11}$ and magnetic dipole DM would be possible. For $\mathcal{O}_4$ (spin-dependent), 2σ discrimination is possible at high DM mass even down to small cross sections, when both Xenon and Argon experiments are used. The spin-zero Argon nucleus has no spin-dependent coupling, so we can discriminate well as long as the Argon detector has sensitivity ($m_\chi > \mathcal{O}(20\text{ GeV})$, below which most recoils are below the 32 keV threshold).

For $\mathcal{O}_{11}$, discrimination is possible at high mass because of the different spectral shapes of $\mathcal{O}_1$ and $\mathcal{O}_{11}$, though cross sections around $\sim 10^{-46}$ cm^2 are required to obtain enough events to map out the spectra precisely. At low mass, the peak in the $\mathcal{O}_{11}$ spectrum falls below the threshold of the experiments; for both $\mathcal{O}_1$ and $\mathcal{O}_{11}$ the exponentially falling tail of the DM velocity distribution dominates the spectral shape [219], making discrimination impossible.

For Magnetic Dipole interactions, discrimination is also possible at high mass, given enough signal events. We note a 'kink' in the boundary for Magnetic Dipole interactions around $m_\chi \sim 20$ GeV. For large DM masses, the short-range spin-dependent dipole-dipole contribution begins to dominate [203]. In this case, discrimination prospects are good with the inclusion of the spin-zero Argon detector.

For the mock detectors we consider, SI interactions cannot be distinguished from Millicharged DM, which is not shown in Fig. 5.3. The recoil spectrum for Millicharged DM is similar to $\mathcal{O}_1$, but has an extra E_R^{-2} suppression. This more rapidly falling recoil spectrum can be mimicked by an SI interaction with smaller DM mass. As

demonstrated in Refs. [181, 201], low-threshold semi-conductor experiments are required to distinguish between the two interactions.

Finally, Fig. 5.4 shows, for various operators, the regions of parameter space where a closed contour for the DM mass would be possible at the 2σ level. At large DM masses, the kinematics of the interaction mean that the recoil spectrum becomes independent of the DM mass, meaning that to the right of the curves in Fig. 5.4, it is not possible to obtain an upper limit on m_χ [220, 221]. For $\mathcal{O}_1$ we show the regions for Xenon-only, as well as for Xenon and Argon combined without halo uncertainties to demonstrate the improvement in mass reconstruction when including a second detector. When the two detectors are combined halo uncertainties make little difference to the mass discrimination, as changes in the velocity distribution affect the spectra in the two detectors differently [173]. Operator $\mathcal{O}_4$ contains the largest region in which the mass cannot be constrained due to the lack of signal in Argon.

Even in the most optimistic case of cross sections just below the current bounds, it is not possible to pin down the DM mass for $m_\chi \gtrsim 100$ GeV. Previous works have demonstrated, typically using a small number of benchmarks [221, 222, 181], that DM mass reconstruction worsens for large masses; here, we have mapped out precisely where this mass reconstruction fails as a function of mass and cross section.

5.7 Discussion

The methods introduced in this paper allow us to efficiently characterize and visualize the phenomenological distinctiveness of direct DM signal models in infometric Venn diagrams, as shown in Fig. 5.2. Furthermore, these methods allow for an efficient exploration of the phenomenology of complex models, and hence allow us to make 'benchmark-free' statements like those shown in Figs. 5.3 and 5.4. In Fig. 5.3 we see that ruling out non-standard interactions is harder for light DM, while in Fig. 5.4 we see that we cannot pin down the DM mass for masses larger than ~ 100 GeV This leaves only a small region of parameter space – for $m_\chi \in [20, 100]$ GeV and cross sections a factor of a few below current bounds – in which the DM mass and interaction can both be constrained at the 2σ level with near-future Xenon and Argon detectors. Such general statements would not have been possible without an efficient exploration of the Dark Matter parameter space, made feasible with the tools presented here. Third generation experiments such as DARWIN [216] will have a far greater sensitivity. More events would dramatically improve our ability to constrain different models, particularly for models at the current XENON1T bound.

This work paves the way for a more complete exploration of the direct detection parameter space and a deeper understanding of the complementarity between detec-

tors. Future work should explore the possibility to discriminate between a wider range of interactions, beyond the subset of five we include here. In addition, the techniques we present may be used to optimize detector properties (target material, thresholds, etc.) in order to understand how operator discrimination can be improved at low DM mass.

Our 'benchmark-free' method rests on the 'euclideanised signal' introduced in [6], and works for any Poisson (and hence Gaussian) likelihood function, as long as background uncertainties are sufficiently Gaussian. Euclideanised signals therefore provide a useful forecasting tool for a wide range of New-Physics signals, including those in cosmology, indirect DM detection, and collider searches. As we have shown, using direct detection alone may not allow us to completely constrain the DM properties. Combining complementary information from other search strategies, coupled with new techniques for efficient forecasting, will provide essential guidance in the future of Dark Matter detection.

5.8 Appendix

5.8.1 Technical details

First, we generate a large number of points in the parameter space of model $\mathcal{M}$, $\vec{\theta}_i \in \Omega_{\mathcal{M}}$ (typically of the order 10^5). The details of the distribution of the points do not matter in the limit of a large number of parameter points. Each point is projected onto the corresponding euclideanized signal, $\vec{x}_i = \vec{x}(\vec{\theta}_i)$. This projection depends on the details of the detector. If multiple experiments are used, the corresponding vectors are concatenated. Euclideanized signals are generated using `swordfish` [6] (see paper for technical details). This process essentially provides a sample of the model parameter space $\mathcal{M}$ embedded in the (usually higher dimensional) euclideanized signal space. In addition, the mapping between these spaces is known if $\mathcal{M}$ is sufficiently sampled. This sample of parameter points and corresponding euclideanized signals are the basis for the various estimation techniques used in this work.

Confidence contours. Given the sample of projected points, $\vec{x}_i$, it is now straightforward to generate expected contours around any parameter $\vec{\theta}_0 \in \Omega_{\mathcal{M}}$ in the model parameter space. Such regions are for instance shown in Fig. 5.1. To this end, we first calculate the projected signal $\vec{x}_0 = \vec{x}(\vec{\theta}_0)$. Then, using standard nearest neighbor finder algorithms [223], we identify the set of points $\vec{\theta}_i$ that are within a radius $r_\alpha(\mathcal{M})$ of point $\vec{\theta}_0$. Here, $r_\alpha(\mathcal{M})$ depends on the dimensionality of the parameter space $\Omega_{\mathcal{M}}$ as well as the significance level of interest, $r_\alpha(\mathcal{M}) = \sqrt{\chi^2_{k=d,\mathrm{ISF}}(1-\alpha)}$ where $\chi^2_{k=d,\mathrm{ISF}}$ is the inverse survival function of the Chi-squared distribution with

$k = d$ degrees of freedom. For the 3-dim models that we consider in this paper, and a significance level of $\alpha = 0.045$, we have for instance $r_\alpha(\mathcal{M}) = 2.84$. Now, the points in the model parameter space $\Omega_\mathcal{M}$ that belong to the confidence region can be simply identified by back-projecting the nearest neighbors in euclideanized signal space (obviously this back-projecting just requires a look-up in the original list of model parameters). In this way, the generation of confidence regions around arbitrary signal points is efficiently achieved.

Number of distinct signals. For Fig. 5.2 we are interested in the (maximum) number of points that can populate the model parameter space $\Omega_\mathcal{M}$ such that the model points can be discriminated in the sense of Eq. 5.1. This is equivalent to finding the maximum number of points in the euclideanized signal space that can populate the embedding of $\Omega_\mathcal{M}$ such that their mutual distance is at least $r_\alpha(\mathcal{M})$. We derive an estimate for this number with the following procedure: For each projected sample point $\vec{x}_i$, we calculate the number of nearest neighbors N_i within a distance $r_\alpha(\mathcal{M})/2$. This can be rather efficiently done using standard clustering algorithms. Now, we assign a weight to point $\vec{x}_i$, which is simply given by $w_i = 1/N_i$. It hence corresponds to the 'fractional contribution' of a single parameter point to a confidence region. The number of distinguishable signals is given by the sum

$$\nu^\alpha_{\mathcal{M},X}(\Omega_\mathcal{M}) = c_{ff} \sum_i w_i \, . \tag{5.4}$$

Here, c_{ff} is a filling factor correction related to the packing density of hyperspheres in a d-dim parameter space. For the 3-dim models in which we are often interested, this number is given by $c_{ff} = 0.74$ [224]. We find that this prescription provides an efficient and reliable way to estimate the number of distinct signals of a model. The main requirement in the calculation is that each of the potential confidence regions contains a sufficiently large (typically at least ten) number of samples. Adding more points to the original list would then not affect the result anymore. We tested the stability of our results by doubling the number of sampled points in various examples from the text. The results remained unchanged in the limit of ten points per distinct region.

Distinct signals compatible with $\mathcal{H}_0$. We are interested in the fraction of the distinct signal points that are consistent with a null hypothesis that is defined as a lower dimensional subspace of the full model parameter space, $\Omega_\mathcal{S} \in \Omega_\mathcal{M}$. Here, we call a point in $\Omega_\mathcal{M}$ 'consistent' with $\Omega_\mathcal{S}$ if the composite null hypothesis $\Omega_\mathcal{S}$ cannot be excluded against the alternative hypothesis $\Omega_\mathcal{M}$. To estimate this number, we first generate a large number of points in $\Omega_\mathcal{S}$. We then collect all points from the original sample of Ω_M whose minimum distance to any of the points from $\Omega_\mathcal{S}$ is smaller than the threshold values $r_\alpha(\mathcal{M}, \mathcal{S})$. Here, the threshold is derived from a χ^2_k distribution with k degrees of freedom, where k is the difference in the dimensionality of $\Omega_\mathcal{M}$ and

Ω_S (for the examples in the paper, we usually have $k = 1$, and hence $r_\alpha(\mathcal{M}, \mathcal{S}) = 2$). The number of distinct signals that are compatible with the null hypothesis $\Omega_\mathcal{S}$ is then simply obtained by restricting the sum in Eq. 5.4 to the points within the shell around $\Omega_\mathcal{S}$.

Parameter ranges and nuisance parameters. Finally, the contours in Fig. 5.3 and Fig. 5.4 are generated by identifying all points that are consistent with the indicated null hypotheses, as described in the previous paragraph. However, in these figures we also take into account the effects of DM halo uncertainties, as described in the main text (this is not easily possible when calculating the number of dimensions). To this end, we generate for each point $\vec{x}_i \in \Omega_\mathcal{M}$ several euclideanized signals corresponding to various randomly selected DM halo configurations. Again, the specific distribution of these points does not matter as long as the number is large enough to sufficiently cover the various halo configurations. In order to incorporate external constraints on the DM halo parameters, we add an additional contribution to the euclideanized distance calculation, which is just given by $(\eta_i - \bar{\eta})^2/\sigma_\eta^2$, where $\bar{\eta}$ and σ_η^2 are the mean and variance of the nuisance parameter, and η_i is the value of the nuisance parameter for a specific point i. The contribution to the $\vec{x}(\vec{\theta})$ is a simple concatenation of $(\eta_i - \bar{\eta})/\sigma_\eta$ with the Euclideanised signal.

For a large number of sampled points this approach exactly matches a profile log-likelihood analysis. We check this limit is saturated by increasing the number of sampled points until our results do not noticeably change.

5.8.2 Dark matter signal modeling

Halo Uncertainties. We incorporate halo uncertainties [214] by assuming Gaussian likelihood distributions for three parameters of the Maxwellian velocity distribution of DM: the Sun's speed $v_\mathrm{e} = (242 \pm 10)\,\mathrm{km/s}$ [225], the local circular speed $v_\mathrm{c} = (220 \pm 18)\,\mathrm{km/s}$ [226], and the Galactic escape speed $v_\mathrm{esc} = (533 \pm 54)\,\mathrm{km/s}$ [227]. We assume that these uncertainties are uncorrelated [228], though in general correlations coming from the modeling of the Milky Way halo can be included [229, 230]. We sample from these distributions as nuisance parameters in our signal calculation and include an additional penalization term to the euclideanized signal in Eq. 5.1.

The Maxwellian velocity distirbution comes from the framework of the Standard Halo Model (SHM) [214, 231]. A steady state collection of collionsionless particles can be modelled through the Boltzmann equation, from which their phase space distribution can be calculated. In the Earth's rest frame this is given by,

$$f(\mathbf{v}) = \frac{1}{(2\pi\sigma_v^2)^{3/2}\, N_\mathrm{esc}} \exp\left(-\frac{(\mathbf{v}+\mathbf{v}_\mathrm{e})^2}{2\sigma_v^2}\right) \Theta\left(v_\mathrm{esc} - |\mathbf{v}+\mathbf{v}_\mathrm{e}|\right) , \tag{5.5}$$

where

$$N_{\rm esc} = \text{erf}\left(\frac{v_{\rm esc}}{\sqrt{2}\sigma_v}\right) - \sqrt{\frac{2}{\pi}}\frac{v_{\rm esc}}{\sigma_v}\exp\left(-\frac{v_{\rm esc}^2}{2\sigma_v^2}\right) . \tag{5.6}$$

Here, we defined $\sigma_v = v_c/\sqrt{2}$, $\mathbf{v}_e$ is the Earth's velocity, and Θ is the heavyside step function. By integrating over all directions, the final speed distribution is given by

$$f(v) = \frac{v}{\sqrt{2\pi}\sigma_v v_e N_{\rm esc}} \begin{cases} 0 & \text{if } v > v_{\rm esc} + v_e\,, \\ \exp[-\frac{(v-v_e)^2}{2\sigma_v^2}] - \exp[-\frac{(v+v_e)^2}{2\sigma_v^2}] & \text{if } v < v_{\rm esc} - v_e\,, \\ \exp[-\frac{(v-v_e)^2}{2\sigma_v^2}] - \exp[-\frac{v_{\rm esc}^2}{2\sigma_v^2}] & \text{otherwise.} \end{cases} \tag{5.7}$$

This satisfies the constraint $\int f(v)\mathrm{d}v = 1$. The speed distribution can be seen in Fig. 5.5, where each panel shows how the different errors effect the distribution. The thick blue line in each panel shows the best fit value.

Detector specifications. We implement a simplified XENON1T for which we adopt an S1-only analysis, full details of which are given in Sec. IIIB of [6]. For the recoil spectrum $\mathrm{d}R/\mathrm{d}S_1$, we use 19 bins linearly spaced between 3 and 70 PE (corresponding to nuclear recoil energies $E_R \in [5, 40]\,\text{keV}$). The number of bins was chosen for computational efficiency with no noticeable loss in accuracy. We have checked that including a 20% energy resolution [232] and increasing the number of bins should have no substantial effect on our results. Background distributions as a function of S_1 are described in Fig. 3 of [172] for which we assume 1% uncertainty on all components separately. We also sum over different Xenon isotopes, weighting by their naturally-occurring mass fractions [233].

For our future Xenon detector we scale up the observation time of XENON1T-2017 by a factor of 100, assuming that background rates stay constant. This exposure roughly corresponds to that expected for the full run of the XENONnT experiment [217], so we will refer to this future detector as *XENONnT*.

DarkSide20k: We directly use the recoil energy spectrum $\mathrm{d}R/\mathrm{d}E_R$ as our signal, assuming that the only relevant isotope is Argon-40. We follow the specifications of the Darkside50 detector [234], with the nuclear recoil efficiency taken from Fig. 6 in [234]. The background is assumed to be flat across the entire energy range with an expected 0.1 events (with 10% uncertainty) over 1422 kg days of observation. We assume 19 linearly spaced bins between 32 and 200 keV.

For our future detector, we assume an exposure of 7.3×10^6 kg days, corresponding to a 1-year exposure with a 20-tonne detector. We assume that the background will remain at 0.1 events in the total exposure. This detector configuration roughly resembles *DarkSide20k* [137].

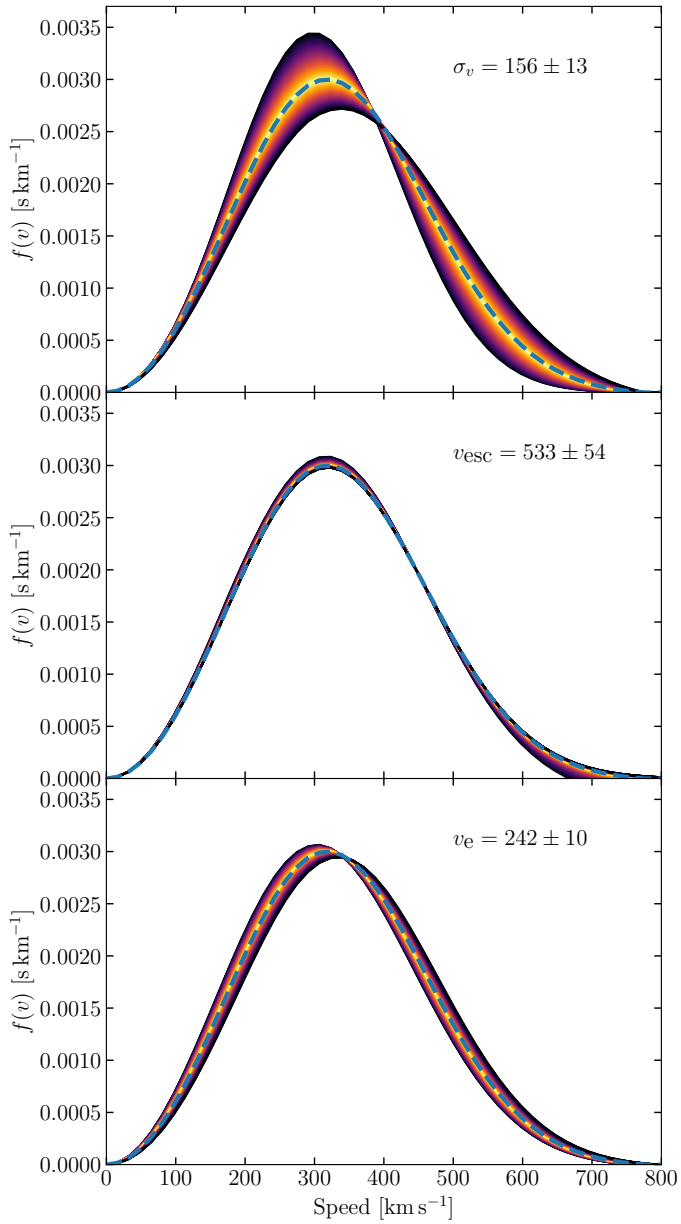

Figure 5.5: **Standard Halo Model Velocity Distribution of DM.** The three panels show the uncertainty on the different parameters of the SHM (as indicated within the panel). The blue dashed line represents the central value for each parameter. For each panel the other two parameters are kept at their central values. These variations show the possible distortions to the SHM accounted for in the modelling uncertainties, as explained in the text.

6 Dark Matter Discovery Potential of Paleo-Detectors

This chapter is based on work from Ref. [3]

6.1 Introduction

Direct detection experiments search for Weakly Interacting Massive Particle (WIMP) Dark Matter (DM) by looking for rare, low-energy nuclear recoils induced by the scattering of DM particles off a target's nuclei [235, 170, 171, 236]. As yet, no conclusive evidence for DM has been reported by direct detection experiments. The leading upper limits on DM–nucleus interactions for WIMPs with masses greater than $\sim 10\,\mathrm{GeV}$ stem from ton-scale liquid noble gas experiments [136, 20, 237, 18, 238]. Meanwhile, cryogenic bolometric detectors have started to probe DM–nucleus interactions for lighter DM candidates [239, 240, 241, 242, 243]. The noteworthy exception is DAMA, which has been reporting evidence [244, 245, 246] for an annually modulated signal [171, 247] compatible with WIMP DM [248, 249, 250, 251, 252, 253, 254, 255, 256] for more than a decade, although this claim is in tension with upper limits from other direct detection experiments [252, 257, 258, 259, 260].

Progress in the search for DM heavier than $\sim 10\,\mathrm{GeV}$ relies on maximizing the exposure (the product of target mass and integration time) of the experiment. Instead, to probe ever lighter DM, experiments must achieve sensitivity to lower and lower nuclear recoil energies. In addition, both mass regimes require exquisite control of a variety of possible background sources, from cosmic rays to intrinsic radioactivity. A number of experiments with lower energy thresholds, larger exposures or, ideally, both will continue the search for lighter and more weakly-interacting DM in the next 5–10 years [261, 262, 216, 217, 137, 218].

Recently, Refs. [263, 264] proposed *paleo-detectors* for the direct detection of DM[1].

[1] For brevity, we use the term 'Dark Matter' (DM) in this work, but will be considering specifically WIMP DM throughout.

In paleo-detectors, one would examine ancient minerals extracted from $\mathcal{O}(10)\,\mathrm{km}$ below the surface of the Earth for traces of DM interactions with atomic nuclei. In lieu of the multi-ton target masses used in conventional direct detection experiments, paleo-detectors take advantage of the fact that DM may have been interacting with the target material for as long as $\sim 1\,\mathrm{Gyr}$. In certain minerals, including those long used as Solid State Track Detectors (SSTDs) [265, 266, 267, 268], DM-induced nuclear recoils would give rise to $1 - 500\,\mathrm{nm}$ long damage tracks. In many materials, such damage tracks would be preserved over timescales much longer than 1 Gyr. Paleo-detectors could thus obtain much larger exposures than conventional detectors even if only a small amount of target material can be investigated. For example, reading out 100 g of material which has been recording DM-induced events for 1 Gyr provides an exposure which could only be matched in the laboratory by observing 10^4 tons of target mass for 10 yr. Further, the relatively small target masses required for paleo-detectors can be obtained from depths much greater than those of underground laboratories in which conventional direct detection experiments are usually operated, providing an unprecedented level of shielding from cosmic rays.

In conventional direct detection experiments, nuclear recoils are detected by scintillation, ionization, and phonon signals in the detector [134]. In paleo-detectors, the observational signature would be nano-scale defects in the minerals. These may be observed through a variety of read-out methods such as X-ray microscopy [269, 270, 271] or ion-beam microscopy [272]. Mineral-based searches, initially for monopoles and then for DM, have been proposed and performed before [273, 274, 275, 276, 277, 278, 279, 280]. However, modern high-resolution imaging techniques [281, 282, 283, 284, 285, 286, 287] as well as the availability of rocks from deeper underground may significantly improve the prospects for DM detection. In particular, measurements of nanometre-length tracks could provide sensitivity to recoil energies as low as 100 eV. A more detailed study of backgrounds, target minerals and read-out methods for paleo-detectors was recently presented in Ref. [264].

In this work, we explore the prospects for excluding or discovering DM with paleo-detectors. In Refs. [263, 264], a simple cut-and-count analysis with a sliding signal window in track length was employed to estimate the sensitivity of paleo-detectors. Here, we adopt more sophisticated statistical techniques, making use of information contained in the spectral shape of the track length distributions. Using realistic distributions for backgrounds induced by neutrinos [288] and radioactivity (developed in Refs. [263, 264]), this approach allows us to explore the impact of different systematic background uncertainties on projected sensitivities. Further, we examine how well the DM properties (mass and cross section) could be reconstructed by paleo-detectors in the case of a discovery.

As in Refs. [263, 264] we consider two fiducial read-out scenarios. The first assumes a high track-length resolution but a relatively small exposure (referred to as

high resolution in the following). The second assumes worse track-length resolution, which should in principle allow more material to be analyzed, facilitating a larger exposure (referred to as *high exposure*). The high resolution configuration is particularly well suited to probing DM with masses below $\sim 10\,\mathrm{GeV}$ while the high exposure configuration is geared more towards heavier DM.

A wide range of minerals are well suited to be paleo-detectors. As described in the discussion of mineral optimization in Ref. [264], minerals can be broadly divided into classes suitable for different applications of paleo-detectors, based on their chemical composition. We consider 4 different minerals in this work, chosen to represent paleo-detectors suitable for probing spin-independent DM–nucleus interactions:

- halite - NaCl,
- olivine - $\mathrm{Mg_{1.6}Fe^{2+}_{0.4}(SiO_4)}$,
- sinjarite - $\mathrm{CaCl_2{\cdot}2\,(H_2O)}$,
- nchwaningite - $\mathrm{Mn^{2+}_2SiO_3(OH)_2{\cdot}(H_2O)}$.

These particular target materials are also selected for their low levels of natural radioactive contamination, helping to suppress radioactivity-induced backgrounds. Average uranium concentrations in the Earth's crust are a few parts per million (ppm) by weight, which would lead to unacceptably high levels of background due to radioactivity. Minerals formed in *ultra-basic rock* deposits, stemming from the Earth's mantle, are much more radiopure. Examples of such minerals investigated here are olivine and nchwaningite, for which we assume uranium concentrations of 0.1 parts per billion. Even less contaminated by radioactive elements may be minerals in *marine evaporite* deposits formed at the bottom of evaporating oceans. We assume uranium concentrations of 0.01 parts per billion for such minerals and use halite and sinjarite as examples. Halite and olivine are very common minerals. In contrast, sinjarite and nchwaningite are less abundant but contain hydrogen. Although hydrogen makes up only a small fraction of these minerals by mass, its presence plays an important role in reducing neutron-induced backgrounds, as we discuss later.

We note that Refs. [263, 264] studied both halite and olivine, allowing a straightforward comparison with our results. Nchwaningite was also used in [264] and our results are comparable to nickelbischofite $[\mathrm{NiCl_2{\cdot}6\,(H_2O)}]$, used in [263], due to its similar chemical composition. Sinjarite gives similar results to epsomite $[\mathrm{Mg(SO_4){\cdot}7\,(H_2O)}]$ used in [263, 264].

The rest of this paper is organized as follows. In Sec. 6.2, we discuss paleo-detectors in more detail, including the calculation of signal spectra, track lengths, and the most relevant background sources. In Sec. 6.3, we present the projected upper limits and discovery reach for the minerals listed above. In Sec. 6.4, we use

benchmark-free techniques to determine contours in the (DM mass)–(cross section) plane which could be reconstructed with paleo-detectors in the hypothetical case of a future discovery. We discuss a number of challenges for paleo-detectors in Sec. 6.5. Our conclusions are presented in Sec. 6.6.

Code for performing all calculations presented in this paper is publicly available here [289].

6.2 Theory

6.2.1 Signal from WIMP Scattering

For elastic scattering of DM with mass m_χ off nuclei with mass m_N, the differential recoil rate as a function of recoil energy E_R per unit target mass is given by [192]:

$$\frac{\mathrm{d}R}{\mathrm{d}E_R} = \frac{\rho_\chi}{m_N m_\chi} \int_{v_{\min}}^{\infty} v f(\mathbf{v}) \frac{\mathrm{d}\sigma}{\mathrm{d}E_R} \, \mathrm{d}^3\mathbf{v} \,. \tag{6.1}$$

The integral is over DM velocities $\mathbf{v}$, with $v = |\mathbf{v}|$ and $v_{\min} = \sqrt{m_N E_R / 2\mu_{\chi N}^2}$. We assume standard spin-independent (SI) interactions, with equal couplings to protons and neutrons. In this case, the differential cross section can be re-written in terms of the DM–nucleon cross section at zero momentum transfer σ_n^{SI} as:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_R} = \sigma_n^{\mathrm{SI}} \frac{m_N}{2\mu_{\chi p}^2 v^2} A^2 F^2(E_R) \,. \tag{6.2}$$

Here, $\mu_{\chi N} \equiv m_\chi m_N / (m_\chi + m_N)$ is the reduced mass of the DM–nucleus system (and similarly for the DM-proton reduced mass $\mu_{\chi p}$). The factor of A^2 corresponds to the coherent enhancement for a nucleus composed of A nucleons. The internal structure of the nucleus is encoded in the form factor $F^2(E_R)$, for which we assume the Helm parametrization [290, 219, 291]. The differential recoil rate then takes the standard form:

$$\frac{\mathrm{d}R}{\mathrm{d}E_R} = \frac{\rho_\chi \sigma_n^{\mathrm{SI}}}{2 m_\chi \mu_{\chi p}^2} A^2 F^2(E_R) \int_{v_{\min}}^{\infty} \frac{f(\mathbf{v})}{v} \, \mathrm{d}^3\mathbf{v} \,. \tag{6.3}$$

For the DM distribution, we assume the Standard Halo Model (SHM), fixing a benchmark value for the local density of $\rho_\chi = 0.3\,\mathrm{GeV\,cm^{-3}}$ [231, 214] in order to compare directly with other direct detection experiments. However, we note that observational estimates of ρ_χ vary substantially [72]. In the SHM, the DM velocity distribution $f(\mathbf{v})$ follows a truncated Maxwell-Boltzmann distribution, for which we

fix values for the Sun's speed $v_\odot = 248\,\mathrm{km/s}$ [225], the local circular speed $v_c = 235\,\mathrm{km/s}$ [226], and the Galactic escape speed $v_{esc} = 550\,\mathrm{km/s}$ [227]. We do not consider here uncertainties on the speed distribution [214] or more recently suggested refinements to the SHM [292].

Recoil spectra for DM and neutrino scattering are calculated using the publicly available `WIMpy` code [212].

6.2.2 Paleo-Detector Rates

The rate of tracks produced with length x_T per unit target mass is given by:

$$\frac{\mathrm{d}R}{\mathrm{d}x_T} = \sum_i^{\text{nuclei}} \xi_i \frac{\mathrm{d}R_i}{\mathrm{d}E_R} \left(\frac{\mathrm{d}E_R}{\mathrm{d}x_T}\right)_i , \tag{6.4}$$

where the index i runs over the different target nuclei which make up the mineral. The rate of nuclear recoils (with initial energy E_R) per unit mass is given by $\mathrm{d}R_i/\mathrm{d}E_R$ and we weight by the mass fraction ξ_i of each nucleus i. The track length as a function of initial recoil energy is calculated as:

$$x_T(E_R) = \int_0^{E_R} \left|\frac{\mathrm{d}E}{\mathrm{d}x_T}\right|^{-1} \mathrm{d}E\,. \tag{6.5}$$

The stopping power $\mathrm{d}E/\mathrm{d}x_T$ as a function of energy must be calculated for each of the recoiling nuclei in a given target material. We use the publicly available SRIM package (**S**topping and **R**ange of **I**ons in **M**atter) [293, 294] to tabulate the stopping power, although analytic estimates are also possible [295, 264]. A more detailed discussion of the calculation of track lengths can be found in Ref. [264].

In this paper we assume that recoiling hydrogen nuclei and α-particles do not produce tracks which can be reconstructed. Whether such low-Z tracks are observable will depend on the target material and read-out method and is a question which must be determined empirically. A discussion of this issue as well as a comparison of results with and without low-Z tracks can be found in Ref. [264].

The resolution at which track lengths can be measured depends on the read-out technique used. For a true track length of x', we assume that the measured track length x is Gaussian distributed[2] with track length resolution σ_{x_T}:

$$P(x|x') = \frac{1}{\sqrt{2\pi\sigma_{x_T}^2}} \exp\left(-\frac{(x-x')^2}{2\sigma_{x_T}^2}\right) . \tag{6.6}$$

[2]This assumption will depend on the imaging technique used and in practice it may be necessary to quantify the probability distribution of the measured track length experimentally.

The number of tracks with lengths in the range $[x_a,\, x_b]$ is then:

$$N(x_a, x_b) = \epsilon \int_0^\infty W(x_T; x_a, x_b)\, \frac{\mathrm{d}R}{\mathrm{d}x_T}(x_T)\, \mathrm{d}x_T\,, \tag{6.7}$$

where ϵ is the exposure, given by the product of the age of the mineral and the total mass of the sample analyzed. The window function W captures resolution effects and is given by:

$$W(x_T; x_a, x_b) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{x_T - x_a}{\sqrt{2}\sigma_{x_T}}\right) - \mathrm{erf}\left(\frac{x_T - x_b}{\sqrt{2}\sigma_{x_T}}\right)\right]. \tag{6.8}$$

We assume that the smallest measurable track length is $\sigma_{x_T}/2$ and consider tracks as long as 1000 nm. As we will see in Sec. 6.2.5, DM-induced recoils do not typically induce tracks longer than this.

We will consider two scenarios for the analysis of paleo-detectors, as in Ref. [264]:

- *High resolution* - we assume a track length resolution of $\sigma_{x_T} = 1\,\mathrm{nm}$, which may be achievable with helium ion beam microscopy [272] (using a focused ion beam [296, 285] and/or pulsed lasers [297, 298, 299] to remove layers of material which have already been imaged). In this case, we assume an exposure of $\epsilon = 0.01\,\mathrm{kg\,Myr}$, which would correspond to a few $\mathcal{O}(1\,\mathrm{mm}^3)$ mineral samples, each with an age of 1 Gyr.

- *High exposure* - we assume a track length resolution of $\sigma_{x_T} = 15\,\mathrm{nm}$. Small angle X-ray scattering has been demonstrated to achieve such spatial resultions in three dimensions [286, 287, 271]. However, such resolutions have not yet been demonstrated when imaging damage tracks arising from nuclear recoils. Here we assume an exposure of $\epsilon = 100\,\mathrm{kg\,Myr}$, corresponding to the analysis of larger samples of $\mathcal{O}(10\,\mathrm{cm}^3)$.

This is not an exhaustive list of possible scenarios, see Ref. [264] for a discussion of a variety of read-out techniques.

We note that a number of techniques (both stratigraphic and radiometric) are used for dating rock samples [300]. Perhaps most relevant for $\mathcal{O}$(Gyr)-aged rocks is fission track dating, which should allow an age estimate which is accurate to $\sim 10\%$ [301, 302], though we will neglect dating uncertainties in our analysis.

Given the target materials we analyze, we note that DM candidates with $m_\chi \lesssim 500\,\mathrm{MeV}$ do not give rise to a significant number of recoil tracks longer than $\sim 1\,\mathrm{nm}$, the best track length resolution assumed in our analysis. Thus, we only consider DM with $m_\chi \gtrsim 500\,\mathrm{MeV}$, though it may be possible to probe lower mass DM with either

better track length resolution or with target materials which allow for longer tracks.

6.2.3 Backgrounds

Here, we summarize the most problematic backgrounds for DM searches with ancient minerals. While follow-up studies and direct calibration will likely lead to refined background modeling, we expect the estimates presented here to be representative. A more detailed discussion can be found in Ref. [264].

We note that cosmogenic backgrounds should be negligible for materials obtained from a depth below $\sim 5\,\mathrm{km}$, meaning that the dominant backgrounds will be neutrino interactions and intrinsic radioactive backgrounds in the target materials themselves.

Neutrinos — Being weakly interacting particles, neutrinos represent an irreducible background to (non-directional) DM searches [303, 304, 305, 306]. Neutrinos with MeV-energies can produce keV-scale nuclear recoils, thereby mimicking DM signatures. We calculate the neutrino-nucleus scattering rate following Ref. [306] (and references therein). In our analysis, we include solar, atmospheric and diffuse supernova background (DSNB) neutrino fluxes, with spectra and normalizations as compiled in Ref. [288]. Uncertainties on the present day neutrino fluxes vary substantially, from $\mathcal{O}(1\,\%)$ for ^{8}B and hep neutrinos [307] to as much as $50\,\%$ for DSNB neutrinos [308]. Paleo-detectors probe neutrino fluxes over the past $\mathcal{O}(1)\,\mathrm{Gyr}$, which may differ from the current values. We therefore conservatively assume a Gaussian systematic uncertainty of $100\,\%$ on the normalization of each neutrino component independently[3], including each of the components from different nuclear processes in the Sun.

Backgrounds from α-decays — One possible background is from the 'uranium series' of uranium-238, a decay chain which proceeds via a series of α and β decays. With each α-decay in the series, the child nucleus recoils against the α particle with $\mathcal{O}(10-100)\,\mathrm{keV}$ energy. The half-life of ^{238}U is $T_{1/2} \sim 4.5 \times 10^9\,\mathrm{yr}$, while the subsequent decays occur much more quickly ($T_{1/2} \lesssim 2.5 \times 10^5\,\mathrm{yr}$). Thus, the vast majority of ^{238}U nuclei which decay will have completed the entire decay chain over the age of the mineral (see Ref. [264] for a more in-depth discussion). Even if the α tracks are not observable, the numerous decays in the chain will lead to a characteristic pattern of tracks. We assume that all such track arrangements can be rejected as background. However, we note that in a $10\,\mathrm{mg}$ sample of sinjarite there would be $\mathcal{O}(10^7)$ completed decay chains and further work is required to estimate whether such large rejection factors will be achievable in a real experiment.

[3] In principle, this allows negative normalization for the background, but in practice the data is constraining enough to exclude this situation. The result is that the normalizations of the neutrino fluxes are effectively free.

A more problematic background comes from uranium-238 nuclei which have only undergone a single α decay ($^{238}\mathrm{U} \to {}^{234}\mathrm{Th} + \alpha$). In this case, the thorium-234 child has a characteristic recoil energy of 72 keV and (assuming that the α track is not seen) is indistinguishable from a DM-induced recoil. The number of such 1α-thorium tracks depends on the relative half-lives of the $^{238}\mathrm{U}$ and $^{234}\mathrm{U}$ decays [($^{234}\mathrm{U} \to {}^{230}\mathrm{Th} + \alpha$) is the second α-decay in the uranium-238 decay chain] and is roughly

$$n_{\mathrm{Th}} \approx 10^9\,\mathrm{kg}^{-1} \left(\frac{C^{238}}{0.01\,\mathrm{ppb}} \right) , \tag{6.9}$$

where C^{238} is the uranium-238 concentration (by weight).

It should be possible to estimate the normalization of radioactive backgrounds to a high precision, for example by measuring the number of full $^{238}\mathrm{U}$ decay chains in the sample. We therefore assume a 1 % systematic uncertainty on the normalization of the 1α-thorium background, though as we will see in Sec. 6.3, the projected sensitivity of paleo-detectors is limited more by uncertainties in background shapes.

Neutron-Induced Backgrounds — Fast neutrons produced in or around the target minerals will scatter elastically with nuclei. Such neutrons have a mean free path of a few centimetres and typically give rise to $10 - 100$ nuclear recoil events with energies comparable to those caused by DM. This large number of neutron-induced tracks is therefore difficult to reject as background.

Fast neutrons may be produced in the spontaneous fission (SF) of $^{238}\mathrm{U}$. This accounts for roughly 1 in every 2×10^6 decays of $^{238}\mathrm{U}$, producing ~ 2 fast neutrons with MeV energies per SF decay. Neutrons may also be produced in (α, n) reactions, in which nuclei absorb an incident α particles and emit fast neutrons. The neutron-induced recoil spectra are estimated using the `SOURCES-4A` code [309], including both SF and (α, n) contributions, as described in Ref. [264]. As in the case of 1α-thorium backgrounds, the normalization of the neutron-induced background scales with uranium-238 contamination. For minerals found in ultra-basic rocks (nchwaningite and olivine), we assume a uranium-238 contamination of $C^{238} = 0.1\,\mathrm{ppb}$ by weight, while for those found in marine evaporites (halite and sinjarite) we assume $C^{238} = 0.01\,\mathrm{ppb}$.

We assume a 1 % systematic uncertainty on the normalization of the neutron-induced backgrounds. We discuss the sensitivity of our results to the assumed background uncertainty and explore more extended shape systematics in Sec. 6.3.

6.2.4 Analysis Theory

To estimate projected upper limits and discovery reaches for paleo-detectors, we use the statistical techniques developed in Ref. [1]. These extend traditional Fisher forecasting methods to the Poisson regime, which allows one to approximate the median result obtained via a Monte-Carlo simulation with minimal computational overhead. We briefly summarize the technique below.

Traditional Fisher forecasting methods are sufficient to accurately calculate expected exclusion limits as well as the discovery reach when in the Gaussian regime. Typically, in direct detection experiments, however, the small number of signal and background events means that the number of expected counts is not well described by a Gaussian distribution. To remain accurate in this Poissonian regime, we adopt the *equivalent counts method*, as developed in Ref. [1]. The basic procedure for the equivalent counts method is to define a mapping between the expected background counts, their associated uncertainty, and the expected signal such that the full profile log-likelihood is approximated by the Poisson log-likelihood ratio. This mapping is given by,

$$-2\ln\frac{\mathcal{L}_{\rm p}(\mathcal{D}_{\mathcal{A}}(\mathbf{S}_0)|\mathbf{S}_0)}{\mathcal{L}_{\rm p}(\mathcal{D}_{\mathcal{A}}(\mathbf{S}_0)|\mathbf{S})} \simeq -2\ln\frac{P(b_{\rm eq}|b_{\rm eq})}{P(b_{\rm eq}|s_{\rm eq}+b_{\rm eq})}\,, \tag{6.10}$$

where $\mathcal{D}_{\mathcal{A}}(\mathbf{S}_0)$ is the Asimov data set [17] given no expected signal events and $\mathbf{S}$ is the expected number of signal events. $P(a|a')$ represents the Poisson probability mass function, *i.e.*, the probability of a events given a' expected events. The *equivalent* signal and background events are denoted by $s_{\rm eq}$ and $b_{\rm eq}$, respectively. They are defined such that the log-likelihood ratio of a simple one-bin Poisson process approximates the full log-likelihood ratio. We found expressions for $s_{\rm eq}$ and $b_{\rm eq}$ in terms of the Fisher matrix of the full problem, which are given in Eq. (6) and Eq. (7) of Ref. [1]. The procedure leads – per definition – to exact results in the limits where the full problem is a one-bin Poisson process or Gaussian, and approximates very well Monte Carlo results in the intermediate range.

The discovery reach and exclusion limit can then, trivially, be calculated by solving

$$-2\ln\frac{P(s_{\rm eq}+b_{\rm eq}|b_{\rm eq})}{P(s_{\rm eq}+b_{\rm eq}|s_{\rm eq}+b_{\rm eq})} = Z^2\,, \tag{6.11}$$

and

$$s_{\rm eq} = Z\sqrt{s_{\rm eq}+b_{\rm eq}}\,, \tag{6.12}$$

respectively, and mapping this back on to the signal parameters of the full model. The significance level α determines the critical value Z, which is derived from the

inverse of the standard normal cumulative distribution, denoted $F_{\mathcal{N}}$, as

$$Z(\alpha) \equiv F_{\mathcal{N}}^{-1}(1-\alpha) \,. \tag{6.13}$$

For example, $Z(0.05) = 1.64$ represents a 95% confidence level. Reference [6] showed that for most cases the equivalent counts method is accurate to the percent level. In some exceptional cases the derived upper limit exhibits a maximum deviation of 40% from a fully coverage corrected Monte Carlo calculation, for more detailed discussion see Ref. [6]. These deviations are most relevant when two distinct regions of the spectrum are present; one in which the signal dominates and the background is in the Poissonian regime, and the other in which the background dominates over the signal. In almost all our cases we expect a significant background so this is unlikely to be a problem.

The *information flux* [1] provides an intuitive illustration of which region of signal space provide most information about the DM-induced signal. It is a generalized signal-to-noise ratio which allows for the inclusion of extended systematics and covariances. Technically, the information flux is obtained by taking functional derivatives of the Fisher information matrix with respect to the exposure at different track lengths. It can be thought of as the rate at which the error bar on the parameter of interest will be reduced from an infinitesimal increase in exposure (for a particular track length). We stress that the information flux is used for illustration only and does not enter directly into the calculation of projected upper limits or the discovery reach.

The statistical techniques outlined above are implemented in the software package `swordfish`[4] [6]. This package provides a straightforward interface, allowing the user to input signal and background spectra, and efficiently computes the projected exclusion limits and discovery reach. Although the forecasting techniques developed in Ref. [1] are applicable to unbinned data, the implementation in `swordfish` requires, for practical reasons, that the data be binned. We use 70 log-spaced bins throughout this work, with the range defined by the resolution as $\sigma_{x_T}/2 \lesssim x_T \lesssim 1000\,\text{nm}$. The number of bins was chosen by incrementally increasing the bin width until the projected constraints on the cross section begin to weaken. This way we use the minimum number of bins required to resolve all features in the track length spectra. We do this both to minimize computation time and to ensure that the systematics study in Sec. 6.3 is conservative (using a small number of bins enhances the impact of bin-to-bin variations in the backgrounds). The same number of bins is used in both the high resolution and high exposure cases to allow for an easier comparison between the two read-out methods.

[4] https://github.com/cweniger/swordfish

6.2.5 Track Length Spectra

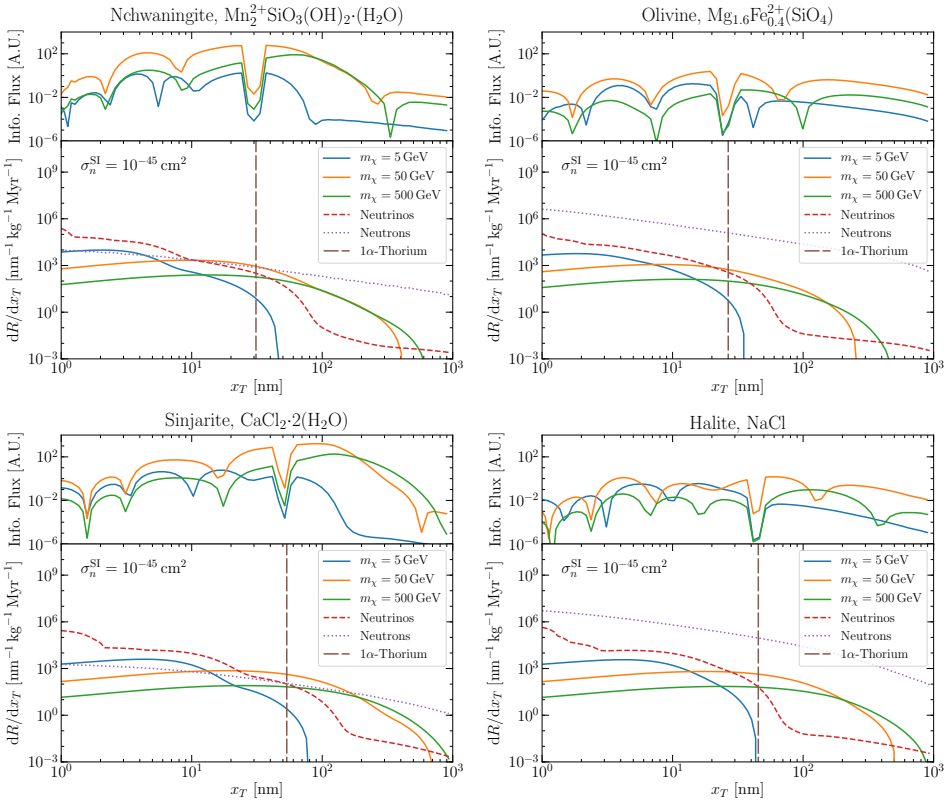

Figure 6.1: **Track-length spectra and information flux for three different WIMP masses.** Each panel shows the spectrum of track lengths x_T expected for a different mineral target. WIMP signals are shown as solid lines, while background distributions are shown as dotted and dashed lines. We fix the signal normalization to $\sigma_n^{\rm SI} = 10^{-45}\,{\rm cm}^2$. The information flux, shown above each set of spectra, is a generalized signal-to-noise ratio discussed in more detail in Sec. 6.2.4. The information flux has been calculated for the high-resolution case, with a track-length resolution of $\sigma_{x_T} = 1\,{\rm nm}$.

We now present the distribution of track lengths expected for DM signals as well as the backgrounds already described in this section. These are shown in Fig. 6.1, with each panel showing a different mineral. We fix the DM–nucleon cross section to $\sigma_n^{\rm SI} = 10^{-45}\,{\rm cm}^2$ and show the DM signal spectra for three different DM masses, $m_\chi = 5$, 50 and 500 GeV, as solid lines. We see that at short track lengths, $x_T \lesssim 10\,{\rm nm}$, where signals of lighter WIMPs would appear, the dominant background often comes from solar neutrinos. At longer track lengths, where the signal from heavier

DM typically peaks, the dominant backgrounds are radioactive: mono-energetic 1α-thorium recoils and neutron-induced recoils.

In the upper part of each panel, we also plot the information flux, as described in Sec. 6.2.4. The most pronounced feature is a sharp drop for track lengths corresponding to 1α-thorium recoils. Little information about a DM signal can be gained by studying tracks of this length, as signal events are here degenerate with 1α-thorium tracks. The various peaks in the information flux indicate the track lengths that provide the most constraining power for the DM signal. These maxima appear either when the signal is large, corresponding to a large signal-to-background ratio, or when one of the backgrounds is prominent. In the latter case, track lengths corresponding to peaks in the information flux allow us to constrain the normalization of a particular background component. This in turn leads to improved constraints on the signal. In between these peaked regions, the information flux is typically suppressed. Note that the detailed shape of the information flux depends significantly on the specific assumptions that are made about the correlation of background systematics.

We discuss the track length spectra, information fluxes and their impact on paleo-detector sensitivity in more detail in the next section.

6.3 Projected Sensitivity

Here we present the results of the sensitivity analysis. First we discuss a simple benchmark case in which we consider systematic errors only on the normalization of the individual background components. We also use this scenario for the mass reconstruction projections in Sec. 6.4. In addition, we also consider bin-to-bin systematics in order to assess the impact of shape uncertainties of the background spectra on the projected sensitivity.

6.3.1 Background Normalization Systematics

The backgrounds described in Sec. 6.2 all have an associated uncertainty which must be accounted for within the analysis framework. For our background normalization systematics scenario we assume that the shapes of the signal and backgrounds are fixed with only a systematic uncertainty on the normalization of each background component. The systematic uncertainties we assign to the normalization of different backgrounds are detailed in Sec. 6.2.3. We ignore covariances between the signal and background and between individual backgrounds components. With careful calibration, we may be able to understand the shape of the background to a high degree of precision. In practice, it should be straightforward to produce target samples with high levels of radioactivity-induced tracks in the laboratory, though such an approach

is more challenging for the neutrino-induced backgrounds.

Uncertainties in the normalization of the backgrounds can be mitigated when a good 'control region' is available, where the signal is sub-dominant and the overall background rate can be well-constrained. This is typically the case for the broad distribution of neutron-induced recoils; even for heavy WIMPs, the signal drops off well below 1000 nm, providing a good control region at large track lengths. Instead, neutrino-induced backgrounds may mimic the DM signal for certain DM masses. If this is the case, no control region is available and limits are severely weakened.

We show in Fig. 6.2 the projected sensitivity for the four minerals we consider in this work. The top two panels show the projected 90% confidence exclusion limits while the bottom two panels show the projected $5\,\sigma$ discovery reach [310], which we define as the line above which the paleo-detector setups we consider would have at least 50 % chance of achieving a $5\,\sigma$ discovery of DM. In gray, we show current bounds from conventional direct detection experiments, coming predominantly from XENON1T [18] and SuperCDMS [240] in this mass and cross section range.

In the left panels, we show the high resolution case, with $\sigma_{x_T} = 1\,\text{nm}$ and an exposure of $\epsilon = 0.01\,\text{kg}\,\text{Myr}$. The 'bump' in the limits at $m_\chi \sim 7\,\text{GeV}$ is due to the WIMP signal spectrum near this mass being mimicked by the spectra of ^{8}B and hep neutrinos. This 'Solar neutrino floor' has been studied in detail in, for example, Refs. [306, 311, 288]. The loss in sensitivity is even more pronounced in our case, owing to the larger systematic uncertainty we have assigned to the Solar neutrino flux. In addition, moving away from $m_\chi \sim 7\,\text{GeV}$, the spectra are no longer degenerate, meaning that control regions are rapidly regained and sensitivity restored. In contrast to conventional direct detection experiments, paleo-detectors have large enough exposures to directly constrain the normalization of the ^{8}B/hep Solar neutrino fluxes.

Even accounting for these low energy background neutrinos, in the left panels, the limits at low mass are substantially stronger than limits from any conventional detector. This is due to the ability of modern imaging techniques to measure small track lengths. Tracks of 1 nm in length typically correspond to recoil energies around 100 eV, giving a threshold comparable to the CRESST-2017 Surface Run [241], albeit with a much larger exposure. Thus, for DM lighter than $\sim 10\,\text{GeV}$, paleo-detectors may probe DM–nucleon cross sections much smaller than the projected sensitivity of conventional direct detection experiments.

For larger DM masses, the sensitivity is severely limited by neutron-induced backgrounds. Nchwaningite and sinjarite contain hydrogen, which is an efficient moderator of fast neutrons. Thus, the rate of neutron-induced backgrounds is much smaller than in olivine and halite, which do not contain hydrogen. In addition, for halite and sinjarite we assume an intrinsic contamination of $C^{238} = 0.01\,\text{ppb}$, whereas for olivine

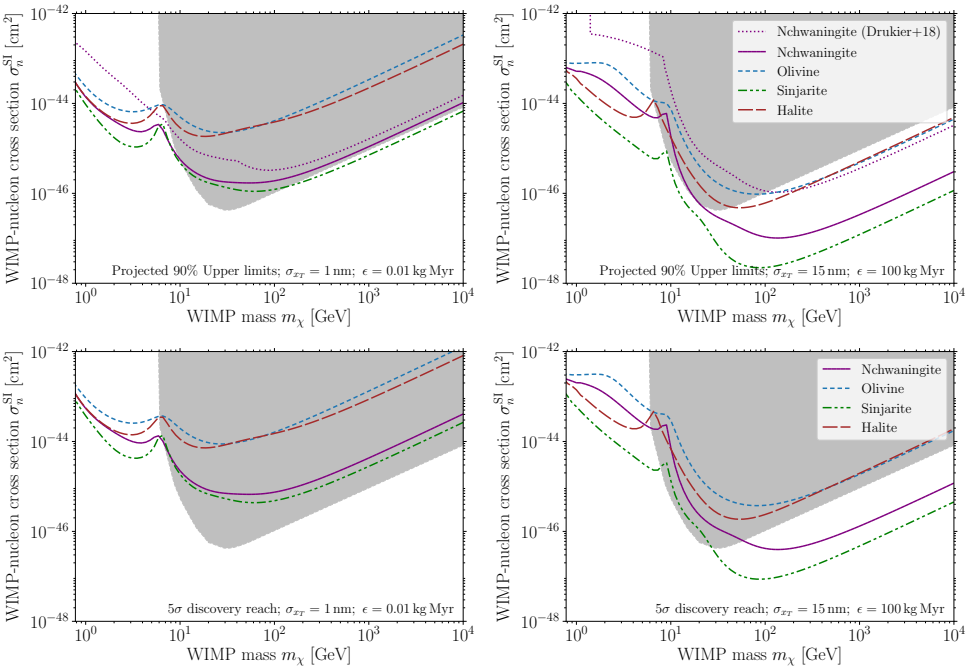

Figure 6.2: **Projected paleo-detector sensitivity using a spectral analysis. Top:** Projected 90% Upper Limits. **Bottom:** $5\,\sigma$ Discovery Reach. In the *left column*, we assume $\sigma_{x_T} = 1\,\text{nm}$; $\epsilon = 0.01\,\text{kg}\,\text{Myr}$ (high resolution case). In the *right column*, we assume $\sigma_{x_T} = 15\,\text{nm}$; $\epsilon = 100\,\text{kg}\,\text{Myr}$ (high exposure case). We assume a 100 % systematic uncertainty on the normalization of each individual neutrino component, as well as a 1 % systematic uncertainty on the normalization of both the neutron background and the 1α-thorium background. *Gray regions* show the parameter space currently excluded by conventional direct detection experiments [243, 18]. In the *upper panels*, we also show the projected limits from nchwaningite using a sliding window analysis (dotted purple), as reported recently in Ref. [264].

and nchwaningite we assume $C^{238} = 0.1\,\text{ppb}$. For nchwaningite and sinjarite, the minerals containing hydrogen, cross sections as low as $10^{-46}\,\text{cm}^2$ could be probed for a DM mass of 50 GeV, assuming an exposure of 0.01 kg Myr, as shown in the left panels. At higher masses, the projected sensitivity of paleo-detectors is comparable to current XENON1T bounds.

In the right panels of Fig. 6.2, we show the projected sensitivity for the lower resolution case, with $\sigma_{x_T} = 15\,\text{nm}$ and a larger exposure of $\epsilon = 100\,\text{kg}\,\text{Myr}$. For DM heavier than 100 GeV, it is possible to probe DM–nucleon cross sections a factor of 30 and 100 smaller than current XENON1T bounds using nchwaningite and sinjarite, respectively. For DM lighter than 10 GeV, the sensitivity is marginally weaker than in

the high resolution case; at low DM mass, the signal spectrum peaks towards shorter track lengths and is thus challenging to resolve with worse resolution.

We note that the projected limits are comparatively weak between 20 and 100 GeV. This is because the peak due to 1α-thorium recoils (vertical dashed lines in Fig. 6.1), typically coincides with the peak in the signal spectra in this mass range[5]. We see in the upper panels of Fig. 6.1 that the broad peak in the information flux for the 50 GeV case typically occurs at the same position as the 'dip' caused by the 1α-thorium background.

We now compare our results with those obtained using a simpler 'sliding window' analysis in Refs. [263, 264], where a cut-and-count analysis is performed over a window chosen to optimize the signal-to-noise ratio. In the top two panels of Fig. 6.2, we show the limit obtained in Ref. [264] for nchwaningite.

In the high resolution case (left panels of Fig. 6.2), we find that the 'sliding window' analysis is only marginally less sensitive at the highest masses. Near $m_\chi \sim 50\,\text{GeV}$, the full spectral analysis is more sensitive by a factor of a few due to better rejection of the sharply peaked 1α-thorium background. Going to lower masses, the 'bump' corresponding to the ^{8}B/hep Solar neutrinos becomes more pronounced in a full spectral analysis. For lighter DM, the neutrino- and DM-induced spectra become distinguishable again. The spectral analysis can effectively reduce the error on the normalization of the neutrino-induced background by using information from track lengths where neutrino-induced events dominate. For example, at masses of 1 GeV this allows the projected sensitivity from the spectral analysis to improve nearly an order of magnitude with respect to that found in the 'sliding window' analysis.

In the high exposure case (right panels of Fig. 6.2), the spectral analysis gains roughly an order of magnitude in sensitivity with respect to the 'sliding window' analysis for DM heavier than $\sim 100\,\text{GeV}$. The lower resolution makes it more difficult to exploit subtle differences in the shape of the signal and background spectra. However, at the longest track lengths considered, the tracks induced by neutrons always dominate over those induced by DM, cf. Fig. 6.1. Because of the larger exposure, this 'control region' has sufficient statistics to tightly constrain the normalization of the neutron-induced background. Thus, the sensitivity to higher mass WIMPs is improved with respect to the 'sliding window' analysis.

[5]For $m_\chi \sim 50\,\text{GeV}$, the typical recoil energy for nuclei in the minerals we consider is $\sim 20\,\text{keV}$, much smaller than the 72 keV 1α-thorium recoil. However, the stopping power for lighter nuclei such as Ca, Cl and Na is smaller, leading to similar track lengths.

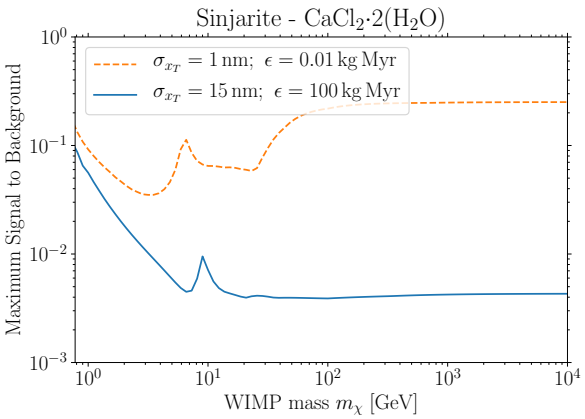

Figure 6.3: **Maximum signal-to-background ratio at the 90 % confidence limit.** Both lines show the maximum signal-to-background ratio (over all track lengths) for sinjarite with σ_n^{SI} set to the projected 90 % confidence exclusion limit at each mass. The *orange (dashed) line* shows the high-resolution case. Here we see the signal-to-background is relatively constant around $10-30\,\%$. The *blue (solid) line* shows the high-exposure case. At low masses the maximum signal-to-background is roughly 10 % whereas at high masses this is reduced to $0.4-0.5\,\%$. This is reflected by the increased sensitivity of the limit to bin-to-bin background shape uncertainties, as shown in Fig. 6.4.

6.3.2 Background Shape Systematics

As discussed in the previous section, our sensitivity to DM at high masses is typically limited by neutron-induced backgrounds. Conversely, neutrinos are the dominant background for low DM mass. As we show in Fig. 6.3, the maximum signal-to-background ratio (as a function of track length) for the high exposure case is typically much smaller than 10 %. This means that the strength of the limits depends in principle on a %-level understanding of the shape of the background distributions. For the high resolution case the maximum signal-to-background is typically closer to 30 %. Therefore, the projected sensitivity should be more robust to shape uncertainties in that case.

In order to explore how the sensitivity of paleo-detectors is affected by such background shape uncertainties, we assign a Gaussian systematic error to the normalization of each bin of each background component. Our analysis therefore allows the backgrounds in individual bins to fluctuate independently. Because we no longer assign systematic uncertainties to the overall normalization of the backgrounds, in some situations (e.g. when the bin-to-bin uncertainty is chosen to be smaller than the normalization uncertainty in Sec. 6.3.1) the projected limit with shape uncertainties may

be stronger than in the normalization systematics case. In such situations, we set the projected limit equal to the normalization systematics case.

In Fig. 6.4, we plot the limits obtained with various bin-to-bin background systematics for a sinjarite paleo-detector. For comparison, we also show the sensitivity obtained in the background normalization systematics case described in the previous sub-section. These results are mostly illustrative since the actual level of shape uncertainties is hard to anticipate *a priori*. However, they indicate the level of uncertainty that can be tolerated in practice.

As expected, allowing some variation in the shape of the backgrounds degrades the sensitivity of paleo-detectors. In the upper panel of Fig. 6.4 we show the high resolution case. The limit is unaffected by bin-to-bin systematics until they are increased to 50 %. The more relevant uncertainty is therefore the overall normalization of the background components.

The high exposure case is shown in the lower panel of Fig. 6.4. We find a much greater dependence on the background shape systematics. For 10 % bin-to-bin systematic uncertainties, the sensitivity is degraded by an order of magnitude compared to the background normalization systematics case for DM heavier than $\sim$ 40 GeV. Another factor of $\sim$ 5 is lost when increasing the bin-to-bin systematics from 10 % to 50 %.

The high resolution case is more robust to background shape systematics primarily because of its larger signal-to-background ratio, as shown in Fig. 6.3. In addition, spectral features in the high resolution case are more pronounced, allowing for easier distinction between signal and background even when there are significant uncertainties in the background shapes.

For comparison with our projections, we also show in Fig. 6.4 the projected 90 % confidence exclusion limits from LUX-Zeplin [218] and SuperCDMS SNOLAB (Ge) [262] (planned for data-taking from 2020 onwards), with respective exposures of 1.5×10^4 kg yr and 44 kg yr. For DM masses below 10 GeV, the high resolution case can improve upon future SuperCDMS SNOLAB constraints by up to one and a half orders of magnitude. For the case of 50 % bin-to-bin shape systematics, sinjarite would still be an order of magnitude more sensitive than SuperCDMS SNOLAB projections at 2 GeV. The high exposure case can achieve the same sensitivity as LUX-Zeplin to higher mass DM only if the background shape uncertainties are kept at the 1 % level.

6.4 Constraining the Dark Matter Mass

In this section, we investigate to what extent the properties of a DM candidate, in particular its mass, could be constrained in the hypothetical case of a DM discovery.

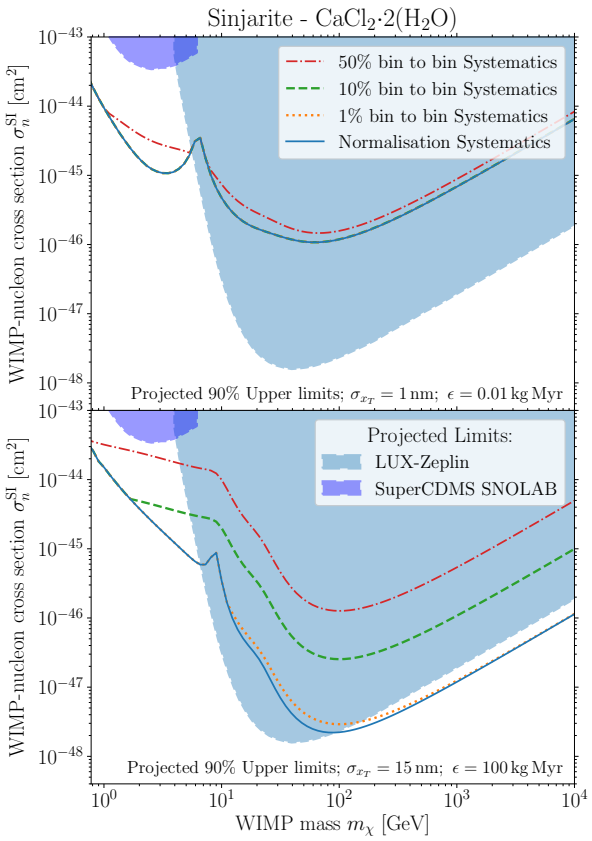

Figure 6.4: **Projected 90 % confidence limits for sinjarite including background shape systematics.** Top and bottom panels show the high resolution and high exposure cases, respectively. *Blue line:* Background normalization systematics case with systematic normalization uncertainties for each background component. The normalization systematic on neutrinos here is set to 100 % whereas for the radioactive backgrounds we assume 1 % normalization error. *Red, Orange, and green lines:* Background shape systematics case where we allow the normalization of each background component in each of the 70 log-spaced track-length bins to fluctuate independently. The red, green, and orange lines show results for 50 %, 10 %, and 1 % systematic uncertainty per bin, respectively. Note that where the bin-to-bin systematics produce a limit stronger than that of the normalization systematics case, we set the projected limit assuming normalization systematics. In the top panel the 1% (orange dotted) and 10 % (green dashed) bin-to-bin systematic lines are therefore not distinguishable from the normalization systematics case (blue solid). Shaded regions show projected limits from LUX-Zeplin [218] and SuperCDMS SNOLAB (Ge) [262].

Thus, we switch from projecting limits, as in Sec. 6.3, to parameter reconstruction. A priori there is no reason for the DM to appear in any particular region of the parameter space. Instead of employing benchmark scenarios as often done in the literature [312, 313, 173, 314], we perform a benchmark-free study using the Euclideanized signal method [6, 2].

The Euclideanized signal method maps points in the model parameter space – here, $(m_\chi, \sigma_n^{\mathrm{SI}})$ – onto points in a 'Euclideanized signal' space, taking into account systematic uncertainties, covariances and nuisance parameters. Likelihood ratios between points in the model space are mapped to Euclidean distances in the 'Euclideanized signal' space. Comparing the Euclidean distances between large numbers of points is computationally fast (using clustering algorithms), allowing us to efficiently map out the reconstruction prospects over a wide range of the parameter space. Full details of the Euclideanized signal method can be found in Refs. [6, 2] and a summary is given in App. 6.7.1.

For regions where a future signal would provide a closed constraint on the DM mass, we calculate the accuracy to which this is possible by defining the fractional uncertainty as

$$\frac{\Delta m_\chi}{m_\chi} = \frac{|m_{\chi,\mathrm{max}} - m_{\chi,\mathrm{min}}|}{m_\chi} \,. \tag{6.14}$$

Here, $m_{\chi,\mathrm{max}}$ and $m_{\chi,\mathrm{min}}$ are the maximum and minimum edges of the two-dimensional $2\,\sigma$ confidence contour around a point with a given $(m_\chi, \sigma_n^{\mathrm{SI}})$. Note that these contours are typically quite asymmetric, usually extending much further towards masses larger than the true mass than towards masses smaller. Thus, a fractional uncertainty $\Delta m_\chi/m_\chi \gtrsim 1$ does not necessarily imply that no information about the DM mass can be obtained. Rather, $\Delta m_\chi/m_\chi \gtrsim 1$ typically implies that $\Delta m_\chi/m_\chi \sim m_{\chi,\mathrm{max}}/m_\chi$, while usually $m_{\chi,\mathrm{min}}$ is not much smaller than m_χ.

In Fig. 6.5, we show the ability of a sinjarite paleo-detector to constrain the DM mass from a future signal. The color scale shows contours in $\Delta m_\chi/m_\chi$; in the following we refer to the colored regions as those where the *mass can be reconstructed*. The gray regions indicate points in the parameter space where the $2\,\sigma$ confidence contours are not closed, i.e. the reconstructed mass would be unbounded from above. Thus, we refer to the gray regions as portions of parameter space where *the mass cannot be reconstructed*. In Fig. 6.5, we quantify how well the mass could be reconstructed for DM–nucleon cross sections between the projected 90 % confidence exclusion limits, cf. Sec. 6.3.1, and cross sections a factor 100 larger than this (the region bounded by the two black curves). Note that some of this region is already ruled out by XENON1T, cf. Fig. 6.2.

Direct detection experiments suffer from an almost exact degeneracy between the mass and cross section at large DM masses [221]. The degeneracy occurs when

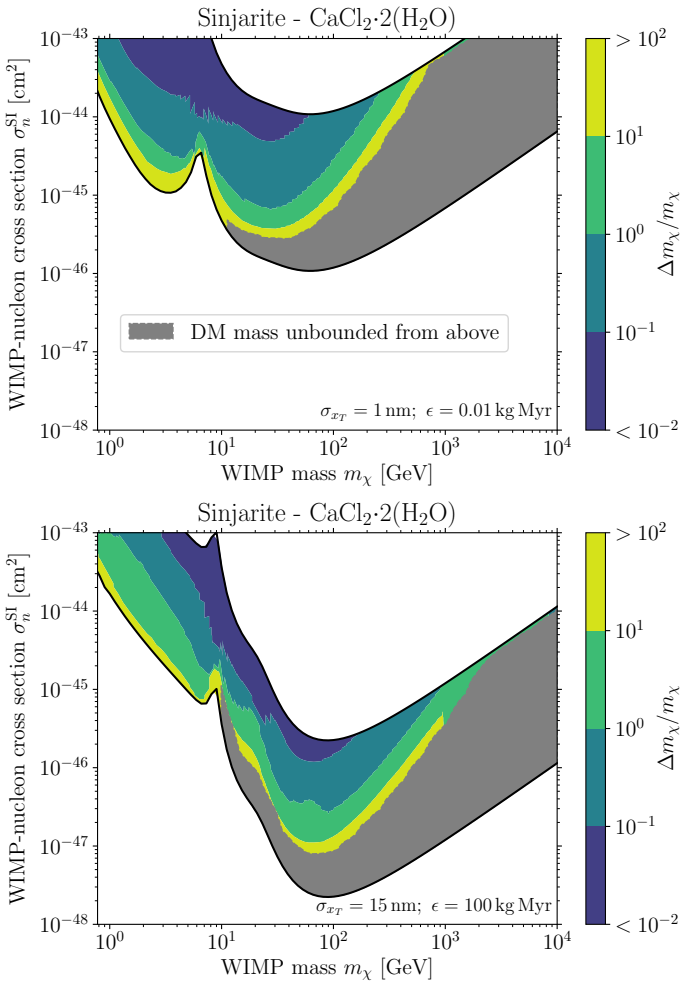

Figure 6.5: **Constraints on the mass of a DM particle from a future signal.** Gray shaded regions correspond to parameter points where the DM mass is unconstrained from above at the $2\,\sigma$-level. The colored contours indicate the fractional uncertainty on the DM mass obtained by constraining a future signal, as defined in Eq. (6.14). Top and bottom panels show the high resolution and high exposure cases, respectively. The lower black lines in both panels correspond to the projected 90 % confidence limit in Fig. 6.2. We consider regions between these lower lines and a factor of 100 larger, indicated by the upper black lines. Note that some of these regions are already excluded by current experiments (see shaded regions in Fig. 6.2).

$m_\chi \gtrsim m_N$. This is because, for a given nuclear recoil energy, v_{min} depends only on the reduced mass of the DM–nucleus system. For $m_\chi \gg m_N$, the reduced mass becomes

independent of the DM mass.

For traditional direct detection experiments, the ability to reconstruct the mass of a hypothetical DM particle has been studied extensively, see e.g. [220, 221, 315, 2, 314]. These studies found that a Xenon experiment would only be able to constrain the DM mass up to $\sim 200\,\mathrm{GeV}$. For paleo-detectors, we instead find that mass reconstruction is possible for DM masses as large as $\sim 1\,\mathrm{TeV}$.

We show results for the high-resolution configuration in the upper panel of Fig. 6.5. In the case of DM–nucleon cross sections just below the current limits, we find that the largest DM mass for which the mass could be reconstructed is $\sim 250\,\mathrm{GeV}$. In the high exposure configuration (Fig. 6.5, lower panel) the DM mass could be reconstructed for masses as high as $\sim 1\,\mathrm{TeV}$ if the DM–nucleon cross section is just below current limits. For cross sections further below current limits, we see that in the high resolution case the colored region extends to slightly larger masses than in the high exposure case.

For a DM candidate with mass below $\sim 10\,\mathrm{GeV}$, the mass could be reconstructed for all cross sections in reach of a sinjarite paleo-detector in either the high resolution or high exposure configuration. Although the precision of the mass reconstruction depreciates with decreasing DM–nucleon cross section, there are large regions of parameter space which are currently unconstrained and where the DM mass could be reconstructed reasonably well by paleo-detectors in case of a discovery.

The potential for paleo-detectors to tightly constraint the mass of a DM candidate stems in part from their large exposure. For example, in the high resolution case, a $6\,\mathrm{GeV}$ DM candidate with cross section at the 90 % confidence exclusion limit would give rise to $\sim 10^5$ signal events. Such large numbers of events would allow us to accurately map out the track length spectrum. For DM masses of $\sim 1\,\mathrm{TeV}$, paleo-detectors would only measure $\mathcal{O}(10^3)$ events at the exclusion limit, making the reconstruction of the associated track length spectrum more challenging. This should be contrasted with exposures possible in conventional direct detection experiments, where at their exclusion limits only $\mathcal{O}$(1-10) events would be detected. Such a low number of signal events would not provide enough information to resolve minute differences in the recoil spectra required for mass reconstruction.

Further, paleo-detectors could probe track lengths over three orders of magnitude, which corresponds to sensitivity spanning a large range of recoil energies. In particular, a $1000\,\mathrm{nm}$ track corresponds roughly to a $\sim 1\,\mathrm{MeV}$ nuclear recoil whereas a $1\,\mathrm{nm}$ track equates to a $\sim 100\,\mathrm{eV}$ recoil[6]. The high energy part of the recoil spectra has a significant dependence on the DM mass [173, 314]. Unlike the traditional energy window of direct detection experiments, we exploit this information by observing a wide variety of track lengths.

[6] This can obviously depend significantly on the recoiling nucleus and target material.

Finally, the target materials we consider here contain a variety of constituent nuclei with different masses. Sinjarite contains nuclei with masses from 1 GeV (H) to 37 GeV (Ca). Since the observed signal is a weighted sum over the contributions from the respective target nuclei, the resulting track length spectrum is richer in features than the simple slope dependence one finds for a single target nucleus. As these features can be exploited efficiently with a spectral analysis, paleo-detectors are particularly well suited for DM mass reconstruction.

We note that we have not considered any uncertainties in the DM velocity distribution in this analysis. For example, the longest tracks (which we leverage to constrain very heavy DM) are produced by the fastest moving DM particles, close to the Galactic escape velocity. The length of these longest tracks is therefore dependent on uncertainties in the escape velocity. More generally, allowing for variations in the DM distribution will widen the constraints on the DM mass. A number of techniques have been developed to incorporate velocity distribution uncertainties in direct detection (see for example Refs. [229, 316, 312, 317, 318, 319, 320]) but we leave this more detailed analysis to future work. We note however that previous studies have shown that using multiple experiments with different target nuclei greatly reduces the impact of these uncertainties [173, 2]. We expect the same to be true for paleo-detectors since the minerals investigated in this work contain a variety of target nuclei.

6.5 Challenges

Throughout this work we have shown that the sensitivity of paleo-detectors may well exceed that of current direct detection experiments. In Sec. 6.3.1, we projected the sensitivity assuming systematic errors on the overall normalization of the different background components only. For the neutrino-induced backgrounds, we assumed 100 % systematics, while we assumed 1 % systematics for backgrounds induced by radioactivity. In order to check the robustness of our results, we increased the normalization systematics on the neutron-induced backgrounds to 5 % and found that the sensitivities are unaffected. In the following, we discuss some of the other potential issues moving forward.

In our background normalization systematics case we assumed no covariance between the normalization of the background components. This assumption should hold for many background components, for example we expect no covariance between the spectra induced by solar neutrinos and diffuse supernova neutrinos. For the radiogenic backgrounds there may exist some covariance since they have a common origin.

For the background shape systematics case, we assume no bin-to-bin covariances. This may be an optimistic or pessimistic assumption depending on the covariances one might expect. The most troublesome scenario would be a bin-to-bin covariance

that makes signal-like variations in the background more likely. Due to the lack of theoretical guidance, we have chosen not to explore bin-to-bin covariances. Instead we attempt to maximize the error introduced by bin-to-bin systematics by using the minimum number of bins required to resolve all features in the spectra. We leave careful study of covariances to future analyses.

One of the primary assumptions throughout our analysis is that we can reject damage features in the minerals that are not tracks arising from nuclear recoils. Further, we do not consider the background produced by a series of linked α-recoil tracks in the uranium-238 decay chain. The assumption is that the characteristic track pattern is easily recognizable and therefore rejected with 100% efficiency. In reality there will be $> \mathcal{O}(10^7)$ tracks within a sample, many of which will exhibit these characteristic track patterns. We therefore require an automated tagging and rejection system. Since the characteristic track pattern is quite distinct from a normal track, it is possible that this system can be very efficient. However, this is yet to be validated for a large data set, a task we leave to future work.

The data produced by scanning significant amounts of material at high precision could present an issue by itself. Naively, scanning 1mm^3 of material at 1nm precision will produce 10^6 terabytes of data. It is a monumental task to analyze such a large data set. Luckily, much of the mineral will contain no track information at all, therefore a suitable compression format can be adopted to make the analysis more tracktable. Analysis of the data will require automated track-recognition, an ideal application of machine learning algorithms. We will also address this in future publications.

Naively, one would expect paleo-detectors to be able to exploit directional information from the orientation of the recoil tracks. However, the target minerals we consider here are $\mathcal{O}(1\,\text{Gyr})$-old, which is comparable to the period of the Sun's revolution around the Galactic center. Also, geological processes occur on timescales shorter than $\mathcal{O}(1\,\text{Gyr})$, further complicating the expected directionality of the DM-induced signal. Reference [279] attempted to quantify the directional dependence of the DM-induced tracks within ancient minerals, showing that there is a preferred direction. Unfortunately for an $\mathcal{O}(1\,\text{Gyr})$ mineral the effect in ancient mica was calculated to be only $\mathcal{O}(1\%)$. Because it is unlikely that we will be able to resolve the head/tail orientation of tracks at the nm scale, the induced anisotropy would need to be much larger than $\mathcal{O}(1\%)$ in order to be statistically observable [321].

Finally, the translation of the range of the nucleus x_T to the reconstructed track length after read-out is a source of uncertainty. Quantifying such an uncertainty requires detailed studies for different combinations of minerals and read-out methods [264]. However, in the case of a claimed detection, we would be able to confirm a signal using minerals with different constituent nuclei and ages, allowing one to mitigate some of these systematic issues.

6.6 Conclusions

In this work, we have explored the prospects for probing Weakly Interacting Massive Particle (WIMP) Dark Matter (DM) with *paleo-detectors*. In particular, we have extended previous studies by performing a full spectral analysis, including information about the expected distributions of track lengths left in the minerals by a DM signal as well as by neutrino-induced and radiogenic backgrounds. We further explored how systematic uncertainties on the normalization and shape of backgrounds impact projected limits. Finally, we have studied how well the DM mass could be measured in case of a future discovery.

We considered 4 minerals in this work: halite [NaCl], olivine [$Mg_{1.6}Fe^{2+}_{0.4}(SiO_4)$], sinjarite [$CaCl_2{\cdot}2\,(H_2O)$] and nchwaningite [$Mn^{2+}_2SiO_3(OH)_2{\cdot}(H_2O)$]. Sinjarite is the most sensitive out of the minerals examined here due to the assumed low levels of radioactive contamination and efficient neutron moderation by hydrogen (Ref. [264] came to a similar conclusion for the mineral epsomite [$Mg(SO_4){\cdot}7\,(H_2O)$]).

For moderate track length resolutions, $\sigma_{x_T} = 15\,\text{nm}$, we find that the full spectral analysis extends the projected paleo-detector sensitivity to DM–nucleon cross sections roughly an order of magnitude smaller than the sliding window analysis of Refs. [263, 264]. This improvement is driven by the fact that a full spectral analysis automatically entails the use of optimal 'control regions', where the signal is sub-dominant, helping to pin down the normalization and shape of the backgrounds.

We find that by analyzing an $\mathcal{O}(10\,\text{cm}^3)$ sample of sinjarite using small angle X-ray scattering, it could be possible to probe DM–nucleon cross sections roughly a factor of 100 smaller than current direct detection experiments for DM heavier than $\sim 100\,\text{GeV}$. Including systematic uncertainties in background shapes at the 10 %-level, projected limits remain a factor of $7-8$ more stringent than current XENON1T bounds [18]. The sensitivity depreciates further if systematics larger than 10 % are assumed for the shapes of backgrounds.

Analyzing smaller samples of $\mathcal{O}(1\,\text{mm}^3)$ at nm-resolution (e.g. using helium ion beam spectroscopy), we find that paleo-detectors may be able to probe DM–nucleon scattering cross sections many orders of magnitude below current limits, for $500\,\text{MeV} \lesssim m_\chi \lesssim 10\,\text{GeV}$. Probing $\mathcal{O}(\text{nm})$ track lengths corresponds to an $\mathcal{O}(100\,\text{eV})$ energy threshold, exploring significant regions of the recoil spectra from low-mass WIMPs. With high-resolution read-out methods, the limits would be robust to systematic uncertainties in the background shapes as large as $\sim 50\,\%$.

In addition, we have investigated the prospects for paleo-detectors to constrain the DM parameters in the case of a future signal. As an example, we calculate the regions in which the mass and cross section become degenerate for a sinjarite paleo-detector. We find that below $m_\chi \lesssim 15\,\text{GeV}$, it would be possible to reconstruct the

mass of the DM particle with a relative error of less than 10 % if the cross section is large enough for a $5\,\sigma$ discovery. For $m_\chi \gtrsim 15\,\mathrm{GeV}$ the signal becomes increasingly insensitive to changes in m_χ, making the mass harder to constrain. In spite of this, paleo-detectors should be able to obtain both a lower and an upper limit on the DM mass for $m_\chi \lesssim 1\,\mathrm{TeV}$ if the cross section is just below current limits. In contrast, conventional direct detection experiments could provide only a lower bound on the DM mass if the true mass is larger than $\sim 200\,\mathrm{GeV}$ [2].

Paleo-detectors could also be used to investigate a number of interesting questions beyond searches for DM. For example, in the absence of a DM signal, the analysis outlined here would straightforwardly allow us to measure neutrino-induced events. It could therefore be possible to use mineral samples of different ages as a unique probe of the neutrino history of our galaxy. This will potentially allow us to study both historical neutrino processes in the Sun and the signal from supernovae.

Paleo-detectors represent an excellent opportunity to probe large areas of the WIMP DM parameter space in the near future. The next steps involve assessing the practical challenges of reaching the required exposures to achieve these sensitivities, as well as more detailed modeling of backgrounds. We leave both of these tasks to future work. However, we note that WIMP-nucleon cross sections much smaller than projected in this work may be probed by paleo-detectors if either novel ideas to control the backgrounds emerge (akin to progress made in conventional direct detection experiments in recent decades) or if target materials with significantly lower levels of radioactivity are available. With uranium concentrations of $\lesssim 10^{-15}$, radioactive backgrounds would no longer dominate at high DM masses. In such a case, paleo detectors could perhaps probe WIMP-nucleon cross sections all the way down to the diffuse supernova and atmospheric neutrino floor.

6.7 Appendix

6.7.1 Euclideanized Signals

Whether an experiment is *a priori* able to constrain a parameter of interest involves calculating the *expected statistical distinctness* between two signals, given a set of backgrounds and their associated uncertainties [9]. Points in the model parameter space are described by a d-dimensional vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_d)$. Two model parameter points $\boldsymbol{\theta}^{(1)}$, $\boldsymbol{\theta}^{(2)}$ can be considered as experimentally distinguishable if the parameter point $\boldsymbol{\theta}^{(2)}$ is inconsistent (at a given significance level) with the Asimov data $\mathcal{D} = \bar{\mathcal{D}}(\boldsymbol{\theta}^{(1)})$. For our application we have a two dimensional model where $\boldsymbol{\theta}^{(i)} = \{m_\chi^{(i)}, (\sigma_n^{\mathrm{SI}})^{(i)}\}$. In order to establish experimental distinguishability, we use

the maximum-likelihood ratio as a test statistic (TS) [146, 17],

$$\mathrm{TS}(\boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(1)}) \equiv -2 \ln \frac{\max_{\boldsymbol{\eta}} \mathcal{L}(\bar{\mathcal{D}}(\boldsymbol{\theta}^{(1)}) | \boldsymbol{\theta}^{(2)}, \boldsymbol{\eta})}{\max_{\boldsymbol{\eta}} \mathcal{L}(\bar{\mathcal{D}}(\boldsymbol{\theta}^{(1)}) | \boldsymbol{\theta}^{(1)}, \boldsymbol{\eta})} \,, \tag{6.15}$$

where $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\eta})$ is the likelihood function for data $\mathcal{D}$. It can depend on some nuisance parameters $\boldsymbol{\eta}$ that are profiled out when calculating TS. For model parameter points with sufficiently similar signals, the value of TS is approximately symmetric under $\boldsymbol{\theta}^{(1)} \leftrightarrow \boldsymbol{\theta}^{(2)}$. Hence, we can write

$$\mathrm{TS}(\boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(1)}) \approx (\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)})^T \mathcal{I} (\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)}) \,, \tag{6.16}$$

where

$$\mathcal{I}_{ij} = - \left\langle \frac{\partial^2 \log \mathcal{L}(\mathcal{D}|\boldsymbol{\theta}^{(1)})}{\partial \theta_i \partial \theta_j} \right\rangle_{\mathcal{D}(\boldsymbol{\theta}^{(1)})} \,, \tag{6.17}$$

is the Fisher information matrix at $\boldsymbol{\theta}^{(1)}$. The derivatives here describe the curvature in the direction of a particular parameter. The Fisher information matrix defines a metric on the space of model parameters, making it accessible to the tools of differential geometry.

The *Euclideanized signal method* is an approximate isometric embedding of a d-dimensional model parameter space (with geometry from the Fisher information metric) into n-dimensional Euclidean space: $\boldsymbol{\theta} \mapsto \mathbf{x}(\boldsymbol{\theta})$ with $\mathbf{x} \in \mathcal{M} \subset \mathbb{R}^n$ and $\boldsymbol{\theta} \in \mathbb{R}^d$. This embedding allows one to estimate differences in the log-likelihood ratio by the Euclidean distance,

$$\mathrm{TS}(\boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(1)}) \simeq \|(\mathbf{x}(\boldsymbol{\theta}^{(1)}) - \mathbf{x}(\boldsymbol{\theta}^{(2)})\|^2 \,. \tag{6.18}$$

Machine learning tools (in particular clustering algorithms, which usually assume Euclidean space) can then be used to efficiently explore the signal phenomenology of different models, and to systematically compare entire model classes, see Ref. [223]. For details on the Euclideanized signal transformation and its accuracy see Ref. [6]. The accuracy of the method (relative to the TS value) is at the $< 20\%$ level, and details can be found in Ref. [2].

We can now estimate the ability of an experiment to constrain the mass in the following way:

- Grid scan the parameter space, calculating signals for each point $\boldsymbol{\theta}^{(i)}$. Here, it is essential that all distinguishable model parameter points are covered down to a specific significance level (this should correspond to approximately 10 points per $1\,\sigma$ region).

- Euclideanize the signals using `swordfish` to produce associated vectors $\mathbf{x}_i$. Note that the transformation is able to account for arbitrary Gaussian background uncertainties and correlations, see Refs. [2, 6] for more details.

- For each parameter point $\boldsymbol{\theta}^{(i)}$ we calculate its associated nearest neighbours within a predefined significance. Here we use $2\,\sigma$ which corresponds to a radius of $\sqrt{2}$ in one dimension. The number of dimensions reflects the difference in dimensionality between the two-parameter model with $\{m_\chi, \sigma_n^{\mathrm{SI}}\}$ and the model living on the high mass boundary where m_χ is fixed. If this set of nearest neighbours contains a parameter point on the high mass boundary[7] the constraint on the mass around $\boldsymbol{\theta}^{(i)}$ is unbounded from above.

[7] Here we define the high mass boundary as 10 TeV.

7 Paleo-Detectors for Galactic Supernova Neutrinos

This chapter is based on work from Ref. [4]

7.1 Introduction

Supernovae (SNe) play an important role in cosmology and astrophysics. For example, SN feedback is thought to be an important ingredient for understanding galaxy formation [322]. While many extragalactic SNe have been observed [323, 324, 325, 326, 327, 328, 329], allowing for a rather precise determination of the cosmic SN rate [330, 331, 332], only a handful of SNe have been observed in the local group [333, 334]. To date, no direct measurement of the SN rate in the Milky Way exists; estimates in the literature suggest a rate of a few SNe per century [335, 336, 337, 338, 339, 340].

In this paper, we explore the potential of *paleo-detectors* to measure the core collapse (CC) SN rate in our galaxy. Paleo-detectors have recently been studied as a method for the direct detection of dark matter [263, 341, 3]. In certain minerals, e.g. those long used as solid state track detectors, recoiling nuclei leave damage tracks [265, 266, 267, 268]. Once created, such tracks are preserved over geological time scales. In paleo-detectors, one would search for damage tracks in minerals as old as $\sim 1\,\mathrm{Gyr}$ using modern nano-technology such as helium-ion beam or X-ray microscopy [263, 341]; see also [342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 277, 275, 356, 279] for related earlier ideas that use ancient minerals to probe rare events.

Besides probing dark matter, paleo-detectors would also detect neutrinos via nuclear recoils induced by coherent neutrino-nucleus scattering. Thus, paleo-detectors could, for the first time, provide a direct measurement of the galactic CC SN rate over the past $\sim 1\,\mathrm{Gyr}$.

SNe are broadly divided into thermonuclear (SNIa) and CC SNe. Only the latter ones are expected to produce a significant flux in neutrinos. The progenitors of CC

SNe are massive stars (heavier than $\sim 8\,M_\odot$). Such stars are short-lived, with lifetimes $\lesssim 50\,\mathrm{Myr}$, see e.g. [357, 358]. Thus, on the time scales relevant for paleo-detectors (order 100 Myr and longer), the CC SN rate closely traces the star formation rate, see e.g. [359, 360, 330, 331]. Considerable uncertainties exist in the estimates of the local star formation rate, see e.g. [361, 362, 363, 364, 365] for recent work. A direct measurement of the galactic CC SN rate would thus provide valuable information for understanding our galaxy.

While paleo-detectors would only provide a coarse-grained time resolution, we demonstrate that some time-dependent information of the galactic CC SNe rate can still be obtained. We consider two distinct cases: (i) we study how well a smooth time evolution of the CC SN rate could be constrained by paleo detectors, and (ii) we investigate if paleo-detectors could be used to find evidence for a starburst period in the Milky Way within the last $\sim 1\,\mathrm{Gyr}$. Both of these cases would provide information about the star formation history of the Milky Way. SN explosions in close proximity to Earth have also been hypothesized to give rise to mass extinction events, see e.g. [366, 367, 368, 369, 370, 371, 372, 373, 374]. We demonstrate that paleo-detectors could probe a single close-by CC SN explosion if it occurred during the exposure time.

The remainder of this paper is organized as follows. In Section 7.2, we discuss the track length spectrum produced in paleo-detectors from galactic CC SN neutrinos. In Section 7.3, we briefly review the relevant sources of backgrounds; a more detailed discussion can be found in [341]. The best read-out technique for the damage tracks induced by galactic CC SN neutrinos appears to be small angle X-ray scattering tomography, which we discuss in Section 7.4. Our projections for the sensitivity of paleo-detectors to galactic CC SNe as well as the time-evolution of the CC SN rate are discussed in Section 7.5. In Section 7.6 we summarize and discuss. Appendices 7.7.1 and 7.7.2 contain additional details about uranium-238 concentrations in typical target materials and the statistical techniques used in this work, respectively. All relevant code can be found online at DOI:10.5281/zenodo.3066206.

7.2 Galactic Core Collapse Supernova Signal

CC SNe are amongst the brightest astrophysical sources of neutrinos. In fact, SN 1987A (which occurred in the Large Magellanic Cloud) is the only astrophysical object, besides the Sun [and the recently-claimed flaring blazar TXS 0506+056 [375]], to be directly observed in neutrinos. Despite the important role neutrinos play in SN explosions [376, 377, 378], the precise shape and normalization of the emitted neutrino spectra are poorly understood. The only experimental knowledge stems from the emission of SN 1987A: the 20 events observed by Kamiokande-II [379], 8 events by IMB [380], 5 events by LSD [381], and 5 events by the Baksan Neutrino Observatory

ν	$E_\nu^{\rm tot}$ [erg]	$\langle E_\nu \rangle$ [MeV]	α
ν_e	6×10^{52}	13.3	3.0
$\nu_{\bar{e}}$	4.3×10^{52}	14.6	3.3
ν_x	2×10^{52}	15	3

Table 7.1: Parameters of the neutrino spectra, Eq. 7.1, for electron neutrinos, anti-electron neutrinos, and $\nu_x \equiv \{\nu_\mu, \nu_{\bar{\mu}}, \nu_\tau, \nu_{\bar{\tau}}\}$ used in our numerical calculations [386].

[382]. Alternatively, neutrino spectra can be predicted from simulations, which are usually well-fitted by a pinched Fermi-Dirac distribution [383]

$$\left(\frac{\mathrm{d}n}{\mathrm{d}E}\right)_{\nu_i} = E_\nu^{\rm tot} \frac{(1+\alpha)^{1+\alpha}}{\Gamma(1+\alpha)} \frac{E^\alpha}{\langle E_\nu \rangle^{2+\alpha}} e^{\left[-(1+\alpha)\frac{E}{\langle E_\nu \rangle}\right]} , \tag{7.1}$$

where $E_\nu^{\rm tot}$ is the energy radiated in the neutrino species ν_i, $\langle E_\nu \rangle$ is the average neutrino energy (approximately given by the core-temperature of the SN), and α is the spectral shape parameter. However, sizable differences remain between parameter values inferred from simulations, see for example [383, 384, 385, 386]. Here, we use the values suggested by [386], listed in Table 7.1.

The dominant source of neutrino-induced nuclear recoils are (flavor-blind) neutral current interactions. Thus, the relevant neutrino flux is the sum over all neutrino flavors,

$$\frac{\mathrm{d}n}{\mathrm{d}E_\nu} = \left(\frac{\mathrm{d}n}{\mathrm{d}E}\right)_{\nu_e} + \left(\frac{\mathrm{d}n}{\mathrm{d}E}\right)_{\nu_{\bar{e}}} + 4\left(\frac{\mathrm{d}n}{\mathrm{d}E}\right)_{\nu_x} , \tag{7.2}$$

where $\nu_x \equiv \{\nu_\mu, \nu_{\bar{\mu}}, \nu_\tau, \nu_{\bar{\tau}}\}$. Since neutral current interactions are flavor blind, we do not need to account for flavor oscillations. These are a major source of uncertainty when calculating the neutrino fluxes from SNe, due to the sizable matter effects in the SN environment.

The time-averaged neutrino spectrum from Galactic CC SNe at Earth is obtained by integrating over the probability density $f(R_E)$ describing the likelihood for a CC SN to occur at a distance R_E from Earth,

$$\left(\frac{\mathrm{d}\phi}{\mathrm{d}E_\nu}\right)^{\rm gal} = \dot{N}_{\rm CC}^{\rm gal} \frac{\mathrm{d}n}{\mathrm{d}E_\nu} \int_0^\infty \mathrm{d}R_E \frac{f(R_E)}{4\pi {R_E}^2} , \tag{7.3}$$

where $\dot{N}_{\rm CC}^{\rm gal}$ is the galactic CC SN rate.[1] To obtain $f(R_E)$, we follow [340] and assume

[1] In principle, the integral over R_E should be truncated at some distance corresponding to the size of our galaxy. Here, we instead use a probability density $f(R_E)$ which takes into account only CC SNe within the galactic disk of the Milky Way.

that CC SNe occur predominantly in the stellar disk. In galactocentric cylindric coordinates, the spatial distribution of CC SNe, ρ, can then be modeled by a double exponential

$$\rho \propto e^{-R/R_d} e^{-|z|/H} \,, \tag{7.4}$$

where R is the galactocentric radius, z is the height above the galactic mid-plane, and we set the disk parameters to $R_d = 2.9\,\mathrm{kpc}$ and $H = 95\,\mathrm{pc}$ [340]. From Eq. 7.3, we obtain the probability density as a function of the distance from Earth $f(R_E)$ by performing a coordinate transformation to the position of the Sun with galactocentric radius $R_\odot = 8.7\,\mathrm{kpc}$ and height above the disc $H_\odot = 24\,\mathrm{pc}$. Note that the position of the Sun with respect to the galactic center will change over the timescales that paleo-detectors were exposed to neutrinos from galactic CC SNe, $\mathcal{O}(1)\,\mathrm{Gyr}$[2]. The solar system is thought to follow an approximately circular orbit around the galactic center, oscillating about the galactic disk by $\Delta z_\odot \sim 100\,\mathrm{pc}$ and oscillating in the galactic plane by $\Delta R_\odot \sim 300\,\mathrm{pc}$, see e.g. [387, 388]. Modifying the distance of the solar system to the galactic center by such an amount would change the neutrino flux from CC SNe at Earth by $\Delta\phi \lesssim 10\,\%$, an error much smaller than the uncertainty on the galactic CC SN rate. In the following, we neglect corrections to the neutrino flux from the changing position of the solar system.

In Figure 7.1 we show the neutrino spectrum from galactic CC SNe together with the neutrino spectrum expected from distant CC SNe throughout the Universe, the so-called Diffuse SN Background (DSNB). We follow [308] for the calculation of the DSNB flux, using the parameterization from [331] for the cosmic CC SN rate, see also [330]. Assuming a galactic CC SN rate of $\dot{N}_{\mathrm{CC}}^{\mathrm{gal}} = 2.3 \times 10^{-2}\,\mathrm{yr}^{-1}$ [338], we find that the time-averaged neutrino flux from galactic CC SN at Earth peaks at $\mathrm{d}\phi/\mathrm{d}E_\nu \sim 10^2\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{MeV}^{-1}$ with $E_\nu \sim 10\,\mathrm{MeV}$. Note that the flux is approximately 100 times that of the DSNB flux. Further, the DSNB spectrum is shifted to lower energies by approximately a factor of two. This shift is due to the peak cosmic CC SN rate occurring at a redshift of $z \sim 1$ [331]. Note that estimates of the CC SN rate inferred from the cosmic star formation rate yield somewhat larger redshifts of $z \sim 2$ at which the star formation rate (and hence the CC SN rate) peaks, see e.g. [330]. The DSNB neutrino spectrum obtained from such parameterizations of the CC SN rate would be shifted to even smaller energies than the DSNB spectrum shown in Figure 7.1. However, as we will see in Section 7.3, such uncertainties on the DSNB neutrino flux are not important for this work as the dominant background for the signal from galactic CC SNe stems from radiogenic neutrons.

The observable in paleo-detectors is damage tracks caused by nuclear recoils. Neutrinos with energies $E_\nu \lesssim \mathcal{O}(100)\,\mathrm{MeV}$ give rise to nuclear recoils predominantly via coherent neutral current interactions.[3] The differential recoil spectrum per unit

[2] The orbital period of the Sun around the galactic center is $T_\odot \sim 250\,\mathrm{Myr}$.

[3] Additional contributions arise from quasi-elastic charged-current interactions. However, the

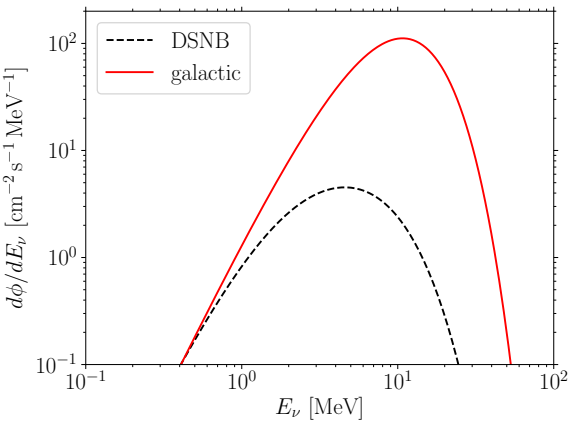

Figure 7.1: **Neutrino flux at Earth from galactic and extra-galactic CC SNe.** Neutrino flux $\mathrm{d}\phi/\mathrm{d}E_\nu$ (solid red) from galactic CC SNe at Earth as a function of neutrino energy E_ν. Here, we assumed a galactic CC SN rate of $\dot{N}_{\mathrm{CC}}^{\mathrm{gal}} = 2.3 \times 10^{-2}\,\mathrm{yr}^{-1}$ [338], a spatial distribution of CC SN as given in Eq. 7.4, the neutrino spectrum per CC SN from Eqs. 7.1, 7.2 with the parameters from Tab. 7.1, and averaged the neutrino flux over time-scales much longer than the inverse galactic CC SN rate $(\dot{N}_{\mathrm{CC}}^{\mathrm{gal}})^{-1} \sim 40\,\mathrm{yr}$. For comparison, the black dashed line shows the neutrino flux from distant CC SNe throughout the Universe, the so-called Diffuse SN Background (DSNB). See [308] for the calculation of the DSNB spectrum; we use the parameterization of the cosmic CC SN rate from [331].

target mass for target nuclei T is given by [306, 288]

$$\left(\frac{\mathrm{d}R}{\mathrm{d}E_R}\right)_T = \frac{1}{m_T}\int_{E_\nu^{\mathrm{min}}} \mathrm{d}E_\nu \, \frac{\mathrm{d}\sigma}{\mathrm{d}E_R}\frac{\mathrm{d}\phi}{\mathrm{d}E_\nu} \,, \tag{7.5}$$

where E_R is the nuclear recoil energy, m_T is the mass of T, $\mathrm{d}\sigma/\mathrm{d}E_\nu$ is the differential neutral current interaction cross section, and $E_\nu^{\mathrm{min}} = \sqrt{m_T E_R/2}$ is the minimum neutrino energy required to induce a nuclear recoil with energy E_R. The differential cross section is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_R}(E_R, E_\nu) = \frac{G_F^2}{4\pi} Q_W^2 m_T \left(1 - \frac{m_T E_R}{2E_\nu^2}\right) F^2(E_R) \,, \tag{7.6}$$

contributions to the recoil spectrum induced by CC SN neutrinos are suppressed at small neutrino energies by the lack of coherent enhancement, and at large neutrino energies by the quickly falling neutrino flux. The more energetic nuclear recoils which may be induced by high energy neutrinos furthermore lead to longer damage tracks than the nuclear recoils induced by CC SN neutrinos.

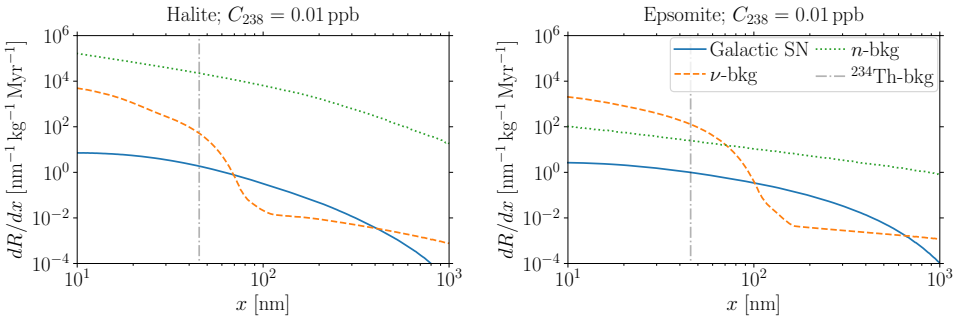

Figure 7.2: **Track length spectra in halite (NaCl; left) and epsomite [$Mg(SO_4)\cdot 7(H_2O)$; right].** In each panel, the blue solid line shows the spectrum from galactic CC SNe, assuming a rate of $\dot{N}_{\rm CC}^{\rm gal} = 2.3 \times 10^{-2}\,{\rm yr}^{-1}$. The dashed orange line indicates the background spectrum induced by coherent scattering of neutrinos from the Sun, the atmosphere, and the DSNB flux, and the dotted green line shows the background spectrum induced by neutrons from spontaneous fission and (α, n) processes in the target material. The vertical gray dash-dotted lines indicates the track length of the $E_R = 72\,{\rm keV}$ ^{234}Th nuclei from ($^{238}{\rm U} \to \alpha + {}^{234}{\rm Th}$) decays. See Section 7.3 for a discussion of the background spectra. For both target minerals, we assumed uranium-238 concentrations of $C_{238} = 0.01\,{\rm ppb}$ by weight. Note that although the signal rate is smaller than the background rate for all track lengths, this does not imply that the signal cannot be measured, see Section 7.5.

with the Fermi coupling constant G_F, the nuclear form factor $F(E_R)$, and

$$Q_W \equiv (A_T - Z_T) - \left(1 - 4\sin^2\theta_W\right) Z_T \,, \tag{7.7}$$

where θ_W is the weak mixing angle and A_T (Z_T) the number of nucleons (protons) in T. In our numerical calculations, we use the Helm nuclear form factor [290, 219, 291]

$$F(E_R) = 3\frac{\sin(qr_n) - qr_n\cos(qr_n)}{(qr_n)^3} e^{(qs)^2/2} \,, \tag{7.8}$$

where $q = \sqrt{2m_T E_R}$ is the momentum transfer and the effective nuclear radius is $r_n^2 \approx c^2 + \frac{7}{3}\pi^2 a^2 - 5s^2$ with $a \approx 0.52\,{\rm fm}$, $c \approx \left(1.23A_T^{1/3} - 0.6\right)\,{\rm fm}$ and $s \approx 0.9\,{\rm fm}$. Note that more refined calculations of the form factors are available, although only for a few isotopes, see e.g. [389, 390, 391, 392].

The recoil spectrum, Eq. 7.5, is converted into a track length spectrum by sum-

ming the stopping power $\mathrm{d}E_R/\mathrm{d}x_T$ over all target nuclei T in a material

$$\frac{dR}{dx} = \sum_T \xi_T \frac{dE_R}{dx_T} \left(\frac{dR}{dE_R}\right)_T \,. \tag{7.9}$$

Here, x (x_T) is the track length (of T), ξ_T is the mass fraction of T in the target material, and the track length for a recoiling nucleus with energy E_R is

$$x_T(E_R) = \int_0^{E_R} \mathrm{d}E \left|\frac{\mathrm{d}E}{\mathrm{d}x_T}\right|^{-1} \,. \tag{7.10}$$

In our numerical calculations, we use the `SRIM` code [393, 394] to calculate the stopping power in composite materials. A more detailed discussion of the calculation of the stopping power (in particular, a description of a semi-analytic calculation of the stopping power and comparison with `SRIM` results) can be found in [341].

In Figure 7.2 we show the track length spectrum from galactic CC SNe together with background spectra for two minerals, halite (NaCl) and epsomite [$\mathrm{Mg(SO_4)}\cdot 7(\mathrm{H_2O})$].

7.3 Backgrounds

A number of sources can give rise to backgrounds in the target sample. The background sources are the same as for dark matter searches with paleo-detectors; see [341] for a detailed discussion. Here, we give only a brief review of the most relevant background sources. Importantly, all the relevant backgrounds stem from nuclear recoils. Natural defects in minerals are either single-site or span across the entire (mono-)crystalline volume and thus do not resemble the damage tracks induced by neutrinos scattering off the nuclei in the target.

7.3.1 Cosmic Ray induced backgrounds

Cosmic rays can lead to both nuclear recoils and direct damage tracks in materials, potentially producing background events. However, cosmic ray induced backgrounds can be mitigated by using target materials obtained from deep below the surface of the Earth. The dominant cosmogenic background source will then be neutrons arising from cosmic ray muons interacting with nuclei in the vicinity of the target. Following [395] we estimate the neutron flux to be $\phi_n = \mathcal{O}(100)\,\mathrm{cm^{-2}\,Gyr^{-1}}$ for an overburden of $\sim 5\,\mathrm{km}$ rock. At a depth of $\sim 6\,\mathrm{km}$, the flux is instead $\phi_n = \mathcal{O}(10)\,\mathrm{cm^{-2}\,Gyr^{-1}}$ and for an overburden of $\sim 7\,\mathrm{km}$ we estimate $\phi_n = \mathcal{O}(0.1)\,\mathrm{cm^{-2}\,Gyr^{-1}}$. We envisage that target samples for paleo-detectors will have masses of order $100\,\mathrm{g}$, corresponding to

geometric cross sections of $\sim 10\,\mathrm{cm}^2$. Thus, for minerals obtained from depths larger than $\sim 6\,\mathrm{km}$, backgrounds due to cosmic ray induced neutrons will be negligible.[4]

7.3.2 Radioactive decays

The natural minerals used for paleo-detectors will be contaminated by trace amounts of radioactive elements which in turn give rise to background events. Thus, it is crucial to use materials containing as little radioactive contamination as possible. Minerals formed close to the surface of the Earth from the crust's material have prohibitively large uranium-238 and thorium-232 concentrations. Minerals formed in Ultra-Basic Rock (UBR) deposits and Marine Evaporites (MEs) are much cleaner. UBRs and MEs are comprised of material from the Earth's mantle and form at the bottom of evaporating bodies of water respectively. Here, we assume benchmark uranium-238 concentrations of 0.1 parts per billion (ppb) in weight for UBRs and 0.01 ppb for MEs. See Appendix 7.7.1 for further discussion.

The most relevant radioactive contaminant in UBR and ME minerals is uranium-238. The half-life of uranium-238 is 4.5 Gyr, while the accumulated half-life of all subsequent decays in the uranium-238 decay chain, until it reaches the stable lead-206, is $\sim 0.3\,\mathrm{Myr}$. Thus, almost all uranium-238 nuclei which undergo the first decay after the target mineral was formed will complete the decay chain. Due to the kinematics, the most problematic decays are α-decays. β- and γ-decays lead to the emission of fast electrons, photons, and neutrinos which do not themselves give rise to observable damage tracks in minerals. The associated recoils of the daughter nuclei from the decays are also too soft to produce observable tracks. α-decays on the other hand give rise to $10-100\,\mathrm{keV}$ recoils of the daughter nuclei and an α-particle with energy of order a few MeV. Here, we assume that the damage track from the α-particle itself is not directly observable, though see [341] and references therein for a discussion. Thus, the remaining signatures from α-decays are the $10-100\,\mathrm{keV}$ recoils of the daughter nuclei which give rise to damage tracks similar to those induced by scattering of CC SN neutrinos off the target nuclei. However, the typical decays of uranium-238 lead to a complete decay chain, which contains eight α-decays. This will lead to eight spatially connected tracks from the various daughter nuclei in the chain. Such signatures are straightforward to distinguish from the isolated damage tracks induced by neutrinos, and we will assume that all such track patterns can be rejected.

However, the second α-decay in the uranium-238 decay chain ($^{234}\mathrm{U} \to {}^{230}\mathrm{Th} + \alpha$) has a half-life of 0.25 Myr. This will lead to a non-negligible population of events which have undergone a single α-decay only. For minerals with ages long compared

[4] Note that for depths larger than $\sim 6\,\mathrm{km}$, in addition to neutrons from cosmogenic muons, neutron production from atmospheric neutrons interacting with nuclei in the vicinity of the target must be taken into account as well, see e.g. [396].

to the half-life of uranium-234 and short compared to the half-life of uranium-238, the number of such single-α events per unit target mass is well approximated by

$$n_{1\alpha} \approx 10^9\,\mathrm{kg}^{-1} \left(\frac{C_{238}}{0.01\,\mathrm{ppb}}\right) , \tag{7.11}$$

where C_{238} is the uranium-238 concentration per weight in the target sample. The energy of the ^{234}Th daughter nucleus from $^{238}\mathrm{U} \to {}^{234}\mathrm{Th} + \alpha$ decays is 72 keV, leading to a population of events with the corresponding (target-dependent) track length indicated by the dash-dotted vertical gray lines in Figure 7.2. The characteristic track length of such events allows for straightforward modeling of this background, leading to negligible effects on the sensitivity to CC SN neutrinos, as we will see below.

7.3.3 Neutron induced backgrounds

The two dominant sources of fast neutrons (in target minerals obtained from depths where cosmic ray induced neutrons are negligible) are spontaneous fission of heavy radioactive elements such as uranium-238[5] and neutrons produced by (α, n)-reactions of α-particles from radioactive decays with the nuclei in the target sample. Depending on the precise chemical composition of the target sample, either neutrons from spontaneous fission or from (α, n)-reactions dominate; we use the `SOURCES-4A` [397] code to obtain the neutron spectrum from both sources, including α-particles from the entire uranium-238 decay chain. Note that the (α, n) cross sections differ substantially between different elements and isotopes; thus, it is difficult to make general statements. However, light nuclei such as lithium or beryllium display particularly large (α, n) cross sections, making minerals containing sizable mass fractions of these elements not well suited for paleo-detectors due to the resulting large neutron fluxes.

Neutrons lose their energy predominantly via elastic scattering off nuclei, giving rise to nuclear recoils that are indistinguishable from those induced by neutrinos. Because of the mismatch between the neutron mass and those of most nuclei, neutrons lose only a small fraction of their energy in a single scattering event. For example, a $\sim 2\,\mathrm{MeV}$ neutron would give rise to ~ 200 nuclear recoils with $E_R \gtrsim 1\,\mathrm{keV}$ in a target material comprised of $m_T \sim 100\,\mathrm{GeV}$ nuclei. This background is highly suppressed in target materials containing hydrogen: since neutrons and hydrogen nuclei (i.e. protons) have approximately the same mass, neutrons lose a large fraction of their energy in a single collision with a hydrogen nucleus. Together with the relatively large elastic scattering cross section between a neutron and hydrogen, this makes hydrogen

[5] Note that the tracks from the fission fragments themselves are easily distinguished from neutrino-induced tracks. This is because fission fragments have energies of order 100 MeV, leading to much longer tracks than the $\lesssim 100\,\mathrm{keV}$ recoils induced by neutrinos from CC SNe.

an efficient moderator of fast neutrons, even if hydrogen makes up only a relatively small mass fraction of the target mineral. For each target mineral, we compute the nuclear recoil spectrum from the neutron spectra using a Monte Carlo simulation with neutron-nucleus cross sections as tabulated in the `JANIS4.0` database [398].[6] The corresponding track length spectrum is indicated by the dotted green lines in Figure 7.2.

7.3.4 Neutrino induced backgrounds

Neutrinos from sources other than galactic CC SNe induce nuclear recoils via the same scattering processes as neutrinos from galactic CC SNe. We take the neutrino flux $\mathrm{d}\phi_\nu/\mathrm{d}E_\nu$ for solar and atmospheric neutrinos from [288]. Extragalactic CC SNe throughout the Universe also give rise to a neutrino flux at Earth; we follow the same prescription as [308], with the parameterization of the cosmic CC SNe rate from [331] to calculate this so-called Diffuse SN Neutrino Background (DSNB), shown in Figure 7.1. There are three separate source regimes. Firstly, neutrinos with energies $E_\nu \lesssim 20\,\mathrm{MeV}$ are dominantly produced by solar emission. Secondly, neutrinos with $20\,\mathrm{MeV} \lesssim E_\nu \lesssim 30\,\mathrm{MeV}$ are primarily from the DSNB. Finally, atmospheric neutrinos produce the dominant flux for larger energies, $E_\nu \gtrsim 30\,\mathrm{MeV}$. The corresponding nuclear recoil and track length spectrum, shown by the dashed orange lines in Figure 7.2, is computed in the same way as for neutrinos from galactic CC SNe. Note that although we will investigate the sensitivity of paleo-detectors to potential variations of the galactic CC SN rate over geological time scales, we keep the background neutrino fluxes fixed at the values which are measured today in our background modeling. Instead, we account for potential variations in the background neutrino fluxes by assuming a large systematic uncertainty on the normalization of the neutrino-induced backgrounds.

7.3.5 Background uncertainties

The relevant background quantity is not only the total number of events, but the uncertainty on the number of background events in the signal region. While the statistical uncertainty is simply given by the square root of the number of background events in the signal region, we need to make assumptions about the systematic uncertainty for each background component. Here, we use the same values for each background component as in [263, 341, 3].

Radiogenic backgrounds, including neutrons induced by radioactivity, are well understood and straightforward to calibrate in the laboratory. For example, by plac-

[6] We use values from `TENDL-2017` [399, 400, 401, 402] for the neutron-nucleus cross sections.

ing radioactive sources in the vicinity of a test sample or by studying samples with relatively large concentrations of uranium-238, one can obtain samples with enhanced radiogenic backgrounds. Furthermore, the normalization of radiogenic backgrounds is determined by the concentration of heavy radioactive elements in the vicinity of the target sample and the age of the target sample only. Thus, we assume that radiogenic backgrounds can be well understood and assume a 1 % relative systematic uncertainty on the corresponding normalizations.

Neutrino induced backgrounds on the other hand are much harder to characterize; their normalization depends on the flux of neutrinos through the target sample. Although the present day neutrino fluxes are relatively well understood, paleo-detectors would measure neutrino-induced backgrounds integrated over geological timescales, $\lesssim 1\,\mathrm{Gyr}$. In this time the flux of atmospheric neutrinos may change substantially. Furthermore, creating target samples in the laboratory with enhanced backgrounds from neutrinos is challenging since this would require a strong neutrino source with a spectrum matching the neutrino spectrum from the Sun, the atmosphere, and CC SNe. Thus, we assume a large systematic uncertainty of 100 % on the normalization of neutrino-induced backgrounds.

7.4 Track Reconstruction

We refer the reader to [341] and references therein for a detailed discussion of damage tracks from nuclear recoils and possible read-out methods. As in [341] we will assume that the entire range of a recoiling nucleus will give rise to an observable damage track. Our studies with `SRIM` indicate that this is a reasonable assumption for the target materials and recoil energies considered here. The effect of possible corrections to this assumption cannot yet be answered quantitatively[7]; detailed experimental studies are required for each combination of target material and read-out method, which are beyond the scope of this work.

Further, we assume that low-Z nuclei, in particular α-particles (He ions) and protons (H ions) do not give rise to observable damage tracks. For further discussion, see [341].

From the track length spectra in Figure 7.2 we can see that the signal-to-noise ratio for CC SN neutrino induced events is largest for track lengths of $\mathcal{O}(100)\,\mathrm{nm}$. An optimal read-out method requires the resolution to which track lengths can be measured, σ_x, to be $\ll \mathcal{O}(100)\,\mathrm{nm}$. This in turn allows for an accurate measurement of the associated recoil energies. Unfortunately, the feasible size of target samples decreases with increasing spatial resolution. We will assume the use of Small Angle X-

[7] To the best of our knowledge, reliable estimates exist only for the particular case of reconstructing tracks in muscovite mica after cleaving and chemical etching [275].

ray scattering (SAXs) tomography at synchrotron facilities as our benchmark read-out scenario. SAXs allows for the three-dimensional read-out of bulk samples with spatial resolution of $\sigma_x \sim 15\,\text{nm}$ [287] and minimal sample preparation [271]. Note that as yet, damage tracks from ions have not been demonstrated to be reconstructible in three-dimensional SAXs tomography; however, damage tracks have been demonstrated to be observable with SAXs (without prior chemical etching) along the direction of the tracks [286]. While we are proposing a challenging application of SAXs, we estimate that it should be feasible to image $\mathcal{O}(100)\,\text{g}$ of target material at synchrotron facilities, with spatial resolutions of $\sigma_x = 15\,\text{nm}$. We note, as in [3], that this will present a significant data storage and analysis challenge. Naively, scanning at this level of precision will provide $\sim 10^7$ terabytes of data for a $\mathcal{O}(100)\,\text{g}$ sample. These issues are beyond the scope of this work and will be addressed in future publications.

7.5 Results

In this section, we present the projected sensitivity of paleo-detectors to neutrinos from galactic CC SNe. A key parameter which determines the sensitivity of paleo-detectors is the mineral age. Throughout this work, we use the term 'mineral age' for the age of the oldest nuclear recoil tracks which persist in the mineral. This should loosely correspond to the time since the formation of the mineral, though we note that on very long time-scales ($\gtrsim 1\,\text{Gyr}$ for many minerals) tracks will eventually disappear.

We begin by investigating the minimum time-averaged galactic CC SN rate to which paleo-detectors would be sensitive, both as function of the concentration of uranium-238 in the target sample and as a function of mineral age. We then investigate the ability of paleo-detectors to decipher the history of galactic CC SNe if one were to study a series of target minerals with ages $100\,\text{Myr} \leq t_{\text{age}} \leq 1\,\text{Gyr}$ and $\Delta t_{\text{age}} = 100\,\text{Myr}$. In each individual sample one would deduce the CC SN rate (within experimental uncertainties) integrated over $0 \lesssim t \lesssim t_{\text{age}}$. Using a series of samples with different t_{age} then allows to reconstruct the time dependence of the CC SN rate. In particular, we investigate the extent to which paleo-detectors could be used to measure the time-dependence of the galactic CC SNe rate. Finally we study the sensitivity of paleo-detectors to both a single near-by CC SN and a localized (in space and time) enhancement of the CC SN rate (as would be expected from a starburst event in the Milky Way or the local group).

Throughout, we use a spectral analysis similar to the procedure used in [3] for dark matter sensitivity forecasts. The analysis is performed using the `swordfish` python package [1, 6][8]. The main difference to the dark matter analysis is that now the signal arises from damage tracks induced by neutrinos from galactic CC SNe. To

[8] github.com/cweniger/swordfish

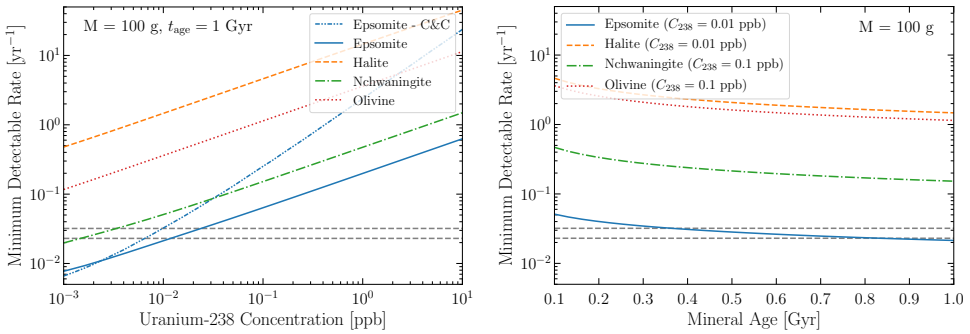

Figure 7.3: **Smallest galactic core collapse supernova rate which could be detected in paleo-detectors as a function of the uranium-238 concentration in the target sample (left) and time the mineral has been recording damage tracks (right).** Here, we assume that $M = 100\,\mathrm{g}$ of target material can be read out with a spatial resolution of $\sigma_x = 15\,\mathrm{nm}$. The different colored lines are for different target materials as indicated in the legend. For epsomite, we show in addition to the results from the spectral analysis also the sensitivity projections obtained in the sliding-window Cut-and-Count (C&C) analysis. The horizontal dashed gray lines indicate estimates for the galactic core collapse supernova rate, $\dot{N}_{\mathrm{CC}}^{\mathrm{gal}} = 2.3\times10^{-2}\,\mathrm{yr}^{-1}$ [403] and $\dot{N}_{\mathrm{CC}}^{\mathrm{gal}} = 3.2\times10^{-2}\,\mathrm{yr}^{-1}$ [340]. In the left panel we assume that the target sample has been recording damage tracks for $t_{\mathrm{age}} = 1\,\mathrm{Gyr}$. In the right panel, we assume a uranium-238 concentration of $C_{238} = 0.01\,\mathrm{ppb}$ in weight for the halite (NaCl) and epsomite [$Mg(SO_4)\cdot7(H_2O)$], which are examples of marine evaporites. For olivine [$Mg_{1.6}Fe^{2+}_{0.4}(SiO_4)$] and nchwaningite [$Mn^{2+}_2SiO_3(OH)_2\cdot(H_2O)$], minerals found in ultra basic rocks, we assume $C_{238} = 0.1\,\mathrm{ppb}$.

cross check, we also perform a sliding-window cut-and-count analysis analogous to the procedure used in [263, 341]. While the cut-and-count analysis is transparent and intuitive, the spectral analysis is more sensitive. The latter allows for the characterization of the backgrounds in control regions, leading to an effective reduction in uncertainties of the background in the signal region. Furthermore, `swordfish` allows one to straightforwardly explore projected confidence regions for reconstructed signal parameters. For all analyses, we consider systematic uncertainties on the normalization of each background component only. See Section 7.3.5 for a discussion of our assumptions and Appendix 7.7.2 for further details of our statistical methodology.

Throughout this section, we will assume that a mass $M = 100\,\mathrm{g}$ of target material can be read out with a spatial resolution of $\sigma_x = 15\,\mathrm{nm}$ (see discussion in Section 7.4).

7.5.1 Galactic CC SN rate

In Figure 7.3, we show the minimum galactic CC SN rate which could be observed in paleo-detectors[9] both as a function of the uranium-238 concentration in the target sample (left panel) and the mineral age (right panel). We consider four minerals, halite [NaCl], epsomite [$Mg(SO_4)\cdot 7(H_2O)$], nchwaningite [$Mn_2^{2+}SiO_3(OH)_2\cdot(H_2O)$], and olivine [$Mg_{1.6}Fe_{0.4}^{2+}(SiO_4)$]. Out of these four, epsomite is most promising. This is due to its chemical composition. First, epsomite contains hydrogen, which effectively suppresses the neutron induced backgrounds as described in Section 7.3. Further, epsomite does not contain any elements with large (α, n) cross sections and is a ME for which we expect low concentrations of uranium-238, $C_{238} \sim 0.01\,\text{ppb}$. Finally, epsomite's particular chemical composition emphasizes the difference between the signal and background spectra: for target nuclei lighter than $\sim 10\,\text{GeV}$, i.e. lighter than C, the spectrum from supernova induced neutrinos becomes increasingly similar to the background induced by solar neutrinos. For target elements heavier than $\sim 30\,\text{GeV}$, i.e. heavier than Si, both the signal and background track length spectra become increasingly compressed to shorter lengths. The finite spatial resolution of any given read out method makes it more difficult to distinguish signal from background for such compressed track spectra. In epsomite, the majority of nuclei lie between C and Si in mass, allowing a better separation of signal and background tracks. Thus, we will focus on epsomite as a target mineral for paleo-detectors in the remainder of this paper.

From the left panel of Figure 7.3 we find that reading out $M = 100\,\text{g}$ of epsomite which has been exposed to neutrinos from CC SNe for $t_{\text{age}} = 1\,\text{Gyr}$ should allow for a measurement of the average galactic CC SN rate. Current estimates for the galactic CC SNe rate suggest $\dot{N}_{\text{CC}}^{\text{gal}} \sim 2 \times 10^{-2}\,\text{yr}^{-1}$ and, as discussed in Section 7.3, we expect samples of ME minerals (e.g. epsomite) with uranium-238 concentrations of $C_{238} = 0.01\,\text{ppb}$ to be readily available in nature.

The right panel of Figure 7.3, where we fix the uranium-238 concentration of epsomite to $C_{238} = 0.01\,\text{ppb}$, indicates that measuring the galactic CC SN rate with an $M = 100\,\text{g}$ epsomite paleo-detector requires target samples which have recorded damage tracks for at least $t_{\text{age}} \sim 0.35 - 0.8\,\text{Gyr}$ (depending on the true rate). Note however that if only younger target samples were available, the sensitivity could be recovered by reading out a somewhat larger target sample. This is because the sensitivity depends on the exposure $\varepsilon = M \times t_{\text{age}}$; the numbers of signal events and the most relevant background events (i.e. recoils induced by other neutrinos and radiogenic neutrons) scale linearly with ε.

In the left panel of Fig. 7.3, we also show, for comparison, the sensitivity forecast

[9] Technically defined as the threshold for a $3\,\sigma$ detection, as in [6].

from the sliding-window cut-and-count analysis for epsomite. It should first be noted that the sensitivities are not directly comparable: for the sliding-window cut-and-count analysis, the signal is considered to be within reach when the Signal-to-Noise Ratio (SNR) (the ratio of the number of signal events and the quadratic sum of the systematic and statistical errors of all background components in the signal region) satisfies SNR > 3. For the spectral analysis, the signal is considered to be within reach if 50% of experiments would return a $3\,\sigma$ preference for the signal+background hypothesis over background-only [404]. The significance is evaluated from the Poisson likelihood ratio [6, 1, 3].

We note that for uranium-238 concentrations $C_{238} \gtrsim 10^{-2}\,\text{ppb}$, the smallest detectable galactic CC SN rate $\dot{N}_{\text{CC}}^{\text{gal}}$ scales as $\dot{N}_{\text{CC}}^{\text{gal}} \propto \sqrt{C_{238}}$ for the spectral analysis, while for the sliding-window cut-and-count analysis the scaling is $\dot{N}_{\text{CC}}^{\text{gal}} \propto C_{238}$. The scaling of the sensitivity in the cut-and-count analysis can be understood from the fact that for $C_{238} \gtrsim \mathcal{O}(0.01)\,\text{ppb}$, the sensitivity is limited by the systematic error on the number of background events induced by radiogenic neutrons in the signal region. For the spectral analysis, the error on the number of background events can be reduced by making use of control regions at longer track lengths. The error on the number of background events in the signal region then scales as $\sqrt{C_{238}}$, since it is given by the statistical error on the number of background events in the control regions. For $C_{238} < \mathcal{O}(0.01)\,\text{ppb}$, the number of events in the control regions becomes too small to allow such an approach. Simultaneously, the number of background events in the signal region becomes smaller, finally causing both analyses to be limited by the statistical error on the number of background events in the signal region. The remaining differences in the sensitivity are due to the different definitions of sensitivity as discussed above.

Before moving on to estimates of how well the time dependence of the galactic CC SN rate could be constrained by paleo-detectors, it is interesting to ask how precisely the time-averaged CC SN rate could be determined. As before, we consider a benchmark scenario of a 100 g epsomite paleo-detector which could be read out with spatial resolution of $\sigma_x = 15\,\text{nm}$, e.g. by SAXs. We assume that the target mineral has been recording events for 1 Gyr and that the true average galactic CC SN rate is $\dot{N}_{\text{CC}}^{\text{gal}} = 3 \times 10^{-2}\,\text{yr}^{-1}$. For a uranium-238 concentration of $C_{238} = 0.01\,\text{ppb}$, the reconstructed rate could be constrained, at 1σ, to $\dot{N}_{\text{CC}}^{\text{gal}} = (3.0 \pm 0.7) \times 10^{-2}\,\text{yr}^{-1}$. For uranium-238 concentrations of $C_{238} = 10^{-3}\,\text{ppb}$, the reconstructed rate could instead be constrained to $\dot{N}_{\text{CC}}^{\text{gal}} = (3.0 \pm 0.3) \times 10^{-2}\,\text{yr}^{-1}$. Thus, with sufficiently low uranium-238 concentrations it may be possible to constrain the time-averaged galactic CC SN rate to within 10%, allowing for a discrimination between different estimates in the literature [403, 340].

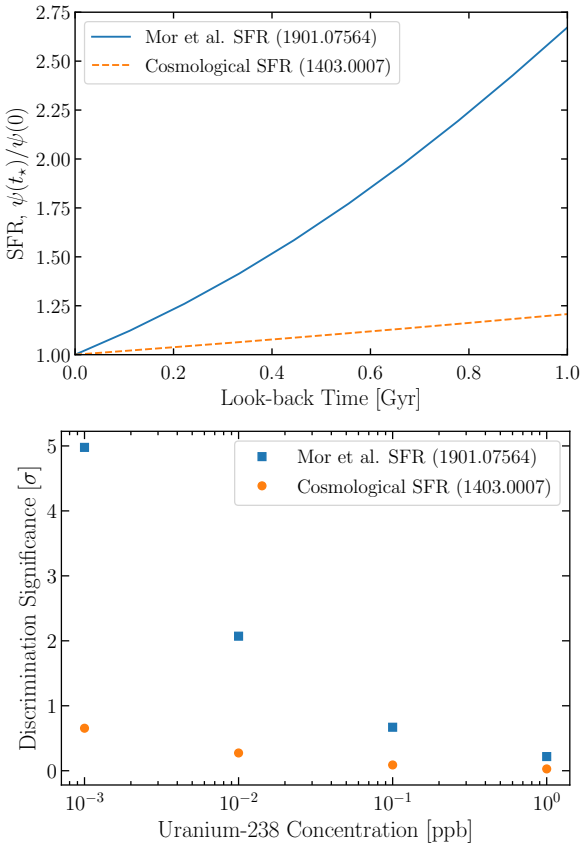

Figure 7.4: **Ability of paleo-detectors to distinguish between different time varying galactic SFRs.** *Left:* Benchmark scenarios for the time-dependence of the galactic Star Formation Rate (SFR) rate $\psi(t_\star)/\psi(t_\star = 0)$, as a function of look-back time, $t_\star$, considered in Section 7.5.2. The CC SN rate $\dot{N}_{\rm CC}^{\rm gal}$ is thought to be directly proportional to the SFR, $\dot{N}_{\rm CC}^{\rm gal} = k_{\rm CC}\psi$, and we use $k_{\rm CC} = 0.0068 M_\odot^{-1}$ [330]. The blue solid line shows the time-evolution of the galactic SFR of the Milky Way as estimated from Gaia data by [365], and the dashed orange line the time evolution of the cosmic SFR as estimated by [330]. *Right:* Assuming different uranium-238 concentrations $C_{238} = \{10^{-3}, 0.01, 0.1, 1\}$ ppb, we show the discrimination significance with which a time-constant galactic CC SN rate could be rejected if the true galactic CC SN rate evolves with look-back time as the corresponding benchmark scenario shown in the left panel. For both cases, we entertain an experimental scenario where 10 epsomite samples with $M = 100\,\text{g}$ each, which have been recording events for different times $t_{\rm age} = \{0.1, 0.2, 0.3, \ldots, 1.0\}$ Gyr, are read out with track length resolution of $\sigma_x = 15\,\text{nm}$.

7.5.2 Time dependence of the CC SN rate

In the previous subsection we investigated the smallest detectable time-constant CC SN rate. Here and in the following subsection, we instead investigate how paleo-detectors can be used to understand the time-evolution of the galactic CC SN rate.

We entertain two benchmark scenarios for the time-dependence of the galactic CC SN rate: (i) a rate increasing with look-back time according to the best-fit evolution of the galactic Star Formation Rate (SFR) obtained by [365] from Gaia data, and (ii) a rate increasing with look-back time proportional to the cosmic SFR as parameterized by [330], cf. the left panel of Figure 7.4. Note that scenario (i) is based on information from the Milky Way, while scenario (ii) is not (it relies purely on cosmological information).

To quantify the significance at which such scenarios could be distinguished using paleo-detectors, we consider an experimental scenario using 10 epsomite samples weighing $M = 100\,\mathrm{g}$ each, which have been recording events for different times $t_{\mathrm{age}} = \{0.1, 0.2, 0.3, \ldots, 1.0\}\,\mathrm{Gyr}$. We assume that each target sample is read out with track length resolution of $\sigma_x = 15\,\mathrm{nm}$. We then simulate expected signal (and background) events in each sample [for scenarios (i) and (ii)] and calculate the best fit value and error bars for the reconstructed time-integrated CC SN rate in each target sample. Assuming the error bars for the reconstructed rates are described by a Gaussian distribution, we then attempt to fit a time-constant galactic CC SN rate to the mock data, and quantify the statistical significance with which the hypothesis of a constant CC SN rate would be rejected.

In the right panel of Figure 7.4, we show the statistical significance with which a constant CC SN rate would be rejected in both scenarios as a function of the uranium-238 concentration in the target sample $C_{238} = \{10^{-3}, 0.01, 0.1, 1\}\,\mathrm{ppb}$. For scenario (ii), where we assume that the galactic CC SN rate increases with look-back time as the cosmic SFR, we see that it is difficult to distinguish such a time evolution from a constant CC SN rate even if the uranium-238 concentration in the target samples is $C_{238} = 10^{-3}\,\mathrm{ppb}$. This is because the cosmic SFR evolves quite slowly in time. For a look-back time of 1 Gyr, the cosmic SFR is only increased by a factor of $\psi(t = 1\,\mathrm{Gyr})/\psi(t = 0) \approx 1.2$, using the estimate of the cosmic SFR from [330]. In scenario (i) on the other hand, where the galactic CC SN rate evolves like the estimate for the galactic SFR from [365], we find that the hypothesis of a constant galactic CC SN rate could be rejected at more than $3\,\sigma$ if the uranium-238 concentration in the target samples is $C_{238} \lesssim 5 \times 10^{-3}\,\mathrm{ppb}$. This is because the estimate for the galactic SFR from Gaia data by [365] indicates a much faster increase of the SFR with look-back time than the cosmic SFR, $\psi(t = 1\,\mathrm{Gyr})/\psi(t = 0) \sim 3$.

7.5.3 Constraining burst-like CC SNe

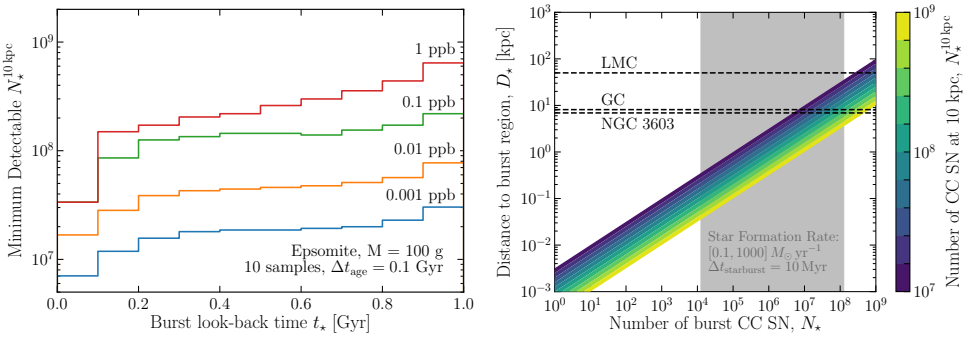

Figure 7.5: **Ability of paleo-detectors to distinguish a burst like event of neutrinos from a constant background.** *Left:* Smallest number of CC SN events $N_\star$ in a burst-like event at a distance of $D_\star = 10\,\mathrm{kpc}$, $N_\star^{10\,\mathrm{kpc}}$, which could be discovered with paleo-detectors as a function of the look-back time of the burst-like event $t_\star$. The different colored lines are for different uranium-238 concentrations, C_{238}, as indicated by the labels. Throughout, we assume that 10 samples of $M = 100\,\mathrm{g}$ epsomite each, which have been recording recoil events for different times $t_{\mathrm{age}} = \{0.1, 0.2, 0.3, \ldots, 1.0\}\,\mathrm{Gyr}$, can be read out with track length resolution of $\sigma_x = 15\,\mathrm{nm}$. The smallest number of detectable CC SN in a burst like event at any given distance $D_\star$ can be directly obtained by $N_\star \geq N_\star^{10\,\mathrm{kpc}} \left(10\,\mathrm{kpc}/D_\star\right)^2$. Similarly, the largest distance at which a single close-by CC SN could be discovered can be obtained by $D_\star \leq 10\,\mathrm{kpc}/\sqrt{N_\star^{10\,\mathrm{kpc}}}$. *Right:* The colored region indicates the range of $N_\star^{10\,\mathrm{kpc}}$ within reach of paleo-detectors, cf. the left panel, in the plane of the number of CC SNe in the burst-like event, $N_\star$, and the distance of Earth to the burst-like region, $D_\star$. The vertical gray band indicates typical values of $N_\star$ which could occur in a starburst, assuming a duration of the starburst of 10 Myr and an average star formation rate of $\psi = 0.1 \div 10^3\,M_\odot\,\mathrm{yr}^{-1}$. The horizontal dashed lines indicate distances to NGC 2603 (a nebula containing the dense open cluster HD 97950), the Galactic Center (GC), and the Large Magellanic Cloud (LMC), respectively.

After investigating the sensitivity of paleo-detectors to a smooth time evolution of the galactic CC SN rate, we now switch to asking if paleo-detectors could be sensitive to time- and space-localized enhancements in the local CC SN rate. The simplest example of such a *burst-like* event would be a single near-by CC SN. While such a single near-by CC SN would truly be localized in space and time, an enhancement to the CC SN rate (in a particular region of Milky Way or the local group) for a duration of time significantly smaller than the anticipated timing resolution of $\Delta t = 100\,\mathrm{Myr}$ would effectively also be a localized event. A starburst event, as described in [405, 406, 407, 408], in which the star formation rate (and hence the

SN rate) can exceed the average star formation rate of the Milky Way by a factor of $\sim 10^3$ for a period of $\Delta t_{\rm starburst} \lesssim 10\,{\rm Myr}$, is an example of an effectively localized event which could also be probed by paleo-detectors.

We parameterize such burst-like events by three parameters, $\{N_\star, D_\star, t_\star\}$. $N_\star$ is the number of CC SNe in the burst-like event, $D_\star$ is the distance to the burst region from Earth, and $t_\star$ is the look-back time to the burst event. For a single close-by CC SN, $N_\star = 1$. For a starburst event, $N_\star$ is given by the average star formation rate ψ during the length of the starburst $\Delta t_{\rm starburst}$ and the number of stars which explode as CC SNe per unit mass, $k_{\rm CC}$. For $k_{\rm CC} = 0.0068\,M_\odot^{-1}$ [330], a typical duration of a starburst of $\Delta t_{\rm starburst} = 10\,{\rm Myr}$, and SFRs of $0.1 \lesssim \psi/(M_\odot\,{\rm yr}^{-1}) \lesssim 10^3$ [see e.g. [405]], we find $N_\star = k_{\rm CC}\psi\Delta t_{\rm starburst} \sim 10^4 \div 10^8$. For reference, the number of CC SNe expected over the entire Milky Way within 1 Gyr is 2.3×10^7, assuming a constant rate of $\dot{N}_{\rm CC}^{\rm gal} = 2.3\times10^{-2}\,{\rm yr}^{-1}$.

To estimate the sensitivity of paleo-detectors to burst events, we follow a similar approach as in the previous subsection. We again assume that 10 samples of $M = 100\,{\rm g}$ epsomite detectors have been recording events for different times $t_{\rm age} = \{0.1, 0.2, 0.3, \ldots, 1.0\}\,{\rm Gyr}$ and can be read out with track length resolution of $\sigma_x = 15\,{\rm nm}$. We simulate mock data, assuming a time- and space-localized injection of additional neutrinos from CC SNe in a burst-like event on top of a constant galactic CC SN rate. Assuming Gaussian errors on the reconstructed CC SN rate in each target sample, we then attempt to fit the null-hypothesis of a time-constant galactic CC SN rate to the mock data and quantify the statistical significance with which this null-hypothesis is disfavored. The number of additional signal events from the burst-like event is proportional to $N_\star/D_\star^2$. We show results for the minimum value of $N_\star/D_\star^2$ required for a $3\,\sigma$ rejection of the null hypothesis as a function of the look-back time to the burst-like event $t_\star$ and the uranium-238 concentration in the target samples C_{238}.

For clarity of discussion we parameterize burst-like events with the three parameters $\{N_\star, D_\star, t_\star\}$, although our analysis is only sensitive to the combination $N_\star/D_\star^2$ as a function of $t_\star$. Hence, our benchmark scenarios of a starburst event or a single close-by CC SN are degenerate, although we discuss the results for both cases separately. We leave the exploration of discriminating such signals for future work. One possibility would be to study the anisotropy of damage track directions. The tracks from an individual close-by CC SN would all arise within $\mathcal{O}(10)\,{\rm s}$, the duration of the SN neutrino burst. On such time-scales, the target mineral would be virtually stationary in space and hence the signal tracks would have a preferred direction. For a starburst event, the signal events are expected to be generated over a time scale of a few tens of Myr, on which the rotation of the Earth, its orbit around the Sun, and the solar system's movement through the galaxy would wash out the directional preference of the tracks. Note that the directional preference would also allow for

additional background suppression when searching for signatures from an individual SN, potentially leading to increased sensitivity.

In the left panel of Figure 7.5, we show the minimum number of CC SNe in a burst like event at $D_\star = 10\,\mathrm{kpc}$[10], $N_\star^{10\,\mathrm{kpc}}$, for which the null-hypothesis of a time-constant galactic CC SN rate would be disfavored by at least $3\,\sigma$. We show these results for various assumptions on the uranium-238 concentration in the target samples, $C_{238} = \{10^{-3}, 0.01, 0.1, 1\}\,\mathrm{ppb}$. Trivially, we find that the smaller the uranium-238 concentration and hence the number of background events induced by radiogenic neutrons, the smaller $N_\star^{10\,\mathrm{kpc}}$ for which the null-hypothesis of a constant CC SN rate could be rejected. Further, we find that the smaller the look-back time to the burst-like event, $t_\star$, the smaller the number of $N_\star^{10\,\mathrm{kpc}}$ required to reject the null hypothesis. This is because the same number of signal events induced by the burst-like event would be present in all target samples with $t_\mathrm{age} < t_\star$, while the number of tracks from the time-constant galactic CC SN rate and from radiogenic neutrons, the most relevant background source, increases linearly with t_age.

In the right panel of Figure 7.5, we show our results in the $N_\star$–$D_\star$ plane. The colored band indicates different values of $N_\star^{10\,\mathrm{kpc}}$, the edges of the band approximately correspond to the range of values for $N_\star^{10\,\mathrm{kpc}}$ that would allow for the rejection of a time-constant galactic CC SN rate for the different assumptions on $t_\star$ and C_{238}, as shown in the left panel of Figure 7.5. To interpret these results, we indicate a range of values for the number of CC SNe, $N_\star$, in typical starburst events as well as the distance to various regions within the local group where starbursts are likely to occur. We find that for a range of possible starburst parameters, such a burst-like event would be detectable in paleo-detectors if it occurred at a distance corresponding to NGC 6303 (a nebula containing the open cluster HD 97950) or the galactic center. Detection of a starburst-event in the Large Magellanic Cloud (LMC), on the other hand, would require a starburst with a SFR a factor of a few higher than the typical range.

From Figure 7.5 we can also read off the minimal distance $D_\star$ for which an individual close-by CC SN could be detected with paleo-detectors. Depending on the uranium-238 concentrations in the target samples and the look-back time to the close-by CC SN, the null-hypothesis of a constant galactic CC SN rate could be rejected if the distance to the CC SN was smaller than $D_\star \lesssim 1 \div 10\,\mathrm{pc}$. For a spatial distribution of the galactic CC SNe as discussed in Section 7.2 and an average CC SN rate of $\dot{N}_\mathrm{CC}^\mathrm{gal}$, the probability that a CC SN has occurred within a distance $D_\star \lesssim 10\,\mathrm{pc}$ from Earth within 100 Myr is only $\sim 5\,\%$. However, despite the rather small statistical chance of such an event, close-by SNe are of particular interest since they may be related to mass extinction events, see [366, 369, 370, 371, 372, 373]. Although the time-resolution of paleo-detectors is rather coarse-grained, valuable information about possible close-by

[10] Here we use 10 kpc as a simple illustration rather than a distance of physical significance.

CC SNe may still be gained.

Furthermore, measurements of ^{60}Fe (and other isotopes produced in CC SNe) in sediments from the Earth and the Moon [367, 368, 409, 410, 411, 412, 413, 374] as well as the effects of such isotopes on the cosmic ray spectra [414, 415, 416, 417] suggest the explosion of at least one CC SN within $D_\star \lesssim 100\,\text{pc}$ from Earth $t_\star \sim 2 \div 3\,\text{Myr}$ ago. These claims could be tested with paleo-detectors by studying samples of minerals with $t_{\text{age}} \lesssim 10\,\text{Myr}$, much younger than what we discussed above. Since in paleo-detectors the signal would arise from the CC SN neutrinos, paleo-detectors would allow for a more direct characterization of these nearby SNe than measurements relying on ^{60}Fe and similar elements (slowly) propagating in cosmic rays.

7.6 Discussion

Paleo-detectors are a proposed experimental technique where one would search for the traces of nuclear recoils recorded in ancient minerals. In minerals that can be used as solid-state track detectors, ions traveling through the crystal lattice give rise to damage tracks which, once created, persist for geological time scales. With modern read-out technology it should be feasible to reconstruct such damage tracks with track length resolutions of order $1 \div 10\,\text{nm}$. Ions giving rise to such short tracks have kinetic energies $E_R \sim 1\,\text{keV}$. Combined with the retention of damage tracks over long times, paleo-detectors would represent a method to probe nuclear recoils down to energy thresholds of order keV whilst obtaining exposures as large as $\varepsilon \sim 100\,\text{g}\,\text{Gyr} = 10^5\,\text{t}\,\text{yr}$ with current read-out technology. [263, 341, 3] explored the potential of paleo-detectors for the direct detection of dark matter. Here, we studied how paleo-detectors can be used to understand galactic Core Collapse (CC) Supernovae (SNe) through the nuclear recoils induced via coherent scattering of neutrinos from CC SNe.

In Section 7.3 we discussed the most relevant background sources when searching for recoils induced by neutrinos from galactic CC SNe. At small track lengths (corresponding to less energetic nuclear recoils), the dominant background is solar neutrinos. At larger track lengths (i.e. more energetic nuclear recoils) the main background comes from nuclear recoils induced by fast neutrons from the radioactive processes of the trace amounts of uranium-238 and other heavy radioactive elements. Both of these will be present in target materials for paleo-detectors. Note that cosmogenic backgrounds, including neutrons induced by cosmogenic muons interacting in the vicinity of the target materials, can be mitigated by using target samples obtained from depths larger that $\sim 6\,\text{km}$ rock overburden, e.g. from the cores of deep boreholes. Unless the concentration of uranium-238 in the target material is less than $C_{238} \lesssim 10^{-14}$ in weight, the sensitivity of paleo-detectors to neutrinos from galactic CC SNe will be limited by radiogenic neutrons. As discussed in detail in

Appendix 7.7.1, we expect to be able to find target samples with uranium-238 concentrations of $C_{238} \sim \mathcal{O}(10^{-11}) = \mathcal{O}(0.01)\,\text{ppb}$.

For these concentrations of heavy radioactive elements, we showed in Section 7.5 that one could measure the time-averaged galactic CC SNe rate using paleo-detectors if the true rate is within the range of current estimates of $2 \div 3$ CC SNe per century in the Milky Way. We also investigated how paleo-detectors could be used to understand the time-evolution of the galactic CC SNe rate. To this end, we considered an experimental scenario where one would use 10 target samples which have been recording nuclear recoil tracks for different times $t_{\text{age}} = \{0.1, 0.2, 0.3, \ldots, 1\}\,\text{Gyr}$. If the galactic CC SNe rate was a factor of ~ 3 higher 1 Gyr ago than today, as indicated by Gaia data [365], paleo-detectors would allow one to reject the hypothesis of a time-constant galactic CC SN rate to high statistical significance. If, on the other hand, the galactic CC SN rate increases with look-back time similarly to the cosmic star formation rate (corresponding to a galactic CC SN rate a factor ~ 1.2 higher 1 Gyr ago than today) the data obtainable through this experimental scenario would not suffice to distinguish such a time evolution from a time-constant galactic CC SN rate.

Finally, we investigated how paleo-detectors could be used to learn about an enhancement of the local CC SN rate on time scales short compared to the time resolution of paleo-detectors, which is of order 100 Myr. Such a *burst-like* enhancement of the CC SN rate could arise from a single close-by CC SN, or from a period of significantly enhanced star formation activity in some region of the local group, i.e. a starburst period. For the latter, we have demonstrated that paleo-detectors could detect a sizable starburst period in the galactic center or a region of our galaxy of comparable distance, e.g. the nebula NGC 3603, if it occurred less than $\sim 1\,\text{Gyr}$ ago. Paleo-detectors could also be sensitive to a starburst in the Large Magellanic Cloud, but only in the case of an exceptionally strong starburst with star formation rates $\psi \gtrsim 10^4\,M_\odot\,\text{yr}^{-1}$ sustained for $\Delta t_{\text{starburst}} \sim 10\,\text{Myr}$.[11] Similarly, a close-by individual CC SN during the last $\sim 1\,\text{Gyr}$ could leave a detectable signature in paleo-detectors if it occurred at a distance of $\lesssim 10\,\text{pc}$ from Earth. In the analysis carried out here, we only considered the number of nuclear recoils induced by neutrinos from CC SNe. In such an analysis, enhancements in the signal rate from a starburst period or a single close-by SN would be indistinguishable; see Section 7.5.3 for a discussion of how this degeneracy could be broken using the directionality of the signal tracks.

In conclusion, this paper demonstrates that paleo-detectors are a promising experimental technique to obtain information about the rate of galactic CC SNe. The long timescales $t_{\text{age}} \lesssim \mathcal{O}(1)\,\text{Gyr}$ over which paleo-detectors could have recorded nuclear recoils induced by neutrinos from CC SNe would furthermore offer the unique ability

[11] Note though that Large Magellanic Cloud is not large enough to sustain such an exceptionally strong starburst, which would require $\gtrsim 10^{11}\,M_\odot$ of baryonic mass.

for a direct determination of the history of the galactic CC SN rate over geological time-scales. Because the star formation rate is thought to be directly proportional to the CC SN rate, this would allow for a measurement of the star formation history of our galaxy, providing important information for the understanding of the Milky Way.

7.7 Appendix

7.7.1 Uranium-238 Concentrations

The concentration of uranium-238, C_{238}, in target samples plays an important role in paleo-detectors because radioactive processes are one of the most relevant sources of backgrounds. As discussed in Section 7.3, see also [341], nuclear recoils induced by radiogenic neutrons are of particular relevance. Fast neutrons produced by spontaneous fission and (α, n) reactions lose their energy predominantly via elastic scattering off nuclei within the target material. The mean free path of fast neutrons in typical minerals is $\mathcal{O}(1)\,\mathrm{cm}$. Furthermore, fast neutrons undergo $\sim 10 \div 10^3$ elastic interactions before losing enough of their energy to no longer give rise to nuclear recoils similar to those induced by neutrinos from CC SNe or dark matter.

Because of the range of the neutrons, the relevant uranium-238 concentration is not necessarily in the target volume itself, but rather the average uranium-238 concentration in an $\mathcal{O}(\mathrm{m}^3)$ volume around the target sample. Modeling of radiogenic neutron backgrounds in an inhomogeneous environment would require knowledge of the geometry and composition of the rock surrounding the target samples. In our background modeling, we assume an infinitely-sized mineral of constant chemical composition. Note that inhomogeneities may lead to either higher or lower neutron-induced backgrounds in the target material. For example, the neutron-induced background in a relatively uranium-rich target sample not comprising hydrogen could be lower by orders of magnitude compared to the background calculated using the infinite mineral approximation, if such an $\mathcal{O}(\mathrm{cm}^3)$ target sample was located in a surrounding $\mathcal{O}(\mathrm{m}^3)$ volume of material where the uranium concentration is lower and hydrogen is present.

Further, we would like to note that the theoretical estimation of uranium concentrations in natural minerals is notoriously difficult. This is because the concentration depends not only on the average uranium concentration of the material in which the mineral forms, but also on the details of how uranium in incorporated into particular minerals during the formation process. For example, many minerals are rather robust to the introduction of heavy radioactive elements into the crystal lattice and thus chemically expel uranium (and similar heavy elements) during their growth. However, the effect of such purification cannot be quantified in general, see e.g. [418] for a discussion. In the remainder of this section, we will motivate our choices of

benchmark values for the uranium-238 concentrations in possible target materials for paleo-detectors. Ultimately, experimental efforts are required to determine the range of uranium concentrations in the most relevant target materials; we are currently coordinating such an effort. Note that once obtained, concentrations of radioactive trace elements in target samples of interest can be measured reliably to levels as low as $\sim 10^{-15}$ in weight, e.g. using inductively coupled plasma mass spectroscopy [419, 420].

A rather comprehensive discussion of the concentration of uranium-238 and other heavy radioactive contaminants in natural minerals can be found in [418]. Typical concentrations of uranium-238 in minerals formed from material in the Earth's crust are of the order of parts per million (ppm) in weight, which would lead to prohibitively large numbers of radiogenic background events in paleo-detectors. However, much lower uranium concentrations are found in minerals which compose Ultra Basic Rocks (UBRs)[12] and Marine Evaporites (MEs)[13]. This is because UBRs form from material in the Earth's mantle and MEs from salt deposits at the bottom of evaporated bodies of water. Both the Earth's mantle and its seas have uranium-238 concentrations a few orders of magnitude below the material in the Earth's crust and therefore minerals in UBRs and MEs are much better suited as paleo-detectors.

[418] quote values for uranium concentrations in UBRs of $1 \div 30\,\text{ppb}$ (parts per billion) and uranium concentrations of $\lesssim 100\,\text{ppb}$ in MEs. However, the aim of [418] was not to find the most radiopure rocks; in particular for MEs the quoted values represent upper limits of uranium-238 concentrations. While the ranges of uranium-238 concentrations given in the literature are typically representative of the most likely values for a given UBR or ME, [418] note that variations of up to an order of magnitude outside of such ranges are common. While experimental efforts are under way to better characterize the distributions of uranium-238 concentrations in representative target materials for paleo-detectors, in particular for the case of MEs with $C_{238} \lesssim$ ppb, our benchmark values represent roughly an order of magnitude downward variation from the most likely ranges given in the literature.

Ultra Basic Rocks

More detailed discussions of UBRs with $C_{238} \lesssim \mathcal{O}(1)\,\text{ppb}$ can be found, for example, in [421]. They reported uranium-238 concentrations of $\mathcal{O}(0.1)\,\text{ppb}$ uniformly distributed in (clino)pyroxenes, minerals which, along with olivine, constitute most of UBRs. Note however that these concentrations can vary upward by a factor of ~ 100 from

[12] Olivine $[Mg_{1.6}Fe^{2+}_{0.4}(SiO_4)]$ is very common in UBRs. We also show results for nchwaningite $[Mn^{2+}_2SiO_3(OH)_2 \cdot (H_2O)]$ in this work, a less common mineral found in UBRs which contains hydrogen.

[13] Halite (NaCl) is one of the most common MEs. We also present results for epsomite $[Mg(SO_4) \cdot 7(H_2O)]$ in this work, a less common example of MEs.

rock to rock, with the upper end of the range of C_{238} consistent with the values reported in [418]. [422] found a similar range for the uranium-238 concentration in UBRs, with some minerals having concentrations as low as $\mathcal{O}(0.1)$ ppb. Further, [422] suggest that the large variation of uranium-238 concentration in UBRs stems from different amounts of more uraniferous materials introduced after the original rock had formed. Both [421] and [422] also suggest (but do not conclusively prove) that such alterations are more prevalent in oceanic UBRs than in continental UBRs.

Marine Evaporites

For MEs there is considerably less published data available. In particular, many of the available data sets only provide upper limits on the uranium concentrations since the true level of uranium-238 in MEs is often below the sensitivity threshold of a given measurement technique. Some of the smallest uranium-238 concentrations in MEs have been reported by [423]. They reported uranium-238 concentrations of $\mathcal{O}(0.1)$ ppb in halite, however, their samples exhibited characteristics suggesting significant impurities. A more recent review of trace elements in MEs is given by [424]. They report uranium-238 concentrations ranging form $\mathcal{O}(0.1)$ ppb to $\mathcal{O}(10)$ ppb. Such large variations in uranium-238 concentrations from sample to sample are difficult to explain from first principles.

However, for MEs one can at least estimate the uranium-238 concentration and demonstrate that this is consistent with the observed range. Additionally, one can therefore estimate the lowest uranium-238 concentrations one may expect to find. Let us consider halite as a typical example of a ME and assume it forms in a body of water large enough such that the environment surrounding the water has a negligible effect on the average uranium-238 concentration (e.g. sufficiently deep ocean water). The uniformly distributed uranium concentration in a halite deposit formed under such conditions can be estimated as

$$C_{238}^{\mathrm{ME}} \sim C_{238}^{\mathrm{H_2O}} \times S_{\mathrm{H_2O}}^{-1} \times \alpha_{\mathrm{NaCl}} \,, \tag{7.12}$$

where $C_{238}^{\mathrm{H_2O}}$ is the uranium-238 concentration and $S_{\mathrm{H_2O}}$ the salinity of the original water, and α_{NaCl} is the ratio of the uranium concentration in the halite deposit to that of the residue left over from the original water. While our simple approximation does not necessarily hold for MEs formed in shallower bodies of water in which the deposition environment can significantly impact the uranium-238 concentration, we note that the ranges of C_{238}^{ME} values measured in such environments are similar. For typical values of seawater today, $C_{238}^{\mathrm{H_2O}} = 3\,\mathrm{ppb}$ [418] and $S_{\mathrm{H_2O}} = 35\,\mathrm{g\,kg^{-1}}$, and assuming that uranium from the water enriches the mineral phase of the evaporite and the leftover water residue equally, $\alpha_{\mathrm{NaCl}} = 1$, we find $C_{238}^{\mathrm{ME}} \approx 90\,\mathrm{ppb}$, which is roughly consistent with the upper limit given in [418].

However, much lower uranium concentrations can be accommodated in our estimate. The uranium-238 concentration in seawater does vary and values as low as $C_{238}^{\mathrm{H_2O}} = 0.3\,\mathrm{ppb}$ have been reported [418]. While the typical uranium-238 concentrations of sea water are not expected to have varied much over the relevant geological times scales, the salinity of seawater is generally assumed to have been significantly higher in the past than it is today, see e.g. [425]. Assuming a factor of two increase in the salinity for ancient oceans relative to today, we find $C_{238}^{\mathrm{ME}} \sim 4\,\mathrm{ppb} \times \alpha_{\mathrm{NaCl}}$. As discussed e.g. in [426], typical values of α_{NaCl} are considerably smaller than $\alpha_{\mathrm{NaCl}} = 1$ because uranium can be maintained as a stable complex anion in the water residue without being precipitated. [426] report values of $\alpha_{\mathrm{NaCl}} = 0.006$ and $\alpha_{\mathrm{NaCl}} = 0.011$ for different samples. Taking such effects into account, one may expect the lower range of typical uranium-238 concentrations in MEs to be $C_{238}^{\mathrm{ME}} = \mathcal{O}(0.01)\,\mathrm{ppb}$, which we assume as the benchmark value for our background modeling.

7.7.2 Statistical Techniques

Here we discuss additional details of the spectral analysis used for sensitivity projections. All analyses were performed using `swordfish` (github.com/cweniger/swordfish), an analysis tool developed in [6, 1]. `swordfish` automatically uses the spectral differences between the signal and background models to calculate accurate sensitivity projections, regardless of the statistical regime (Gaussian or Poissonian). This is made possible through the *equivalent counts method*, introduced in Section 2.4 of [6]. In Section 7.5.1 we calculate the minimum rate required to be detectable at $3\,\sigma$ significance. We define this rate to be the discovery threshold, as discussed in [1, 404]. In particular, this is given by the value of the rate that leads (in 50% of the cases) to a rejection of the no-signal hypothesis at $3\,\sigma$. The exact definition is given in Eq. (8) of [1].

In Section 7.5.2 we discuss the ability of paleo-detectors to decipher the time evolution of the galactic CC SN rate. Here we present the procedure used to calculate the model selection statements in more detail. Note that we proceed similarly for the time varying signal and burst search. Importantly, the ten mineral ages we consider can be treated as independent data sets since no two minerals will record tracks induced by the same neutrino. We first simulate the expected rates, R_i^o, from a time varying signal in each age bin i. This expectation is specific to the model under consideration. We then calculate the expected errors on the reconstructed values, σ_i^2, and fit a time-constant rate by minimizing the chi-squared difference,

$$\chi^2 = \sum_i \frac{\left(R_i^c - R_i^o\right)^2}{\sigma_i^2}\,, \tag{7.13}$$

where R_i^c is the time-constant rate which is varied to best fit R_i^o. We then calculate the statistical distinctness between the signals given by R_i^o and R_i^c. For this we make use of *Euclideanised signals* $x_i(R)$, a technique developed in [6, 3]. The Euclideanized signal method is an approximate isometric embedding of a d-dimensional model parameter space (with geometry from the Fisher information metric) into n-dimensional Euclidean space here given by : $\mathbf{R} \mapsto \mathbf{x}(\mathbf{R})$ with $\mathbf{x} \in \mathcal{M} \subset \mathbb{R}^n$ and $\mathbf{R} \in \mathbb{R}^d$. The full definition is given in Eq. (A18) of [1]. This embedding allows one to estimate differences in the log-likelihood ratio by the Euclidean distance (in units of σ) as,

$$d = \sqrt{\sum_i |x_i(R_i^o) - x_i(R_i^c)|^2} \,, \tag{7.14}$$

as shown in Figure 7.4. Here, d quantifies the degree to which a time-constant rate would be disfavored by a data-set consistent with the time-varying rates we considered. A similar procedure is used in the burst search but instead we compute the minimum enhancement to R_i^o required to give $d > 3$, as shown in Figure 7.5.

8 Conclusion

Searching for signs of new physics is a difficult task. Luckily, we live in an age of extreme technological innovation that enables unprecedented experimental capabilities. This thesis is an attempt to contribute to the huge ongoing effort to learn more about the Universe.

We began by developing the necessary statistical techniques to perform rigorous forecasting estimates without sacrificing computational efficiency. For this we borrowed techniques from the field of information geometry, utilising the structure of likelihood spaces to form more efficiently assess how sensitive a detector will be to signs of new physics. Due to the general applicability of these techniques, they will have far reaching consequences for the community. In particular, they will allow for a more general rigorous discussion on experimental design and optimal running procedures.

Initially, we focused on how to efficiently compute upper limits and the discovery reach for experiments that may, a priori, be in the Gaussian of Poissonian regime. Traditionally, the lack of an analytic description of the test statistic in the Poisson regime has proven to be numerically taxing. We alleviate this problem by developing the equivalent signals technique; a method for calculating (to a high degree of accuracy) the exclusion limits and discovery reach of any counting experiment, regardless of the statistical regime. We showed the validity of the method by analysing a set of background and signal models that are common in astro-particle physics.

We then turned our attention to forecasting the discriminating power of future experiments. Again, computational overhead has prevented any rigorous analyses of model discrimination for many models. Moreover, within any given model one is generally required to consider a set of limited benchmark parameter points. For example, if one wants to consider the degree to which N points can be discriminated even in Gaussian regime this would require N^2 matrix inversions. If one considers instead a Poissonian scenario, the situation becomes worse due to the lack of any analytic description of threshold values on the test statistic. To reduce the computational overhead to just N matrix inversions we developed the Euclideanised signals technique;

a mapping of the signals, backgrounds, and their associated errors onto a new vector called the Euclideanised signal. Importantly, the test statistic between two points in the parameter space is now approximated by the Euclidean distance between two of the Euclideanised signals. We can therefore utilise clustering algorithms to compute the distance between all points under consideration with little computational overhead. In particular, the efficiency gains allow for a benchmark-free analysis of the parameter space.

We then turned to the direct detection of dark matter. Here we demonstrated the general applicability of our methods, whilst introducing a novel way to characterise the degeneracy breaking ability of future experiments through *infometric venn diagrams*. We used these to show, for the first time, that if future direct detection experiments find dark matter, they are very unlikely to simultaneously constrain its mass and interaction type. For this we studied the non-relativistic effective theory (NREFT) of dark matter direct detection. This NREFT characterizes all possible elastic contact interactions for direct detection. We showed that for high mass dark matter, the recoil spectra of these interactions become distinct whereas there is no dependence the dark matter mass. For low mass dark matter, the opposite is true. We are therefore left with only a small range of parameters for which it will be possible to disentangle the mass and interaction type. The methods developed here are general, and can be applied to any general counting experiment.

Next we turned to the use of ancient minerals as experimental probes of new physics. These minerals are called *paleo-detectors* and can be used to search for both dark matter by looking for the small deposits of energy from particle interactions. Unlike direct detection, we see the accumulation of one billion years of these events. We showed that despite a significant background from natural radioactivity, it is still possible to find the tracks left by dark matter. Moreover, paleo-detectors would represent the most sensitive dark matter detectors to date. In addition, we showed that their forecast sensitivity is robust to systematic uncertainty in the backgrounds. Finally, we demonstrated that paleo-detectors could constrain the mass of a dark matter particle to much larger masses than a conventional direct detection experiments.

Paleo-detectors are also sensitive neutrinos, the most mysterious sector of the standard model. The neutrinos emitted during stellar explosions, known as supernovae (SNe), are of particular interest. Not only are the processes governing these SN explosions not well understood, the SN themselves play a fundamental role in galaxy evolution and may even be responsible for mass extinction events here on Earth. We showed that paleo-detectors could detect the cumulative emission of neutrinos from SN explosions within our galaxy. In addition, we showed that they are sensitive to particular SN histories providing a potential tracer of star formation in our galaxy. Finally, we also showed that paleo-detectors could provide evidence that a SN occurred near to the Earth in the past one billion years.

Our analyses of paleo-detectors sensitivity represents a primary example of applicability of the statistical methods we developed in Chapter 4. In particular, our use of `swordfish` (the python analysis package associated with the statistical methods) drastically reduced the complexity of our analyses. In addition, the Euclideanised signals technique allowed us to make concrete statements about the ability of these detectors to constrain the dark matter's mass and the degree to which they could discriminate different star formation histories.

8.1 Outlook

Future searches for DM will take a variety of forms, utilising experimental and theoretical progress. In particular, the advent multi-messenger astrophysics will play a pivotal role in searches for BSM physics. Gravitational waves (GWs) and their EM counterparts provide a deep observational window into the most violent environments in the Universe. The possible effects of DM on these signals are still poorly understood, although some progress has been made [427, 428, 429, 430]. A comprehensive study of the effects of DM models (and other motivated BSM candidates) on these multi-messenger signals may provide a new pathway to discover the particle nature of DM.

WIMPs have been the dominant DM paradigm for several decades due to there simple production mechanism in the early universe and generic compatibility with theories of supersymmetry (SUSY). With no detections made to date and lack of clear future progress on the energy frontier from future colliders, this dominance is beginning to wane. Although the motivations of WIMP DM remain true, theorists are looking for new methods to find non-WIMP candidates. Growing in popularity is axion DM, with a particular focus on the QCD axion. Whilst generic BSM particles can be studied, the most well motivated examples of DM candidates come from solutions to (initially) unassociated particle physics problems. Theorists should seek to find solutions to particle physics issues that provide generic DM candidates, whilst experimentalists should develop methods to find them. The reign of the WIMP may be over, but the path ahead is wide open.

Bibliography

[1] T. D. P. Edwards and C. Weniger, *A Fresh Approach to Forecasting in Astroparticle Physics and Dark Matter Searches*, *JCAP* **1802** (2018) 021, [1704.05458].

[2] T. D. P. Edwards, B. J. Kavanagh and C. Weniger, *Dark Matter Model or Mass, but Not Both: Assessing Near-Future Direct Searches with Benchmark-free Forecasting*, *Phys. Rev. Lett.* **121** (2018) 181101, [1805.04117].

[3] T. D. P. Edwards, B. J. Kavanagh, C. Weniger, S. Baum, A. K. Drukier, K. Freese et al., *Digging for dark matter: Spectral analysis and discovery potential of paleo-detectors*, *Phys. Rev.* **D99** (2019) 043541, [1811.10549].

[4] S. Baum, T. D. P. Edwards, B. J. Kavanagh, P. Stengel, A. K. Drukier, K. Freese et al., *Paleo-Detectors for Galactic Supernova Neutrinos*, 1906.05800.

[5] T. D. P. Edwards, M. Chianese, B. J. Kavanagh, S. M. Nissanke and C. Weniger, *A Unique Multi-Messenger Signal of QCD Axion Dark Matter*, 1905.04686.

[6] T. D. P. Edwards and C. Weniger, *swordfish: Efficient Forecasting of New Physics Searches without Monte Carlo*, 1712.05401.

[7] R. T. Bartels, T. D. P. Edwards and C. Weniger, *Bayesian model comparison and analysis of the galactic disc population of gamma-ray millisecond pulsars*, *Mon. Not. Roy. Astron. Soc.* **481** (2018) 3966–3987, [1805.11097].

[8] R. Bartels and T. D. P. Edwards, *Comment on "Understanding the γ-ray emission from the globular cluster 47 Tuc: evidence for dark matter?"*, *PRD (Accepted)* (2018) , [1807.08800].

[9] S. Algeri et al., *Statistical challenges in the search for dark matter*, 1807.09273.

[10] K. Olive, *Review of particle physics*, *Chinese Physics C* **40** (oct, 2016) 100001.

[11] C. M. Will, *The confrontation between general relativity and experiment*, *Living Reviews in Relativity* **17** (Jun, 2014) 4.

[12] P. Astier and R. Pain, *Observational evidence of the accelerated expansion of the universe*, *Comptes Rendus Physique* **13** (2012) 521 – 538.

[13] Supernova Search Team collaboration, A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009–1038, [astro-ph/9805201].

[14] Supernova Cosmology Project collaboration, S. Perlmutter et al., *A Supernova at z = 0.458 and implications for measuring the cosmological deceleration*, *Astrophys. J.* **440** (1995) L41–L44, [astro-ph/9505023].

[15] L. Canetti, M. Drewes and M. Shaposhnikov, *Matter and Antimatter in the Universe*, *New J. Phys.* **14** (2012) 095012, [1204.4186].

[16] K. Freese, *Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars*, *EAS Publ. Ser.* **36** (2009) 113–126, [0812.4005].

[17] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [1007.1727].

[18] XENON collaboration, E. Aprile et al., *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, *Phys. Rev. Lett.* **121** (2018) 111302, [1805.12562].

[19] Fermi-LAT collaboration, M. Ackermann et al., *Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data*, *Phys. Rev. Lett.* **115** (2015) 231301, [1503.02641].

[20] PandaX-II collaboration, X. Cui et al., *Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment*, *Phys. Rev. Lett.* **119** (2017) 181302, [1708.06917].

[21] SuperCDMS collaboration, R. Agnese et al., *Results from the Super Cryogenic Dark Matter Search Experiment at Soudan*, *Phys. Rev. Lett.* **120** (2018) 061802, [1708.08869].

[22] M. Felcini, *Searches for Dark Matter Particles at the LHC*, in *53rd Rencontres de Moriond on Cosmology La Thuile, Italy, March 17-24, 2018*, 2018. 1809.06341.

[23] LUX Collaboration collaboration, D. S. Akerib, S. Alsum, H. M. Araújo, X. Bai, A. J. Bailey, J. Balajthy et al., *Results from a search for dark matter in the complete lux exposure*, *Phys. Rev. Lett.* **118** (Jan, 2017) 021303.

[24] Y. Hochberg, Y. Kahn, M. Lisanti, K. M. Zurek, A. G. Grushin, R. Ilan et al., *Detection of sub-MeV Dark Matter with Three-Dimensional Dirac Materials*, *Phys. Rev.* **D97** (2018) 015004, [1708.08929].

[25] R. M. Geilhufe, B. Olsthoorn, A. Ferella, T. Koski, F. Kahlhoefer, J. Conrad et al., *Materials Informatics for Dark Matter Detection*, 1806.06040.

[26] M. J. Dolan, F. Kahlhoefer and C. McCabe, *Directly detecting sub-GeV dark matter with electrons from nuclear scattering*, *Phys. Rev. Lett.* **121** (2018) 101801, [1711.09906].

[27] R. Essig, J. Mardon and T. Volansky, *Direct Detection of Sub-GeV Dark Matter*, *Phys. Rev.* **D85** (2012) 076007, [1108.5383].

[28] G. Bhattacharyya, *A Pedagogical Review of Electroweak Symmetry Breaking Scenarios*, *Rept. Prog. Phys.* **74** (2011) 026201, [0910.5095].

[29] H.S.Chia. 'Conversation with H.S.Chia', 2019.

[30] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, 1807.06209.

[31] C. Mariani, *Review of Reactor Neutrino Oscillation Experiments*, *Mod. Phys. Lett.* **A27** (2012) 1230010, [1201.6665].

[32] T. Kajita, E. Kearns and M. Shiozawa, *Establishing atmospheric neutrino oscillations with super-kamiokande*, *Nuclear Physics B* **908** (2016) 14 – 29.

[33] SNO Collaboration collaboration, Q. R. Ahmad, R. C. Allen, T. C. Andersen, J. D.Anglin, J. C. Barton, E. W. Beier et al., *Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory*, *Phys. Rev. Lett.* **89** (Jun, 2002) 011301.

[34] Super-Kamiokande collaboration, Y. Ashie et al., *Evidence for an oscillatory signature in atmospheric neutrino oscillation*, *Phys. Rev. Lett.* **93** (2004) 101801, [hep-ex/0404034].

[35] R. Henning, *Current status of neutrinoless double-beta decay searches*, *Reviews in Physics* **1** (2016) 29 – 35.

[36] M. C. Chen, *Neutrinoless double beta decay experiments*, *Journal of Physics: Conference Series* **598** (apr, 2015) 012008.

[37] L. Cardani, *Neutrinoless Double Beta Decay Overview*, *SciPost Phys. Proc.* **1** (2019) 024, [1810.12828].

[38] P. Ntelis, *The Homogeneity Scale of the universe*, 2016. 1607.03418.

[39] F. Avila, C. P. Novaes, A. Bernui and E. de Carvalho, *The scale of homogeneity in the local Universe with the ALFALFA catalogue*, *JCAP* **1812** (2018) 041, [1806.04541].

[40] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13, [1502.01589].

[41] S. Dodelson, *Modern Cosmology*. Academic Press, Elsevier Science, 2003.

[42] H. Sun, *Dark matter searches in jet plus missing energy events in γp collisions at the CERN LHC*, *Phys. Rev.* **D90** (2014) 035018, [1407.5356].

[43] F. Kahlhoefer, *Review of LHC Dark Matter Searches*, *Int. J. Mod. Phys.* **A32** (2017) 1730006, [1702.02430].

[44] UA1 collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, *Phys. Lett.* **B126** (1983) 398–410.

[45] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [1207.7214].

[46] UA1 collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV*, *Phys. Lett.* **B122** (1983) 103–116.

[47] D. J. Bird et al., *Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation*, *Astrophys. J.* **441** (1995) 144–150, [astro-ph/9410067].

[48] H. Chen, C. Dai, L. Ding, Y. Guo, A. Huo, C. Jing et al., *Search for new massive particles in cosmic rays*, *Physics Reports* **282** (1997) 1 – 34.

[49] C. W. Walter, *The Super-Kamiokande Experiment*, 0802.1041.

[50] F. Halzen and S. R. Klein, *IceCube: An Instrument for Neutrino Astronomy*, *Rev. Sci. Instrum.* **81** (2010) 081101, [1007.1247].

[51] P. A. Čerenkov, *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*, *Phys. Rev.* **52** (Aug, 1937) 378–379.

[52] Super-Kamiokande collaboration, K. Abe et al., *Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton · years exposure of the Super-Kamiokande water Cherenkov detector*, *Phys. Rev.* **D95** (2017) 012004, [1610.03597].

[53] J. C. Pati and A. Salam, *Is baryon number conserved?*, *Phys. Rev. Lett.* **31** (Sep, 1973) 661–664.

[54] G. G. Raffelt, *Stars as laboratories for fundamental physics*. 1996.

[55] B. Wallisch, *Cosmological Probes of Light Relics*. PhD thesis, Cambridge U., DAMTP, 2018. 1810.02800. 10.17863/CAM.30368.

[56] D. Baumann, D. Green and B. Wallisch, *Searching for light relics with large-scale structure*, *JCAP* **1808** (2018) 029, [1712.08067].

[57] B. R. Safdi, Z. Sun and A. Y. Chen, *Detecting Axion Dark Matter with Radio Lines from Neutron Star Populations*, 1811.01020.

[58] A. Hook, Y. Kahn, B. R. Safdi and Z. Sun, *Radio Signals from Axion Dark Matter Conversion in Neutron Star Magnetospheres*, *Phys. Rev. Lett.* **121** (2018) 241102, [1804.03145].

[59] Y. B. Zel'Dovich, *Generation of Waves by a Rotating Body*, *Soviet Journal of Experimental and Theoretical Physics Letters* **14** (Aug., 1971) 180.

[60] R. Brito, V. Cardoso and P. Pani, *Superradiance*, *Lect. Notes Phys.* **906** (2015) pp.1–237, [1501.06570].

[61] D. Baumann, H. S. Chia and R. A. Porto, *Probing Ultralight Bosons with Binary Black Holes*, *Phys. Rev.* **D99** (2019) 044001, [1804.03208].

[62] G. Bertone and D. Hooper, *History of dark matter*, *Rev. Mod. Phys.* **90** (2018) 045002, [1605.04909].

[63] U. L. Verrier, "Comptes rendus hebdomadaires des séances de l'académie des sciences." https://archive.org/stream/comptesrendusheb49acad#page/378/mode/2up.

[64] A. Einstein, *Die grundlage der allgemeinen relativitÃ€tstheorie*, *Annalen der Physik* **354** (1916) 769–822, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163540702].

[65] *Forthcoming article*, *The Astrophysical Journal* (2019) .

[66] J. Binney and S. Tremaine, *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.

[67] P. Salucci, *The mass distribution in spiral galaxies*, *IAU Symp.* **244** (2008) 53–62, [0707.4370].

[68] O. Gerhard, *Mass distribution in our galaxy*, *Space Sci. Rev.* **100** (2002) 129, [astro-ph/0203110].

[69] V. C. Rubin, W. K. Ford, Jr. and N. Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/*, *ApJ* **238** (June, 1980) 471–487.

[70] A. Bosma, *21-cm line studies of spiral galaxies. II. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.*, *AJ* **86** (Dec., 1981) 1825–1846.

[71] M. Persic, P. Salucci and F. Stel, *The Universal rotation curve of spiral galaxies: 1. The Dark matter connection*, *Mon. Not. Roy. Astron. Soc.* **281** (1996) 27, [astro-ph/9506004].

[72] J. I. Read, *The Local Dark Matter Density*, *J. Phys.* **G41** (2014) 063101, [1404.1938].

[73] Y. Shan, M. McDonald and S. Courteau, *Revised Mass-to-light Ratios for Nearby Galaxy Groups and Clusters*, *ApJ* **800** (Feb, 2015) 122, [1412.8474].

[74] M. Markevitch, *Chandra observation of the most interesting cluster in the universe*, astro-ph/0511345.

[75] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones et al., *A direct empirical proof of the existence of dark matter*, *Astrophys. J.* **648** (2006) L109–L113, [astro-ph/0608407].

[76] M. Bartelmann and M. Maturi, *Weak gravitational lensing*, 2016. 1612.06535.

[77] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, *Phys. Rept.* **340** (2001) 291–472, [astro-ph/9912508].

[78] S. D. M. White, C. S. Frenk and M. Davis, *Clustering in a neutrino-dominated universe*, *ApJ* **274** (Nov., 1983) L1–L5.

[79] J. de Swart, G. Bertone and J. van Dongen, *How Dark Matter Came to Matter*, 1703.00013.

[80] J. R. Bond and A. S. Szalay, *The collisionless damping of density fluctuations in an expanding universe*, *ApJ* **274** (Nov., 1983) 443–468.

[81] J. R. Bond, G. Efstathiou and J. Silk, *Massive neutrinos and the large-scale structure of the universe*, *Phys. Rev. Lett.* **45** (Dec, 1980) 1980–1984.

[82] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Can sterile neutrinos be ruled out as warm dark matter candidates?*, *Phys. Rev. Lett.* **97** (2006) 071301, [astro-ph/0605706].

[83] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Constraining warm dark matter candidates including sterile neutrinos and light gravitinos with WMAP and the Lyman-alpha forest*, *Phys. Rev.* **D71** (2005) 063534, [astro-ph/0501562].

[84] Planck collaboration, Y. Akrami et al., *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, 1807.06205.

[85] W. Hu and S. Dodelson, *Cosmic microwave background anisotropies*, *Annual Review of Astronomy and Astrophysics* **40** (2002) 171–216, [https://doi.org/10.1146/annurev.astro.40.060401.093926].

[86] MACHO collaboration, C. Alcock et al., *The MACHO project: Microlensing results from 5.7 years of LMC observations*, *Astrophys. J.* **542** (2000) 281–307, [astro-ph/0001272].

[87] EROS collaboration, T. Lasserre, *Not enough stellar mass machos in the galactic halo*, *Astron. Astrophys.* **355** (2000) L39–L42, [astro-ph/0002253].

[88] EROS-2 collaboration, P. Tisserand et al., *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, *Astron. Astrophys.* **469** (2007) 387–404, [astro-ph/0607207].

[89] LIGO Scientific Collaboration and Virgo Collaboration collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley et al., *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (Feb, 2016) 061102.

[90] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz et al., *Did LIGO detect dark matter?*, *Phys. Rev. Lett.* **116** (2016) 201301, [1603.00464].

[91] A. Hektor, G. Hütsi and M. Raidal, *Constraints on primordial black hole dark matter from Galactic center X-ray observations*, *Astron. Astrophys.* **618** (2018) A139, [1805.06513].

[92] B. J. Kavanagh, D. Gaggero and G. Bertone, *Merger rate of a subdominant population of primordial black holes*, *Phys. Rev.* **D98** (2018) 023536, [1805.09034].

[93] D. Gaggero, G. Bertone, F. Calore, R. M. T. Connors, M. Lovell, S. Markoff et al., *Searching for Primordial Black Holes in the radio and X-ray sky*, *Phys. Rev. Lett.* **118** (2017) 241101, [1612.00457].

[94] N. Arkani-Hamed, "Robustness of gr. attempts to modify gravity, part 1."

[95] N. Arkani-Hamed, "Robustness of gr. attempts to modify gravity, part 2."

[96] N. Arkani-Hamed, "Robustness of gr. attempts to modify gravity, part 3."

[97] N. Arkani-Hamed, "Robustness of gr. attempts to modify gravity, part 4."

[98] S. Weinberg, *Photons and gravitons in perturbation theory: Derivation of maxwell's and einstein's equations*, *Phys. Rev.* **138** (May, 1965) B988–B1002.

[99] E. Wigner, *On unitary representations of the inhomogeneous lorentz group*, *Annals of Mathematics* **40** (1939) 149–204.

[100] T. Kobayashi, *Horndeski theory and beyond: a review*, 1901.07183.

[101] S. M. Carroll, *Lecture notes on general relativity*, gr-qc/9712019.

[102] D. Lovelock, *The fourâdimensionality of space and the einstein tensor*, *Journal of Mathematical Physics* **13** (1972) 874–876, [https://doi.org/10.1063/1.1666069].

[103] S. S. McGaugh, J. M. Schombert, G. D. Bothun and W. J. G. de Blok, *The Baryonic Tully-Fisher relation*, *Astrophys. J.* **533** (2000) L99–L102, [astro-ph/0003001].

[104] P. Mollitor, E. Nezri and R. Teyssier, *Baryonic and dark matter distribution in cosmological simulations of spiral galaxies*, *Mon. Not. Roy. Astron. Soc.* **447** (2015) 1353–1369, [1405.4318].

[105] G. R. Blumenthal, S. M. Faber, J. R. Primack and M. J. Rees, *Formation of Galaxies and Large Scale Structure with Cold Dark Matter*, *Nature* **311** (1984) 517–525.

[106] L. Roszkowski, E. M. Sessolo and S. Trojanowski, *WIMP dark matter candidates and searches—current status and future prospects*, *Rept. Prog. Phys.* **81** (2018) 066201, [1707.06277].

[107] J. L. Feng, *Dark Matter Candidates from Particle Physics and Methods of Detection*, *Ann. Rev. Astron. Astrophys.* **48** (2010) 495–545, [1003.0904].

[108] L. Bergstrom, *Dark Matter Candidates*, *New J. Phys.* **11** (2009) 105006, [0903.4849].

[109] R. J. Scherrer and M. S. Turner, *On the relic, cosmic abundance of stable, weakly interacting massive particles*, *Phys. Rev. D* **33** (Mar, 1986) 1585–1589.

[110] N. Bernal, A. Chatterjee and A. Paul, *Non-thermal production of Dark Matter after Inflation*, *JCAP* **1812** (2018) 020, [1809.02338].

[111] N. Bernal, M. Heikinheimo, T. Tenkanen, K. Tuominen and V. Vaskonen, *The Dawn of FIMP Dark Matter: A Review of Models and Constraints*, *Int. J. Mod. Phys.* **A32** (2017) 1730023, [1706.07442].

[112] D. J. E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [1510.07633].

[113] A. Diez-Tejedor and D. J. E. Marsh, *Cosmological production of ultralight dark matter axions*, 1702.02116.

[114] D. J. E. Marsh, *Axions and ALPs: a very short introduction*, in *Proceedings, 13th Patras Workshop on Axions, WIMPs and WISPs, (PATRAS 2017): Thessaloniki, Greece, 15 May 2017 - 19, 2017*, pp. 59–74, 2018. 1712.03018. DOI.

[115] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356.

[116] "Daniel baumann." https://www.dropbox.com/s/5slirav7yd9mocf/Chapter3.pdf?dl=0.

[117] C. A. Baker et al., *An Improved experimental limit on the electric dipole moment of the neutron*, *Phys. Rev. Lett.* **97** (2006) 131801, [hep-ex/0602020].

[118] J. M. Pendlebury et al., *Revised experimental upper limit on the electric dipole moment of the neutron*, *Phys. Rev.* **D92** (2015) 092003, [1509.04411].

[119] K. J. Bae, J.-H. Huh and J. E. Kim, *Update of axion CDM energy*, *JCAP* **0809** (2008) 005, [0806.0497].

[120] M. W. Goodman and E. Witten, *Detectability of certain dark-matter candidates*, *Phys. Rev. D* **31** (Jun, 1985) 3059–3063.

[121] A. K. Drukier, K. Freese and D. N. Spergel, *Detecting cold dark-matter candidates*, *Phys. Rev. D* **33** (Jun, 1986) 3495–3508.

[122] J. Fan, M. Reece and L.-T. Wang, *Non-relativistic effective theory of dark matter direct detection*, *JCAP* **1011** (2010) 042, [1008.1591].

[123] A. L. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers and Y. Xu, *The Effective Field Theory of Dark Matter Direct Detection*, *JCAP* **1302** (2013) 004, [1203.3542].

[124] P. Bhattacharjee, S. Chaudhury, S. Kundu and S. Majumdar, *Sizing-up the WIMPs of Milky Way : Deriving the velocity distribution of Galactic Dark Matter particles from the rotation curve data*, *Phys. Rev.* **D87** (2013) 083525, [1210.2328].

[125] P. Panci, *New Directions in Direct Dark Matter Searches*, *Adv. High Energy Phys.* **2014** (2014) 681312, [1402.1507].

[126] M. Cirelli, E. Del Nobile and P. Panci, *Tools for model-independent bounds in direct dark matter searches*, *JCAP* **1310** (2013) 019, [1307.5955].

[127] T. Lin, *TASI lectures on dark matter models and direct detection*, 1904.07915.

[128] M. Lisanti, *Lectures on Dark Matter Physics*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 399–446, 2017. 1603.03797. DOI.

[129] S. K. Lee, M. Lisanti, S. Mishra-Sharma and B. R. Safdi, *Modulation Effects in Dark Matter-Electron Scattering Experiments*, *Phys. Rev.* **D92** (2015) 083517, [1508.07361].

[130] R. Essig, T. Volansky and T.-T. Yu, *New Constraints and Prospects for sub-GeV Dark Matter Scattering off Electrons in Xenon*, *Phys. Rev.* **D96** (2017) 043017, [1703.00910].

[131] R. H. Helm, *Inelastic and elastic scattering of 187-mev electrons from selected even-even nuclei*, *Phys. Rev.* **104** (Dec, 1956) 1466–1475.

[132] A. Münster et al., *Radiopurity of CaWO4 crystals for direct dark matter search with CRESST and EURECA*, *JCAP* **1405** (2014) 018, [1403.5114].

[133] DAMA collaboration, R. Bernabei et al., *The DAMA/LIBRA apparatus*, *Nucl. Instrum. Meth.* **A592** (2008) 297–315, [0804.2738].

[134] T. Marrodán Undagoitia and L. Rauch, *Dark matter direct-detection experiments*, *J. Phys.* **G43** (2016) 013001, [1509.08767].

[135] J. Cooley, *Overview of Non-Liquid Noble Direct Detection Dark Matter Experiments*, *Phys. Dark Univ.* **4** (2014) 92–97, [1410.4960].

[136] LUX collaboration, D. S. Akerib et al., *Results from a search for dark matter in the complete LUX exposure*, *Phys. Rev. Lett.* **118** (2017) 021303, [1608.07648].

[137] C. E. Aalseth et al., *DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS*, *Eur. Phys. J. Plus* **133** (2018) 131, [1707.08145].

[138] XENON Collaboration 7 collaboration, E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, L. Althueser, F. D. Amaro et al., *Dark matter search results from a one ton-year exposure of xenon1t*, *Phys. Rev. Lett.* **121** (Sep, 2018) 111302.

[139] F. Ruppin, J. Billard, E. Figueroa-Feliciano and L. Strigari, *Complementarity of dark matter detectors in light of the neutrino background*, *Phys. Rev. D* **90** (Oct, 2014) 083510.

[140] K. Cranmer, *Practical Statistics for the LHC*, in *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011): Cheile Gradistei, Romania, September 7-20, 2011*, pp. 267–308, 2015. 1503.07622. DOI.

[141] G. Casella and R. L. Berger, *Statistical Inference*. 2002.

[142] M. Ghosh, *Objective Priors: An Introduction for Frequentists*, *arXiv e-prints* (Aug, 2011) arXiv:1108.2120, [1108.2120].

[143] F. James, *Statistical methods in experimental physics*. 2006.

[144] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys. Rev.* **D57** (1998) 3873–3889, [physics/9711021].

[145] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Annals Math. Statist.* **9** (1938) 60–62.

[146] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, *The Annals of Mathematical Statistics* **9** (1938) 60–62.

[147] S. Yellin, *Finding an upper limit in the presence of an unknown background*, *Phys. Rev. D* **66** (Aug, 2002) 032005, [physics/0203002].

[148] G. Cowan, *Statistical data analysis*. 1998.

[149] H. Chernoff, *On the distribution of the likelihood ratio*, *The Annals of Mathematical Statistics* **25** (1954) 573–578.

[150] L. Wolz, M. Kilbinger, J. Weller and T. Giannantonio, *On the Validity of Cosmological Fisher Matrix Forecasts*, *JCAP* **1209** (2012) 009, [1205.3984].

[151] M. Vallisneri, *Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects*, *Phys. Rev.* **D77** (2008) 042001, [gr-qc/0703086].

[152] E. Sellentin, M. Quartin and L. Amendola, *Breaking the spell of Gaussianity: forecasting with higher order Fisher matrices*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 1831–1840, [1401.6892].

[153] A. Heavens, *Generalisations of Fisher Matrices*, 1606.06455.

[154] J. Brehmer, K. Cranmer, F. Kling and T. Plehn, *Better Higgs Measurements Through Information Geometry*, 1612.05261.

[155] S. Camera, M. Fornasa, N. Fornengo and M. Regis, *Tomographic-spectral approach for dark matter detection in the cross-correlation between cosmic shear and diffuse γ-ray emission*, *JCAP* **1506** (2015) 029, [1411.4651].

[156] T. Bringmann, A. Galea, A. Hryczuk and C. Weniger, *Novel Spectral Features in MeV Gamma Rays from Dark Matter*, 1610.04613.

[157] R. Bartels, D. Gaggero and C. Weniger, *Prospects for indirect dark matter searches with MeV photons*, 1703.02546.

[158] Fermi-LAT collaboration, E. Charles et al., *Sensitivity Projections for Dark Matter Searches with the Fermi Large Area Telescope*, *Phys. Rept.* **636** (2016) 1–46, [1605.02016].

[159] Fermi-LAT collaboration, A. A. Abdo et al., *The Second Fermi Large Area Telescope Catalog of Gamma-ray Pulsars*, *Astrophys. J. Suppl.* **208** (2013) 17, [1305.4385].

[160] Fermi-LAT collaboration, A. Albert, G. A. Gomez-Vargas, M. Grefe, C. Munoz, C. Weniger, E. D. Bloom et al., *Search for 100 MeV to 10 GeV γ-ray lines in the Fermi-LAT data and implications for gravitino dark matter in $\mu\nu$SSM*, *JCAP* **1410** (2014) 023, [1406.3430].

[161] C. Radhakrishna Rao, *Information and the accuracy attainable in the estimation of statistical parameters*, vol. 37. 1945.

[162] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*, vol. 9. Princeton university press, 2016.

[163] S. N. Chiu, D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[164] J.-F. Coeurjolly and J. Møller, *Variational approach for spatial point process intensity estimation*, 1407.0249v1.

[165] J. Neyman, *Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability*, *Phil. Trans. Roy. Soc. Lond.* **A236** (1937) 333–380.

[166] E. Sellentin, M. Quartin and L. Amendola, *Breaking the spell of Gaussianity: forecasting with higher order Fisher matrices*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 1831–1840, [1401.6892].

[167] W. A. Rolke, A. M. Lopez and J. Conrad, *Limits and confidence intervals in the presence of nuisance parameters*, *Nucl. Instrum. Meth.* **A551** (2005) 493–503, [physics/0403059].

[168] G. Jungman, M. Kamionkowski and K. Griest, *Supersymmetric dark matter*, *Phys. Rept.* **267** (1996) 195–373, [hep-ph/9506380].

[169] G. Bertone, D. Hooper and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, *Phys. Rept.* **405** (2005) 279–390, [hep-ph/0404175].

[170] M. W. Goodman and E. Witten, *Detectability of Certain Dark Matter Candidates*, *Phys. Rev.* **D31** (1985) 3059.

[171] A. K. Drukier, K. Freese and D. N. Spergel, *Detecting Cold Dark Matter Candidates*, *Phys. Rev.* **D33** (1986) 3495–3508.

[172] XENON collaboration, E. Aprile et al., *First Dark Matter Search Results from the XENON1T Experiment*, *Phys. Rev. Lett.* **119** (2017) 181301, [1705.06655].

[173] A. H. G. Peter, V. Gluscevic, A. M. Green, B. J. Kavanagh and S. K. Lee, *WIMP physics with ensembles of direct-detection experiments*, *Phys. Dark Univ.* **5-6** (2014) 45–74, [1310.7039].

[174] B. J. Kavanagh, M. Fornasa and A. M. Green, *Probing WIMP particle physics and astrophysics with direct detection and neutrino telescope data*, *Phys. Rev.* **D91** (2015) 103533, [1410.8051].

[175] M. Cahill-Rowley, R. Cotta, A. Drlica-Wagner, S. Funk, J. Hewett, A. Ismail et al., *Complementarity of dark matter searches in the phenomenological MSSM*, *Phys. Rev.* **D91** (2015) 055011, [1405.6716].

[176] A. Alves, A. Berlin, S. Profumo and F. S. Queiroz, *Dark Matter Complementarity and the Z' Portal*, *Phys. Rev.* **D92** (2015) 083004, [1501.03490].

[177] G. Bertone, N. Bozorgnia, J. S. Kim, S. Liem, C. McCabe, S. Otten et al., *Identifying WIMP dark matter from particle and astroparticle data*, *JCAP* **1803** (2018) 026, [1712.04793].

[178] R. Catena, *Prospects for direct detection of dark matter in an effective theory approach*, *JCAP* **1407** (2014) 055, [1406.0524].

[179] M. I. Gresham and K. M. Zurek, *Effect of nuclear response functions in dark matter direct detection*, *Phys. Rev.* **D89** (2014) 123521, [1401.3739].

[180] V. Gluscevic and A. H. G. Peter, *Understanding WIMP-baryon interactions with direct detection: A Roadmap*, *JCAP* **1409** (2014) 040, [1406.7008].

[181] V. Gluscevic, M. I. Gresham, S. D. McDermott, A. H. G. Peter and K. M. Zurek, *Identifying the Theory of Dark Matter with Direct Detection*, *JCAP* **1512** (2015) 057, [1506.04454].

[182] F. Kahlhoefer and S. Wild, *Studying generalised dark matter interactions with extended halo-independent methods*, *JCAP* **1610** (2016) 032, [1607.04418].

[183] R. Catena, J. Conrad, C. Döring, A. D. Ferella and M. B. Krauss, *Dark matter spin determination with directional direct detection experiments*, 1706.09471.

[184] S. Baum, R. Catena, J. Conrad, K. Freese and M. B. Krauss, *Determining Dark Matter properties with a XENONnT/LZ signal and LHC-Run3 mono-jet searches*, 1709.06051.

[185] A. L. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers and Y. Xu, *Model Independent Direct Detection Analyses*, 1211.2818.

[186] N. Anand, A. L. Fitzpatrick and W. C. Haxton, *Weakly interacting massive particle-nucleus elastic scattering response*, *Phys. Rev.* **C89** (2014) 065501, [1308.6288].

[187] J. B. Dent, L. M. Krauss, J. L. Newstead and S. Sabharwal, *General analysis of direct dark matter detection: From microphysics to observational signatures*, *Phys. Rev.* **D92** (2015) 063515, [1505.03117].

[188] R. Catena, *Analysis of the theoretical bias in dark matter direct detection*, *JCAP* **1409** (2014) 049, [1407.0127].

[189] R. Catena and P. Gondolo, *Global fits of the dark matter-nucleon effective interactions*, *JCAP* **1409** (2014) 045, [1405.2637].

[190] R. Catena and P. Gondolo, *Global limits and interference patterns in dark matter direct detection*, *JCAP* **1508** (2015) 022, [1504.06554].

[191] J. Venn, *I. on the diagrammatic and mechanical representation of propositions and reasonings*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **10** (jul, 1880) 1–18.

[192] D. G. Cerdeno and A. M. Green, *Direct detection of WIMPs*, pp. 347–369. 2010. 1002.1912.

[193] S. Chang, A. Pierce and N. Weiner, *Momentum Dependent Dark Matter Scattering*, *JCAP* **1001** (2010) 006, [0908.3192].

[194] R. Catena, *Dark matter directional detection in non-relativistic effective theories*, *JCAP* **1507** (2015) 026, [1505.06441].

[195] B. J. Kavanagh, *New directional signatures from the nonrelativistic effective field theory of dark matter*, *Phys. Rev.* **D92** (2015) 023513, [1505.07406].

[196] E. Del Nobile, G. B. Gelmini and S. J. Witte, *Target dependence of the annual modulation in direct dark matter searches*, *Phys. Rev.* **D91** (2015) 121302, [1504.06772].

[197] E. Del Nobile, G. B. Gelmini and S. J. Witte, *Prospects for detection of target-dependent annual modulation in direct dark matter searches*, *JCAP* **1602** (2016) 009, [1512.03961].

[198] F. D'Eramo, B. J. Kavanagh and P. Panci, *You can hide but you have to run: direct detection with vector mediators*, *JHEP* **08** (2016) 111, [1605.04917].

[199] F. Bishara, J. Brod, B. Grinstein and J. Zupan, *DirectDM: a tool for dark matter direct detection*, 1708.02678.

[200] D. G. Cerdeño et al., *Complementarity of dark matter direct detection: the role of bolometric targets*, *JCAP* **1307** (2013) 028, [1304.1758].

[201] F. Kahlhoefer, S. Kulkarni and S. Wild, *Exploring light mediators with low-threshold direct detection experiments*, *JCAP* **1711** (2017) 016, [1707.08571].

[202] C. M. Ho and R. J. Scherrer, *Anapole Dark Matter*, *Phys. Lett.* **B722** (2013) 341–346, [1211.0503].

[203] E. Del Nobile, G. B. Gelmini, P. Gondolo and J.-H. Huh, *Direct detection of Light Anapole and Magnetic Dipole DM*, *JCAP* **1406** (2014) 002, [1401.4508].

[204] S. D. McDermott, H.-B. Yu and K. M. Zurek, *Turning off the Lights: How Dark is Dark Matter?*, *Phys. Rev.* **D83** (2011) 063509, [1011.2907].

[205] N. Fornengo, P. Panci and M. Regis, *Long-Range Forces in Direct Dark Matter Searches*, *Phys. Rev.* **D84** (2011) 115002, [1108.4661].

[206] K. Sigurdson, M. Doran, A. Kurylov, R. R. Caldwell and M. Kamionkowski, *Dark-matter electric and magnetic dipole moments*, *Phys. Rev.* **D70** (2004) 083501, [astro-ph/0406355].

[207] T. Banks, J.-F. Fortin and S. Thomas, *Direct Detection of Dark Matter Electromagnetic Dipole Moments*, 1007.5515.

[208] J. Bagnasco, M. Dine and S. D. Thomas, *Detecting technibaryon dark matter*, *Phys. Lett.* **B320** (1994) 99–104, [hep-ph/9310290].

[209] T. Banks, J. D. Mason and D. O'Neil, *A Dark matter candidate with new strong interactions*, *Phys. Rev.* **D72** (2005) 043530, [hep-ph/0506015].

[210] V. Barger, W.-Y. Keung and D. Marfatia, *Electromagnetic properties of dark matter: Dipole moments and charge form factor*, *Phys. Lett.* **B696** (2011) 74–78, [1007.4345].

[211] E. Del Nobile, C. Kouvaris, P. Panci, F. Sannino and J. Virkajarvi, *Light Magnetic Dark Matter in Direct Detection Searches*, *JCAP* **1208** (2012) 010, [1203.6652].

[212] B. J. Kavanagh and T. D. P. Edwards, WIMpy_NREFT v1.0 [Computer Software], doi:10.5281/zenodo.1230503. Available at *https://github.com/bradkav/WIMpy_NREFT*, 2018.

[213] R. Catena and B. Schwabe, *Form factors for dark matter capture by the Sun in effective theories*, *JCAP* **1504** (2015) 042, [1501.03729].

[214] A. M. Green, *Astrophysical uncertainties on the local dark matter distribution and direct detection experiments*, *J. Phys.* **G44** (2017) 084001, [1703.10102].

[215] LZ collaboration, D. S. Akerib et al., *LUX-ZEPLIN (LZ) Conceptual Design Report*, 1509.02910.

[216] DARWIN collaboration, J. Aalbers et al., *DARWIN: towards the ultimate dark matter detector*, *JCAP* **1611** (2016) 017, [1606.07001].

[217] XENON collaboration, G. Plante. https://conferences.pa.ucla.edu/dm16/talks/plante.pdf, 2016.

[218] LUX-ZEPLIN collaboration, D. S. Akerib et al., *Projected WIMP sensitivity of the LUX-ZEPLIN (LZ) dark matter experiment*, 1802.06039.

[219] J. D. Lewin and P. F. Smith, *Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil*, *Astropart. Phys.* **6** (1996) 87–112.

[220] A. M. Green, *Determining the WIMP mass using direct detection experiments*, *JCAP* **0708** (2007) 022, [hep-ph/0703217].

[221] A. M. Green, *Determining the WIMP mass from a single direct detection experiment, a more detailed study*, *JCAP* **0807** (2008) 005, [0805.1704].

[222] J. L. Newstead, T. D. Jacques, L. M. Krauss, J. B. Dent and F. Ferrer, *Scientific reach of multiton-scale dark matter direct detection experiments*, *Phys. Rev.* **D88** (2013) 076011, [1306.3244].

[223] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825–2830.

[224] E. W. Weisstein, "Hypersphere packing."

[225] J. Bovy et al., *The Milky Way's circular velocity curve between 4 and 14 kpc from APOGEE data*, *Astrophys. J.* **759** (2012) 131, [1209.0759].

[226] S. E. Koposov, H.-W. Rix and D. W. Hogg, *Constraining the Milky Way potential with a 6-D phase-space map of the GD-1 stellar stream*, *Astrophys. J.* **712** (2010) 260–273, [0907.1085].

[227] T. Piffl et al., *The RAVE survey: the Galactic escape speed and the mass of the Milky Way*, *Astron. Astrophys.* **562** (2014) A91, [1309.4293].

[228] A. H. G. Peter, *Getting the astrophysics and particle physics of dark matter out of next-generation direct detection experiments*, *Phys. Rev.* **D81** (2010) 087301, [0910.4765].

[229] L. E. Strigari and R. Trotta, *Reconstructing WIMP Properties in Direct Detection Experiments Including Galactic Dark Matter Distribution Uncertainties*, *JCAP* **0911** (2009) 019, [0906.5361].

[230] M. Pato, L. E. Strigari, R. Trotta and G. Bertone, *Taming astrophysical bias in direct dark matter searches*, *JCAP* **1302** (2013) 041, [1211.7063].

[231] A. M. Green, *Astrophysical uncertainties on direct detection experiments*, *Mod. Phys. Lett.* **A27** (2012) 1230004, [1112.0524].

[232] M. Schumann, L. Baudis, L. Bütikofer, A. Kish and M. Selvi, *Dark matter sensitivity of multi-ton liquid xenon detectors*, *JCAP* **1510** (2015) 016, [1506.08309].

[233] J. L. Feng, J. Kumar, D. Marfatia and D. Sanford, *Isospin-Violating Dark Matter*, *Phys. Lett.* **B703** (2011) 124–127, [1102.4331].

[234] DarkSide collaboration, P. Agnes et al., *First Results from the DarkSide-50 Dark Matter Experiment at Laboratori Nazionali del Gran Sasso*, *Phys. Lett.* **B743** (2015) 456–466, [1410.0653].

[235] A. Drukier and L. Stodolsky, *Principles and Applications of a Neutral Current Detector for Neutrino Physics and Astronomy*, *Phys. Rev.* **D30** (1984) 2295.

[236] D. N. Spergel, *The Motion of the Earth and the Detection of Wimps*, *Phys. Rev.* **D37** (1988) 1353.

[237] DEAP-3600 collaboration, P. A. Amaudruz et al., *First results from the DEAP-3600 dark matter search with argon at SNOLAB*, *Phys. Rev. Lett.* **121** (2018) 071801, [1707.08042].

[238] DarkSide collaboration, P. Agnes et al., *DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon*, *Phys. Rev.* **D98** (2018) 102006, [1802.07198].

[239] EDELWEISS collaboration, L. Hehn et al., *Improved EDELWEISS-III sensitivity for low-mass WIMPs using a profile likelihood approach*, *Eur. Phys. J.* **C76** (2016) 548, [1607.03367].

[240] SuperCDMS collaboration, R. Agnese et al., *Low-mass dark matter search with CDMSlite*, *Phys. Rev.* **D97** (2018) 022002, [1707.01632].

[241] CRESST collaboration, G. Angloher et al., *Results on MeV-scale dark matter from a gram-scale cryogenic calorimeter operated above ground*, *Eur. Phys. J.* **C77** (2017) 637, [1707.06749].

[242] CRESST collaboration, F. Petricca et al., *First results on low-mass dark matter from the CRESST-III experiment*, in *15th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2017) Sudbury, Ontario, Canada, July 24-28, 2017*, 2017. 1711.07692.

[243] SuperCDMS collaboration, R. Agnese et al., *Search for Low-Mass Dark Matter with CDMSlite Using a Profile Likelihood Fit*, 1808.09098.

[244] DAMA collaboration, R. Bernabei et al., *First results from DAMA/LIBRA and the combined results with DAMA/NaI*, *Eur. Phys. J.* **C56** (2008) 333–355, [0804.2741].

[245] R. Bernabei et al., *Final model independent result of DAMA/LIBRA-phase1*, *Eur. Phys. J.* **C73** (2013) 2648, [1308.5109].

[246] R. Bernabei et al., *First model independent results from DAMA/LIBRA-phase2*, *Universe* **4** (2018) 116, [1805.10486].

[247] K. Freese, M. Lisanti and C. Savage, *Colloquium: Annual modulation of dark matter*, *Rev. Mod. Phys.* **85** (2013) 1561–1581, [1209.3339].

[248] P. Gondolo and G. Gelmini, *Compatibility of DAMA dark matter detection with other searches*, *Phys. Rev.* **D71** (2005) 123520, [hep-ph/0504010].

[249] F. Petriello and K. M. Zurek, *DAMA and WIMP dark matter*, *JHEP* **09** (2008) 047, [0806.3989].

[250] S. Chang, A. Pierce and N. Weiner, *Using the Energy Spectrum at DAMA/LIBRA to Probe Light Dark Matter*, *Phys. Rev.* **D79** (2009) 115011, [0808.0196].

[251] M. Fairbairn and T. Schwetz, *Spin-independent elastic WIMP scattering and the DAMA annual modulation signal*, *JCAP* **0901** (2009) 037, [0808.0704].

[252] C. Savage, G. Gelmini, P. Gondolo and K. Freese, *Compatibility of DAMA/LIBRA dark matter detection with other searches*, *JCAP* **0904** (2009) 010, [0808.3607].

[253] A. Bottino, F. Donato, N. Fornengo and S. Scopel, *Interpreting the recent results on direct search for dark matter particles in terms of relic neutralino*, *Phys. Rev.* **D78** (2008) 083520, [0806.4099].

[254] S. Baum, K. Freese and C. Kelso, *Dark Matter implications of DAMA/LIBRA-phase2 results*, 1804.01231.

[255] S. Kang, S. Scopel, G. Tomar and J.-H. Yoon, *DAMA/LIBRA-phase2 in WIMP effective models*, *JCAP* **1807** (2018) 016, [1804.07528].

[256] J. Herrero-Garcia, A. Scaffidi, M. White and A. G. Williams, *Reproducing the DAMA/LIBRA phase-2 results with two dark matter components*, 1804.08437.

[257] C. Savage, K. Freese, P. Gondolo and D. Spolyar, *Compatibility of DAMA/LIBRA dark matter detection with other searches in light of new Galactic rotation velocity measurements*, *JCAP* **0909** (2009) 036, [0901.2713].

[258] C. Savage, G. Gelmini, P. Gondolo and K. Freese, *XENON10/100 dark matter constraints in comparison with CoGeNT and DAMA: examining the Leff dependence*, *Phys. Rev.* **D83** (2011) 055002, [1006.0972].

[259] C. McCabe, *DAMA and CoGeNT without astrophysical uncertainties*, *Phys. Rev.* **D84** (2011) 043525, [1107.0741].

[260] R. Catena, A. Ibarra and S. Wild, *DAMA confronts null searches in the effective theory of dark matter-nucleon interactions*, *JCAP* **1605** (2016) 039, [1602.04074].

[261] G. Angloher et al., *Eureca conceptual design report*, *Physics of the Dark Universe* **3** (2014) 41 – 74.

[262] SuperCDMS collaboration, R. Agnese et al., *Projected Sensitivity of the SuperCDMS SNOLAB experiment*, *Phys. Rev.* **D95** (2017) 082002, [1610.00006].

[263] S. Baum, A. K. Drukier, K. Freese, M. Górski and P. Stengel, *Searching for Dark Matter with Paleo-Detectors*, 1806.05991.

[264] A. K. Drukier, S. Baum, K. Freese, M. Górski and P. Stengel, *Paleo-detectors: Searching for Dark Matter with Ancient Minerals*, 1811.06844.

[265] R. L. Fleischer, P. B. Price, R. M. Walker and E. L. Hubbard, *Track registration in various solid-state nuclear track detectors*, *Phys. Rev.* **133** (Mar, 1964) A1443–A1449.

[266] R. L. Fleischer, P. B. Price and R. M. Walker, *Tracks of charged particles in solids*, *Science* **149** (1965) 383–393.

[267] R. L. Fleischer, P. B. Price and R. M. Walker, *Solid-state track detectors: applications to nuclear science and geophysics*, *Ann. Rev. Nucl. Part. Sci.* **15** (1965) 1–28.

[268] S.-L. Guo, B.-L. Chen and S. Durrani, *Chapter 4 - solid-state nuclear track detectors*, in *Handbook of Radioactivity Analysis (Third Edition)* (M. F. L'Annunziata, ed.), pp. 233 – 298. Academic Press, Amsterdam, third edition ed., 2012. DOI.

[269] J. Kirz and H. Rarback, *Soft x-ray microscopes*, *Review of Scientific Instruments* **56** (jan, 1985) 1–13.

[270] V. V. Lider, *X-ray microscopy*, *Physics-Uspekhi* **60** (2017) 187.

[271] F. Schaff, M. Bech, P. Zaslansky, C. Jud, M. Liebi, M. Guizar-Sicairos et al., *Six-dimensional real and reciprocal space small-angle x-ray scattering tomography*, *Nature* **527** (nov, 2015) 353–356.

[272] R. Hill, J. A. Notte and L. Scipioni, *Scanning helium ion microscopy*, in *Advances in Imaging and Electron Physics* (P. W. Hawkes, ed.), vol. 70, pp. 65–148. Elsevier, 2012. DOI.

[273] P. B. Price, S.-l. Guo, S. P. Ahlen and R. L. Fleischer, *Search for grand-unified-theory magnetic monopoles at a flux level below the parker limit*, *Phys. Rev. Lett.* **52** (Apr, 1984) 1265–1268.

[274] P. B. Price and M. H. Salamon, *Search for supermassive magnetic monopoles using mica crystals*, *Phys. Rev. Lett.* **56** (Mar, 1986) 1226–1229.

[275] J. I. Collar and F. T. Avignone, III, *Nuclear tracks from cold dark matter interactions in mineral crystals: A Computational study*, *Nucl. Instrum. Meth.* **B95** (1995) 349, [astro-ph/9505055].

[276] J. Engel, M. T. Ressell, I. S. Towner and W. E. Ormand, *Response of mica to weakly interacting massive particles*, *Phys. Rev. C* **52** (Oct, 1995) 2216–2221.

[277] D. P. Snowden-Ifft, E. S. Freeman and P. B. Price, *Limits on dark matter using ancient mica*, *Phys. Rev. Lett.* **74** (1995) 4133–4136.

[278] D. P. Snowden-Ifft and M. K. Y. Chan, *A New track etch model for mica*, *Nucl. Instrum. Meth.* **B101** (1995) 247–251.

[279] D. P. Snowden-Ifft and A. J. Westphal, *Unique signature of dark matter in ancient mica*, *Phys. Rev. Lett.* **78** (1997) 1628–1631, [astro-ph/9701215].

[280] E. A. Baltz, A. J. Westphal and D. P. Snowden-Ifft, *Probing the structure of the cold dark matter halo using ancient mica*, *Phys. Rev.* **D59** (1999) 023510, [astro-ph/9711039].

[281] S. M. Stevens, P. Cubillas, K. Jansson, O. Terasaki, M. W. Anderson, P. A. Wright et al., *Nanometre resolution using high-resolution scanning electron microscopy corroborated by atomic force microscopy*, *Chemical Communications* (2008) 3894.

[282] K. W. Urban, *Is science prepared for atomic-resolution electron microscopy?*, *Nature Materials* **8** (apr, 2009) 260–262.

[283] N. de Jonge, N. Poirier-Demers, H. Demers, D. B. Peckys and D. Drouin, *Nanometer-resolution electron microscopy through micrometers-thick water layers*, *Ultramicroscopy* **110** (aug, 2010) 1114–1119.

[284] R. van Gastel, G. Hlawacek, H. J. Zandvliet and B. Poelsema, *Subsurface analysis of semiconductor structures with helium ion microscopy*, *Microelectronics Reliability* **52** (sep, 2012) 2104–2109.

[285] M. S. Joens, C. Huynh, J. M. Kasuboski, D. Ferranti, Y. J. Sigal, F. Zeitvogel et al., *Helium ion microscopy (HIM) for the imaging of biological samples at sub-nanometer resolution*, *Scientific Reports* **3** (dec, 2013) .

[286] M. Rodriguez, W. Li, F. Chen, C. Trautmann, T. Bierschenk, B. Afra et al., *SAXS and TEM investigation of ion tracks in neodymium-doped yttrium aluminium garnet*, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **326** (may, 2014) 150–153.

[287] M. Holler, A. Diaz, M. Guizar-Sicairos, P. Karvinen, E. Färm, E. Härkönen et al., *X-ray ptychographic computed tomography at 16 nm isotropic 3d resolution*, *Scientific Reports* **4** (jan, 2014) .

[288] C. A. O'Hare, *Dark matter astrophysical uncertainties and the neutrino floor*, *Phys. Rev.* **D94** (2016) 063527, [1604.03858].

[289] T. D. P. Edwards and B. J. Kavanagh, "paleo_detectors v1.0 [Computer Software]." doi:10.5281/zenodo.1491777. Available at https://github.com/tedwards2412/paleo_detectors, 2018.

[290] R. H. Helm, *Inelastic and Elastic Scattering of 187-Mev Electrons from Selected Even-Even Nuclei*, *Phys. Rev.* **104** (1956) 1466–1475.

[291] G. Duda, A. Kemper and P. Gondolo, *Model Independent Form Factors for Spin Independent Neutralino-Nucleon Scattering from Elastic Electron Scattering Data*, *JCAP* **0704** (2007) 012, [hep-ph/0608035].

[292] N. W. Evans, C. A. J. O'Hare and C. McCabe, *SHM^{++}: A Refinement of the Standard Halo Model for Dark Matter Searches*, 1810.11468.

[293] J. F. Ziegler, *The stopping and range of ions in solids*. Pergamon, New York, 1985.

[294] J. F. Ziegler, M. D. Ziegler and J. P. Biersack, *SRIM - The stopping and range of ions in matter (2010)*, *Nuclear Instruments and Methods in Physics Research B* **268** (June, 2010) 1818–1823.

[295] W. D. Wilson, L. G. Haggmark and J. P. Biersack, *Calculations of nuclear stopping, ranges, and straggling in the low-energy region*, *Physical Review B* **15** (mar, 1977) 2458–2468.

[296] J. J. Lombardo, R. A. Ristau, W. M. Harris and W. K. S. Chiu, *Focused ion beam preparation of samples for X-ray nanotomography*, *Journal of Synchrotron Radiation* **19** (Sep, 2012) 789–796.

[297] M. P. Echlin, M. Straw, S. Randolph, J. Filevich and T. M. Pollock, *The tribeam system: Femtosecond laser ablation in situ sem*, *Materials Characterization* **100** (2015) 1 – 12.

[298] M. J. Pfeifenberger, M. Mangang, S. Wurster, J. Reiser, A. Hohenwarter, W. Pfleging et al., *The use of femtosecond laser ablation as a novel tool for rapid micro-mechanical sample preparation*, *Materials & Design* **121** (2017) 109 – 118.

[299] S. J. Randolph, J. Filevich, A. Botman, R. Gannon, C. Rue and M. Straw, *In situ femtosecond pulse laser ablation for large volume 3D analysis in scanning electron microscope systems*, *Journal of Vacuum Science & Technology B* **36** (2018) 06JB01.

[300] F. M. Gradstein, J. G. Ogg, M. D. Schmitz and G. M. Ogg, eds., *The Geologic Time Scale*. Elsevier, 2012, 10.1016/c2011-1-08249-8.

[301] K. Gallagher, R. Brown and C. Johnson, *Fission track analysis and its applications to geological problems*, *Annual Review of Earth and Planetary Sciences* **26** (1998) 519–572, [https://doi.org/10.1146/annurev.earth.26.1.519].

[302] P. van den Haute and F. de Corte, eds., *Advances in Fission-Track Geochronology*. Springer, 1998, 10.1007/978-94-015-9133-1.

[303] J. Monroe and P. Fisher, *Neutrino Backgrounds to Dark Matter Searches*, *Phys. Rev.* **D76** (2007) 033007, [0706.3019].

[304] L. E. Strigari, *Neutrino Coherent Scattering Rates at Direct Dark Matter Detectors*, *New J. Phys.* **11** (2009) 105011, [0903.3630].

[305] A. Gutlein et al., *Solar and atmospheric neutrinos: Background sources for the direct dark matter search*, *Astropart. Phys.* **34** (2010) 90–96, [1003.5530].

[306] J. Billard, L. Strigari and E. Figueroa-Feliciano, *Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments*, *Phys. Rev.* **D89** (2014) 023524, [1307.5458].

[307] Borexino collaboration, M. Agostini et al., *Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure*, 1709.00756.

[308] J. F. Beacom, *The Diffuse Supernova Neutrino Background*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 439–462, [1004.3311].

[309] D. Madland, E. Arthur, G. Estes, J. Stewart, M. Bozoian, R. Perry et al., *SOURCES 4a: A code for calculating (alpha, n), spontaneous fission, and delayed neutron sources and spectra*, tech. rep., sep, 1999. 10.2172/15215.

[310] J. Billard, F. Mayet and D. Santos, *Assessing the discovery potential of directional detection of Dark Matter*, *Phys. Rev.* **D85** (2012) 035006, [1110.6079].

[311] L. E. Strigari, *Neutrino floor at ultralow threshold*, *Phys. Rev.* **D93** (2016) 103534, [1604.00729].

[312] A. H. G. Peter, *WIMP astronomy and particle physics with liquid-noble and cryogenic direct-detection experiments*, *Phys. Rev.* **D83** (2011) 125029, [1103.5145].

[313] C. Strege, R. Trotta, G. Bertone, A. H. G. Peter and P. Scott, *Fundamental statistical limitations of future dark matter direct detection experiments*, *Phys. Rev.* **D86** (2012) 023507, [1201.3631].

[314] N. Bozorgnia, D. G. Cerdeno, A. Cheek and B. Penning, *Opening the energy window on direct dark matter detection*, 1810.05576.

[315] A. M. Green, *Determining the WIMP mass from a single direct detection experiment*, *PoS* **IDM2008** (2008) 108, [0809.1904].

[316] P. J. Fox, J. Liu and N. Weiner, *Integrating Out Astrophysical Uncertainties*, *Phys. Rev.* **D83** (2011) 103514, [1011.1915].

[317] B. J. Kavanagh and A. M. Green, *Model independent determination of the dark matter mass from direct detection experiments*, *Phys. Rev. Lett.* **111** (2013) 031302, [1303.6868].

[318] B. Feldstein and F. Kahlhoefer, *A new halo-independent approach to dark matter direct detection analysis*, *JCAP* **1408** (2014) 065, [1403.4606].

[319] G. B. Gelmini, J.-H. Huh and S. J. Witte, *Unified Halo-Independent Formalism From Convex Hulls for Direct Dark Matter Searches*, *JCAP* **1712** (2017) 039, [1707.07019].

[320] A. Ibarra, B. J. Kavanagh and A. Rappelt, *Bracketing the impact of astrophysical uncertainties on local dark matter searches*, 1806.08714.

[321] C. A. J. O'Hare, B. J. Kavanagh and A. M. Green, *Time-integrated directional detection of dark matter*, *Phys. Rev.* **D96** (2017) 083011, [1708.02959].

[322] A. Dekel and J. Silk, *The origin of dwarf galaxies, cold dark matter, and biased galaxy formation*, *Astrophys. J.* **303** (1986) 39–55.

[323] C. W. Akerlof et al., *The ROTSE-III robotic telescope system*, *Publ. Astron. Soc. Pac.* **115** (2003) 132–140, [astro-ph/0210238].

[324] N. M. Law et al., *The Palomar Transient Factory: System Overview, Performance and First Results*, *Publ. Astron. Soc. Pac.* **121** (2009) 1395, [0906.5350].

[325] J. Leaman, W. Li, R. Chornock and A. V. Filippenko, *Nearby Supernova Rates from the Lick Observatory Supernova Search. I. The Methods and Database*, *Mon. Not. Roy. Astron. Soc.* **412** (2011) 1419, [1006.4611].

[326] O. Graur, D. Poznanski, D. Maoz, N. Yasuda, T. Totani, M. Fukugita et al., *Supernovae in the Subaru Deep Field: the rate and delay-time distribution of Type Ia supernovae out to redshift 2*, *Mon. Not. Roy. Astron. Soc.* **417** (2011) 916–940, [1102.0005].

[327] O. Graur et al., *Type-Ia Supernova Rates to Redshift 2.4 from CLASH: the Cluster Lensing And Supernova survey with Hubble*, *Astrophys. J.* **783** (2014) 28, [1310.3495].

[328] SDSS collaboration, M. Sako et al., *The Data Release of the Sloan Digital Sky Survey-II Supernova Survey*, *Publ. Astron. Soc. Pac.* **130** (2018) 064002, [1401.3317].

[329] K. Krisciunas et al., *The Carnegie Supernova Project I: Third Photometry Data Release of Low-Redshift Type Ia Supernovae and Other White Dwarf Explosions*, *Astron. J.* **154** (2017) 211, [1709.05146].

[330] P. Madau and M. Dickinson, *Cosmic Star Formation History*, *Ann. Rev. Astron. Astrophys.* **52** (2014) 415–486, [1403.0007].

[331] L.-G. Strolger, T. Dahlen, S. A. Rodney, O. Graur, A. G. Riess, C. McCully et al., *The Rate of Core Collapse Supernovae to Redshift 2.5 From The CANDELS and CLASH Supernova Surveys*, *Astrophys. J.* **813** (2015) 93, [1509.06574].

[332] T. Petrushevska et al., *High-redshift supernova rates measured with the gravitational telescope A 1689*, *Astron. Astrophys.* **594** (2016) A54, [1607.01617].

[333] W. D. Arnett, J. N. Bahcall, R. P. Kirshner and S. E. Woosley, *SUPERNOVA SN1987A*, *Ann. Rev. Astron. Astrophys.* **27** (1989) 629–700.

[334] D. A. Green and F. R. Stephenson, *The Historical supernovae*, *Lect. Notes Phys.* **598** (2003) 7, [astro-ph/0301603].

[335] E. Cappellaro, R. Barbon and M. Turatto, *Supernova statistics*, *Springer Proc. Phys.* **99** (2005) 347–354, [astro-ph/0310859].

[336] R. Diehl et al., *Radioactive Al-26 and massive stars in the galaxy*, *Nature* **439** (2006) 45–47, [astro-ph/0601015].

[337] A. Strumia and F. Vissani, *Neutrino masses and mixings and...*, *arXiv: hep-ph/0606054* (2006) , [hep-ph/0606054].

[338] W. Li, R. Chornock, J. Leaman, A. V. Filippenko, D. Poznanski, X. Wang et al., *Nearby supernova rates from the Lick Observatory Supernova Search - III. The rate-size relation, and the rates as a function of galaxy Hubble type and colour*, *Mon. Not. Roy. Astron. Soc.* **412** (2011) 1473–1507, [1006.4613].

[339] M. T. Botticella, S. J. Smartt, R. C. Kennicutt, Jr., E. Cappellaro, M. Sereno and J. C. Lee, *A comparison between star formation rate diagnostics and rate of core collapse supernovae within 11 Mpc*, *Astron. Astrophys.* **537** (2012) A132, [1111.1692].

[340] S. M. Adams, C. S. Kochanek, J. F. Beacom, M. R. Vagins and K. Z. Stanek, *Observing the Next Galactic Supernova*, *Astrophys. J.* **778** (2013) 164, [1306.0559].

[341] A. K. Drukier, S. Baum, K. Freese, M. Górski and P. Stengel, *Paleo-detectors: Searching for Dark Matter with Ancient Minerals*, *Phys. Rev.* **D99** (2019) 043014, [1811.06844].

[342] E. Goto, *On the Observation of Magnetic Monopoles*, *J. Phys. Soc. Japan* **13** (1958) 1413–1418.

[343] E. Goto, H. H. Kolm and K. W. Ford, *Search for Ferromagnetically Trapped Magnetic Monopoles of Cosmic-Ray Origin*, *Phys. Rev.* **132** (1963) 387.

[344] R. L. Fleischer, I. S. Jacobs, W. M. Schwarz, P. B. Price and H. G. Goodell, *Search for multiply charged dirac magnetic poles*, *Phys. Rev.* **177** (1969) 2029–2035.

[345] R. L. Fleischer, H. R. Hart, I. S. Jacobs, P. B. Price, W. M. Schwarz and F. Aumento, *Search for magnetic monopoles in deep ocean deposits*, *Phys. Rev.* **184** (1969) 1393–1397.

[346] R. L. Fleischer, P. B. Price and R. T. Woods, *Search for tracks of massive, multiply charged magnetic poles*, *Phys. Rev.* **184** (1969) 1398–1401.

[347] L. W. Alvarez, P. H. Eberhard, R. R. Ross and R. D. Watt, *Search for Magnetic Monopoles in the Lunar Sample*, *Science* **167** (1970) 701–703.

[348] H. H. Kolm, F. Villa and A. Odian, *SEARCH FOR MAGNETIC MONOPOLES*, *Phys. Rev.* **D4** (1971) 1285.

[349] P. H. Eberhard, R. R. Ross, L. W. Alvarez and R. D. Watt, *SEARCH FOR MAGNETIC MONOPOLES IN LUNAR MATERIAL*, *Phys. Rev.* **D4** (1971) 3260.

[350] R. R. Ross, P. H. Eberhard, L. W. Alvarez and R. D. Watt, *SEARCH FOR MAGNETIC MONOPOLES IN LUNAR MATERIAL USING AN ELECTROMAGNETIC DETECTOR*, *Phys. Rev.* **D8** (1973) 698.

[351] P. B. Price, S.-l. Guo, S. P. Ahlen and R. L. Fleischer, *Search for GUT Magnetic Monopoles at a Flux Level Below the Parker Limit*, *Phys. Rev. Lett.* **52** (1984) 1265.

[352] J. M. Kovalik and J. L. Kirschvink, *New Superconducting Quantum Interface Device Based Constraints on the Abundance of Magnetic Monopoles Trapped in Matter: An Investigation of Deeply Buried Rocks*, *Phys. Rev.* **A33** (1986) 1183–1187.

[353] D. Ghosh and S. Chatterjea, *Supermassive magnetic monopoles flux from the oldest mica samples*, *Europhys. Lett.* **12** (1990) 25–28.

[354] H. Jeon and M. J. Longo, *Search for magnetic monopoles trapped in matter*, *Phys. Rev. Lett.* **75** (1995) 1443–1446, [hep-ex/9508003].

[355] J. I. Collar and K. Zioutas, *Limits on exotic heavily ionizing particles from the geological abundance of fullerenes*, *Phys. Rev. Lett.* **83** (1999) 3097, [astro-ph/9902310].

[356] J. Engel, M. T. Ressell, I. S. Towner and W. E. Ormand, *Response of mica to weakly interacting massive particles*, *Phys. Rev.* **C52** (1995) 2216–2221, [hep-ph/9504322].

[357] R. Hirschi, G. Meynet and A. Maeder, *Stellar evolution with rotation 12: Pre-supernova models*, *Astron. Astrophys.* **425** (2004) 649–670, [astro-ph/0406552].

[358] S. J. Smartt, *Progenitors of core-collapse supernovae*, *Ann. Rev. Astron. Astrophys.* **47** (2009) 63–106, [0908.0700].

[359] P. Madau, M. Della Valle and N. Panagia, *On the evolution of the cosmic supernova rates*, *Mon. Not. Roy. Astron. Soc.* **297** (1998) 17, [astro-ph/9803284].

[360] T. Dahlen et al., *High redshift supernova rates*, *Astrophys. J.* **613** (2004) 189–199, [astro-ph/0406547].

[361] M. Cignoni, S. Degl'Innocenti, P. G. P. Moroni and S. N. Shore, *Recovering the star formation rate in the solar neighborhood*, *Astron. Astrophys.* **459** (2006) 783, [astro-ph/0608654].

[362] M. A. Czekaj, A. C. Robin, F. Figueras, X. Luri and M. Haywood, *The BesanÃ§on Galaxy model renewed*, *Astron. Astrophys.* **564** (2014) A102, [1402.3257].

[363] O. Snaith, M. Haywood, P. Di Matteo, M. D. Lehnert, F. Combes, D. Katz et al., *Reconstructing the star formation history of the Milky Way disc(s) from chemical abundances*, *Astron.Astrophys* **578** (Jun, 2015) A87, [1410.3829].

[364] M. Haywood, M. D. Lehnert, P. Di Matteo, O. Snaith, M. Schultheis, D. Katz et al., *When the Milky Way turned off the lights: APOGEE provides evidence of star formation quenching in our Galaxy*, *Astron.Astrophys* **589** (May, 2016) A66, [1601.03042].

[365] R. Mor, A. C. Robin, F. Figueras, S. Roca-Fabregas and X. Luri, *Gaia DR2 reveals a star formation burst in the disc 2-3 Gyr ago*, *Astron.Astrophys* **624** (2019) L1, [1901.07564].

[366] J. R. Ellis and D. N. Schramm, *Could a nearby supernova explosion have caused a mass extinction?*, *arXiv: hep-ph/9303206s* (1993) , [hep-ph/9303206].

[367] J. R. Ellis, B. D. Fields and D. N. Schramm, *Geological isotope anomalies as signatures of nearby supernovae*, *Astrophys. J.* **470** (1996) 1227–1236, [astro-ph/9605128].

[368] B. D. Fields and J. R. Ellis, *On deep ocean Fe-60 as a fossil of a near earth supernova*, *New Astron.* **4** (1999) 419–430, [astro-ph/9811457].

[369] H. Svensmark, *Evidence of nearby supernovae affecting life on Earth*, *Mon. Not. Roy. Astron. Soc.* **423** (2012) 1234–1253, [1210.2963].

[370] B. C. Thomas, E. E. Engler, M. Kachelrieß, A. L. Melott, A. C. Overholt and D. V. Semikoz, *Terrestrial Effects Of Nearby Supernovae In The Early Pleistocene*, *Astrophys. J.* **826** (2016) L3, [1605.04926].

[371] A. L. Melott, B. C. Thomas, M. Kachelriess, D. V. Semikoz and A. C. Overholt, *A supernova at 50 pc: Effects on the Earth's atmosphere and biota*, *Astrophys. J.* **840** (2017) 105, [1702.04365].

[372] A. L. Melott, F. Marinho and L. Paulucci, *Muon Radiation Dose and Marine Megafaunal Extinction at the end-Pliocene Supernova*, *Astrobiology* **19** (2019) xxx, [1712.09367].

[373] A. L. Melott and B. C. Thomas, *From cosmic explosions to terrestrial fires?*, *arXiv: 1903.01501* (2019) , [1903.01501].

[374] B. D. Fields et al., *Near-Earth Supernova Explosions: Evidence, Implications, and Opportunities*, *arXiv: 1903.04589* (2019) , [1903.04589].

[375] IceCube, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift NuSTAR, VERITAS, VLA/17B-403 collaboration, M. G. Aartsen et al., *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, *Science* **361** (2018) eaat1378, [1807.08816].

[376] N. Langer, *Pre-Supernova Evolution of Massive Single and Binary Stars*, *Ann. Rev. Astron. Astrophys.* **50** (2012) 107–164, [1206.5443].

[377] T. Foglizzo et al., *The explosion mechanism of core-collapse supernovae: progress in supernova theory and experiments*, *Publ. Astron. Soc. Austral.* **32** (2015) e009, [1501.01334].

[378] B. Müller, *The Core-Collapse Supernova Explosion Mechanism*, *IAU Symp.* **329** (2016) 17–24, [1702.06940].

[379] Kamiokande-II collaboration, K. Hirata et al., *Observation of a Neutrino Burst from the Supernova SN 1987a*, *Phys. Rev. Lett.* **58** (1987) 1490–1493.

[380] R. M. Bionta et al., *Observation of a Neutrino Burst in Coincidence with Supernova SN 1987a in the Large Magellanic Cloud*, *Phys. Rev. Lett.* **58** (1987) 1494.

[381] M. Aglietta et al., *Supernova 1987A in the Large Magellanic Cloud*, *IAU Circular* **No. 4323** (1987) .

[382] E. N. Alekseev, L. N. Alekseeva, I. V. Krivosheina and V. I. Volchenko, *Detection of the Neutrino Signal From SN1987A in the LMC Using the Inr Baksan Underground Scintillation Telescope*, *Phys. Lett.* **B205** (1988) 209–214.

[383] M. T. Keil, G. G. Raffelt and H.-T. Janka, *Monte Carlo study of supernova neutrino spectra formation*, *Astrophys. J.* **590** (2003) 971–991, [astro-ph/0208035].

[384] I. Tamborra, B. Muller, L. Hudepohl, H.-T. Janka and G. Raffelt, *High-resolution supernova neutrino spectra represented by a simple fit*, *Phys. Rev.* **D86** (2012) 125031, [1211.3920].

[385] A. Mirizzi, I. Tamborra, H.-T. Janka, N. Saviano, K. Scholberg, R. Bollig et al., *Supernova Neutrinos: Production, Oscillations and Detection*, *Riv. Nuovo Cim.* **39** (2016) 1–112, [1508.00785].

[386] S. Horiuchi, K. Sumiyoshi, K. Nakamura, T. Fischer, A. Summa, T. Takiwaki et al., *Diffuse Supernova Neutrino Background from extensive core-collapse simulations of* 8-100$M_\odot$ *progenitors*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 1363, [1709.06567].

[387] R. Schönrich, J. Binney and W. Dehnen, *Local kinematics and the local standard of rest*, *Mon. Not. Roy. Astron. Soc.* **403** (Apr, 2010) 1829–1833, [0912.3693].

[388] D. Kawata, J. Bovy, N. Matsunaga and J. Baba, *Galactic rotation from Cepheids with Gaia DR2 and effects of non-axisymmetry*, *Mon. Not. Roy. Astron. Soc.* **482** (Jan, 2019) 40–51, [1803.05927].

[389] L. Vietze, P. Klos, J. Menéndez, W. C. Haxton and A. Schwenk, *Nuclear structure aspects of spin-independent WIMP scattering off xenon*, *Phys. Rev.* **D91** (2015) 043520, [1412.6091].

[390] D. Gazda, R. Catena and C. Forssén, *Ab initio nuclear response functions for dark matter searches*, *Phys. Rev.* **D95** (2017) 103011, [1612.09165].

[391] C. Körber, A. Nogga and J. de Vries, *First-principle calculations of Dark Matter scattering off light nuclei*, *Phys. Rev.* **C96** (2017) 035805, [1704.01150].

[392] M. Hoferichter, P. Klos, J. Menéndez and A. Schwenk, *Nuclear structure factors for general spin-independent WIMP-nucleus scattering*, *Phys. Rev.* **D99** (2019) 055031, [1812.05617].

[393] J. F. Ziegler, J. P. Biersack and U. Littmark, *The Stopping and Range of Ions in Solids*. Pergamon Press, New York, 01, 1985.

[394] J. F. Ziegler, M. D. Ziegler and J. P. Biersack, *SRIM â The stopping and range of ions in matter (2010)*, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **268** (2010) 1818 – 1823.

[395] D. Mei and A. Hime, *Muon-induced background study for underground laboratories*, *Phys. Rev.* **D73** (2006) 053004, [astro-ph/0512125].

[396] SNO collaboration, B. Aharmim et al., *Measurement of the Cosmic Ray and Neutrino-Induced Muon Flux at the Sudbury Neutrino Observatory*, *Phys. Rev.* **D80** (2009) 012001, [0902.2776].

[397] D. G. Madland, E. D. Arthur, G. P. Estes, J. E. Stewart, M. Bozoian, R. T. Perry et al., *SOURCES 4A: A Code for Calculating (alpha,n), Spontaneous*

Fission, and Delayed Neutron Sources and Spectra, Tech. Rep. LA-13639-MS, Los Alamos National Lab, 1999. 10.2172/15215.

[398] N. Soppera, M. Bossant and E. Dupont, *JANIS 4: An Improved Version of the NEA Java-based Nuclear Data Information System*, *Nucl. Data Sheets* **120** (2014) 294–296.

[399] A. J. Koning and D. Rochman, *Modern Nuclear Data Evaluation with the TALYS Code System*, *Nucl. Data Sheets* **113** (2012) 2841–2934.

[400] D. Rochman, A. J. Koning, J. C. Sublet, M. Fleming et al., *The TENDL library: hope, reality and future*, in *Proceedings of the International Conference on Nuclear Data for Science and Technology, September 11-16, 2016, Bruges, Belgium*, 2016.

[401] J. C. Sublet, A. J. Koning, D. Rochman, M. Fleming and M. Gilbert, *TENDL-2015: Delivering Both Completeness and Robustness*, in *Advances in Nuclear Nonproliferation Technology and Policy Conference, Sept. 25-30, Santa Fe, NM, USA*, 2016.

[402] M. Fleming, J. C. Sublet, J. Kopecky, D. Rochman and A. J. Koning, *Probing experimental and systematic trends of the neutron-induced TENDL-2014 nuclear data library*, in *CCFE report UKAEA-R(15)29, October 2015*, 2015.

[403] W. Li, R. Chornock, J. Leaman, A. V. Filippenko, D. Poznanski, X. Wang et al., *Nearby supernova rates from the Lick Observatory Supernova Search - III. The rate-size relation, and the rates as a function of galaxy Hubble type and colour*, *Mon. Not. Roy. Astron. Soc.* **412** (Apr, 2011) 1473–1507, [1006.4613].

[404] J. Billard, F. Mayet and D. Santos, *Assessing the discovery potential of directional detection of dark matter*, *Phys. Rev. D* **85** (Feb., 2012) 035006, [1110.6079].

[405] R. C. Kennicutt, Jr., *Star formation in galaxies along the Hubble sequence*, *Ann. Rev. Astron. Astrophys.* **36** (1998) 189–231, [astro-ph/9807187].

[406] R. Genzel et al., *A Study of the Gas-Star Formation Relation over Cosmic Time*, *Mon. Not. Roy. Astron. Soc.* **407** (2010) 2091, [1003.5180].

[407] I. Negueruela, *The Starbursts in the Milky Way*, in *Massive Young Star Clusters Near and Far: From the Milky Way to Reionization*, pp. 9–15, 2014. 1402.3328.

[408] S. Ohm, *Starburst galaxies as seen by gamma-ray telescopes*, *Comptes Rendus Physique* **17** (2016) 585–593, [1601.06386].

[409] N. Benitez, J. Maiz-Apellaniz and M. Canelles, *Evidence for nearby supernova explosions*, *Phys. Rev. Lett.* **88** (2002) 081101, [astro-ph/0201018].

[410] P. Ludwig, S. Bishop, R. Egli, V. Chernenko, B. Deneva, T. Faestermann et al., *Time-resolved 2-million-year-old supernova activity discovered in Earth's microfossil record*, *Proceedings of the National Academy of Science* **113** (Aug, 2016) 9232–9237, [1710.09573].

[411] L. Fimiani, D. L. Cook, T. Faestermann, J. M. Gómez-Guzmán, K. Hain, G. Herzog et al., *Interstellar ^{60}Fe on the Surface of the Moon*, *Phys. Rev. Lett.* **116** (Apr, 2016) 151104.

[412] A. Wallner, J. Feige, N. Kinoshita, M. Paul, L. K. Fifield, R. Golser et al., *Recent near-Earth supernovae probed by global deposition of interstellar radioactive ^{60}Fe*, *Nature* **532** (Apr, 2016) 69–72.

[413] M. M. Schulreich, D. Breitschwerdt, J. Feige and C. Dettbarn, *A Way Out of the Bubble Trouble?âUpon Reconstructing the Origin of the Local Bubble and Loop I via Radioisotopic Signatures on Earth*, *Galaxies* **6** (2018) 26, [1802.09275].

[414] M. Kachelrieß, A. Neronov and D. V. Semikoz, *Signatures of a two million year old supernova in the spectra of cosmic ray protons, antiprotons and positrons*, *Phys. Rev. Lett.* **115** (2015) 181103, [1504.06472].

[415] W. R. Binns et al., *Observation of the 60Fe nucleosynthesis-clock isotope in galactic cosmic rays*, *Science* **352** (2016) 677–680.

[416] D. Breitschwerdt, J. Feige, M. M. Schulreich, M. A. D. Avillez, C. Dettbarn and B. Fuchs, *The locations of recent supernovae near the Sun from modelling ^{60}Fe transport*, *Nature* **532** (Apr, 2016) 73–76.

[417] A. D. Erlykin, S. K. Machavariani and A. W. Wolfendale, *The Local Bubble in the interstellar medium and the origin of the low energy cosmic rays*, *Adv. Space Res.* **59** (2017) 748–750, [1610.07014].

[418] J. A. S. Adams, J. K. Osmond and J. J. W. Rogers, *The geochemistry of thorium and uranium*, *Phys. Chem. Earth* **3** (1959) 298–348.

[419] P. P. Povinec, *New ultra-sensitive radioanalytical technologies for new science*, *J Radioanal Nucl Chem* **316** (2018) 893–931.

[420] P. P. Povinec et al., *Ultra-sensitive radioanalytical technologies for underground physics experiments*, *J. Radioanal. Nucl. Chem.* **318** (2018) 677–684.

[421] K. C. Condie, C. S. Kuo, R. M. Walker and V. R. Murthy, *Uranium Distribution in Separated Clinopyroxenes from Four Eclogites*, *Science* **165** (1969) 57–59.

[422] M. Seitz and S. Hart, *Uranium and boron distributions in some oceanic ultramafic rocks*, *Earth and Planetary Science Letters* **21** (1973) 97 – 107.

[423] S. Thomson and G. Wardle, *Coloured natural rocksalts: a study of their helium contents, colours and impurities*, *Geochimica et Cosmochimica Acta* **5** (1954) 169 – 184.

[424] W. E. Dean, *Section 5 trace and minor elements in evaporites*, in *Marine Evaporites*. SEPM Society for Sedimentary Geology, 1987. DOI.

[425] W. Sanford, M. Doughten, T. Coplen, A. Hunt and T. Bullen, *Evidence for high salinity of Early Cretaceous sea water from the Chesapeake Bay crater*, *Nature* **503** (2013) 252–256.

[426] M. Yui, Y. Kikawada, T. Oi, T. Honda and T. Nozaki, *Abundance of Uranium and Thorium in Rock Salts*, *Radioisotopes* **47** (1998) 488–492.

[427] K. Eda, Y. Itoh, S. Kuroyanagi and J. Silk, *Gravitational waves as a probe of dark matter minispikes*, *Phys. Rev.* **D91** (2015) 044045, [1408.3534].

[428] K. Eda, Y. Itoh, S. Kuroyanagi and J. Silk, *New Probe of Dark-Matter Properties: Gravitational Waves from an Intermediate-Mass Black Hole Embedded in a Dark-Matter Minispike*, *Phys. Rev. Lett.* **110** (2013) 221101, [1301.5971].

[429] E. Barausse, V. Cardoso and P. Pani, *Environmental Effects for Gravitational-wave Astrophysics*, *J. Phys. Conf. Ser.* **610** (2015) 012044, [1404.7140].

[430] E. Barausse, V. Cardoso and P. Pani, *Can environmental effects spoil precision gravitational-wave astrophysics?*, *Phys. Rev.* **D89** (2014) 104059, [1404.7149].

Samenvatting

Het werk van een natuurkundige bestaat in het algemeen uit twee aspecten: een poging wagen om een signaal van een nieuw fenomeen te vinden, of het proberen de lange lijst van onbeantwoorde fenomenen uit te leggen. Dit proefschrift zit op de grens van de twee en neemt ideeën van de ene kant en past ze toe op de andere kant, hopend op succes om iets nieuws te begrijpen, of het vinden van oplossingen op lang bestaande mysteries. Gedurende ons dagelijks leven hebben we interactie met *dingen* van waaruit onze wereld uit bestaat. Deze dingen, de atomen die onze lichamen en alles uit het universum samenstellen, is *materie*. Vreemd genoeg, hebben de observaties van hoe het universum beweegt en ontwikkelt, geleid tot het geloven dat er veel meer dingen zijn die we niet kunnen zien, *donkere materie*. Er is gewoon niet genoeg zichtbare massa om de graviteit uit te leggen die we zien werken op sterren, sterrenstelsels, and nog verder. Deze donkere materie is sinds decenia een mysterie voor natuurkundige. Velen denken dat het een nieuw type deeltje is, met normale materie als bouwstenen, hoewel dit nooit bewezen is. Helaas is het zoeken naar donkere materie extreem moeilijk vanwege zijn aard. Het feit dat we het nog niet hebben gezien betekent dat het zo zelden met normale materie interactie ondergaat, dat het geen effect heeft op aarde. In feite, door het volgen van de bewegingen van sterren die dichtbij aarde zijn, kunnen we de hoeveelheid donkere materie dichtbij ons voorspellen. Het blijkt dat er zo veel donkere materie is dat er bijna 10,000 deeltjes elke seconde door je lichaam gaan! Gelukkig kunnen we deze overvloed gebruiken om donkere materie hier op aarde te studeren. Door detectors diep onder de grond te bouwen, is het mogelijk om kleine energie sporen te traceren die donkere materie zouden kunnen hebben vrijgelaten, gedurende een van deze zeldzame gebeurtenis. Het interpreteren van data afkomstig van een van deze experimenten is een moeilijke taak. Aangezien het aantal keren dat er interactie zal plaats vinden zo laag is, zouden we bij geluk enkel een of meerdere interacties zien. Als we een ontdekking willen claimen en kunnen zeggen dat we werkelijk donkere materie hebben gezien, is het noodzakelijk dat we absoluut zeker zijn dat deze *gebeurtenissen* zijn veroorzaakt door donkere materie. Hoofdstuk 4 gaat precies over dit onderwerp. Hoe kunnen we zeker zijn dat een detector gevoelig

genoeg is om het bewijs te leveren dat we nodig hebben? We ontwikkelden de nodige technieken om deze vragen op een efficiënte manier te beantwoorden. In hoofdstuk 5 zijn we verder gegaan dan dat, we berekenden niet alleen of toekomstige experimenten donkere materie kan *detecteren* maar ook of ze gevoelig genoeg zijn om te leren over zijn aard, en dus om zijn geheimen te onthullen. Helaas zullen we met toekomstige experimentele opstellingen waarschijnlijk niets leren over alles wat we over donkere materie kunnen weten en zouden daarom meer toekomstige experimentele testen nodig zijn. We zijn vervolgens in hoofdstuk 6 naar een nieuwe manier gekeerd om deze energiestorting te detecteren; het gebruiken van oude stenen onder the oppervlakte van de aarde. Door het zoeken naar afwijkingen in de steen structuur, kunnen we zoeken naar historische bewijzen van donkere materie die de aarde passeren. We toonden aan dat deze *paleo-detectors* de meest gevoelige donkere materie detectors tot op heden zijn, en ze liggen enkel onder onze voeten.

Tenslotte zijn we in Hoofdstuk 7 omgedraaid naar een andere set van mysterieuze deeltjes, bekend als *neutrinos*. Deze fundamentele deeltjes worden nog niet goed begrepen maar ze spelen een grote rol in veel astrofysische omgevingen. De explosies die aan het einde van een sterren leven gebeuren, genaamd supernovae, zouden er compleet anders uitzien als er geen neutrinos waren. Net als donkere materie, zijn ze overal om ons heen, en bewegen ze schadeloos door aarde. We lieten zien dat paleo-detectoren neutrinos kan detecteren afkomstig van deze supernovae explosies in ons eigen sterren-stelsel. Als een supernova dichtbij genoeg van de aarde is afgegaan, zou de straling afkomstig van de explosie onze atmosfeer ernstig kunnen beschadigen, en potentieel massa-extinctie kunnen veroorzaken. We lieten zien dat paleo-detectoren bewijs kan verzamelen voor een van deze gebeurtenissen, dat het bewijs kan verstrekken over hoe het leven ontwikkelde op aarde.

Vertaald door Ariane Dekker.

Summary

A physicists job generally takes two forms: attempting to find signs of a new phenomena, or trying to explain the long list of existing unexplained phenomena. This thesis sits at the boundary, taking ideas from one side and applying them to the other, hoping to succeed in understanding something new, or finding solutions to long standing mysteries.

Throughout our daily lives we interact with the *stuff* that makes up the world around us. This stuff, the atoms that make up our bodies and everything throughout the Universe, is called *matter*. Strangely, measurements of how the Universe moves and evolves have led physicists to believe that there is a lot of extra stuff that we cannot see, *dark matter*. There just isn't enough visible matter to explain amount of gravity we see acting on stars, galaxies, and beyond.

This dark matter has been a mystery to physicists for decades. Many think its a new type of particle, the building blocks normal matter, although this has never been proven. Unfortunately, searching for dark matter is extremely difficult by its very nature. The fact that we haven't seen it already means that it must interact with normal matter so rarely that it has no effect on the Earth. In fact, by tracking the motions of stars close to Earth we can predict the amount of dark matter close by. It turns out that there is so much dark matter that almost 10,000 particles pass through your body every second! Luckily we can use this abundance to search for dark matter here on Earth. By building detectors deep underground its possible to look for the tiny traces of energy that dark matter might release when one of these rare events occurs.

To interpret the data from one of these experiments is a arduous task. Since the number of times these interactions will happen is so low, just by chance we might see only one but we could also see more. If we want to make a discovery and claim we truly have seen dark matter, its imperative that we are absolutely certain that these *events* are caused by dark matter. Chapter 4 is about exactly this topic. How can we be sure that a detector will be sensitive enough to find the evidence we need? We

developed the necessary techniques to answer these questions in an efficient way. In Chapter 5 we went further, calculating whether future experiments will be able to not only *detect* the dark matter but learn about its nature, thus revealing its secrets. Unfortunately, future experimental setups will be unlikely to learn all there is to know about dark matter therefore further experimental tests may be required.

We then turned, in Chapter 6, to a new way to detect these minute energy deposits; using ancient rocks from below the surface of the Earth. By searching for abnormalities in the rocks structure, we can search for the historic evidence of dark matter passing through the Earth. We showed that these *paleo-detectors* could be the most sensitive dark matter detectors to date, and they lie just under our feet.

Finally in Chapter 7 we turned to another set of mysterious particles, known as *neutrinos*. These fundamental particles are not well understood but known to play huge roles in many astrophysical settings. The explosions that happen at the end of a stars life, called supernovae, would look completely different if it were not for neutrinos. Like dark matter, they are all around us, harmlessly zooming through the Earth. We showed that paleo-detectors could find the neutrinos from these supernovae explosions in our own galaxy. If a supernova went off close enough to Earth in its history, the radiation from the explosion could seriously damage our atmosphere, potentially causing mass extinction events. We showed that paleo-detectors would record evidence for one of these events within them, providing evidence for how life evolved on Earth.

Acknowledgements

First and foremost I'd like to thank my supervisor Christoph Weniger. You've pushed and guided me when I was unsure of how to progress and given me space to explore when I needed. You've taught me how to do research and how to navigate the hectic world of an being academic. Thank you for your time, patience, fountain of knowledge, and complete support.

I am also hugely grateful to Bradley Kavanagh who has been a constant source of guidance for me. From the moment we began working together his simple approach to physics and insightful ability to communicate has been an inspiration. I could not ask for a better collaborator.

To my promotor Patrick Decowski, thank you for making the time to talk to me. Although my PhD has gone relatively smoothly, his support has given me the confidence that I would not be left alone if there was an unexpected set back. For this I'm extremely grateful. In addition, he has provided a direct link to NIKHEF and given me a wonderful insight into the more experimental side of physics, something I would not have without him. Jan Pieter van der Schaar as my independent supervisor has also been an excellent, ensuring that I always feel comfortable and welcome in the Strings/Cosmology group and free to talk should anything go wrong.

In the final year of my PhD Samaya Nissanke became the newest faculty member of GRAPPA. Her openness to new collaborations and optimism about scientific research has been exceptional to experience. She allowed me to collaborate with her group, despite my lack of any of the prequisite scientific knowledge. I'm extremely grateful for her scientific guidance and fundamental insight into the world of multi-messenger physics. I look forward to working closely together in the future. To her group as well, Tanja Hinderer, David Nichols, and Andrew Williamson, thank you for putting up with my endless questioning and for providing endless answers.

Niki Klop, Richard Bartels, Sebastian Liem, Hamish Silverwood, Suraj Krishnamurthy, and Michael Feyereisen all deserve a special thank you for being wonderful office mates and friends.

Numerous other people along the way have made my PhD a pleasure: Emma Storm, Francesca Calore, Horng Sheng Chia, Fabio Zandanel, and Daniele Gaggero to name just a few. Special thanks also goes to my fellow PhD student Ariane Dekker for translating my English Summary to the Dutch Samenvatting.

Thank you to the faculty for making GRAPPA an engaging, fun, and scientifically outstanding place to work; Gianfranco Bertone, Shin'ichiro Ando, Ben Freivogel, Jacco Vink, and Daniel Baumann.

No department can function or succeed without the tireless effort of the administration staff; Natalie Wells, Astrid Harryson, Jiřina Šálková, Anne-Marieke Crommentuijn, Klaartje Wartenbergh, and Joost van Mameren. The Institute of Physics is blessed with some of the best and I'm hugely grateful for their patience and knowledge as I've tried to navigate the difficult path of completing a PhD.

Thank you to my girlfriend, Helen, for her constant support and understanding in times of stress. On top of being a wonderful partner, she has been a continued source of inspiration to improve my communication skills both academically and to the public. I am forever grateful for everything you have given me.

Finally, I thank my family who have constantly supported me throughout this endeavour, despite finding my choice of profession strange and very disconnected from reality. Particularly to my parents who have supported me unconditionally throughout my life. I would have not completed this without them.