



VRIJE
UNIVERSITEIT
BRUSSEL

Vrije Universiteit Brussel
Faculteit Wetenschappen en
Bio-ingenieurswetenschappen
Departement Natuurkunde

Search for Dark Matter in the Monojet
and Trackless Jets Final States with
the CMS Detector at the LHC

Isabelle De Bruyn

Promotor
Prof. Dr. Steven Lowette

Proefschrift ingediend met het oog op het behalen van de
academische graad van Doctor in de Wetenschappen

December 2017

Doctoral examination commission:

Prof. Dr. Alexander Sevrin (VUB, *chair*)
Prof. Dr. Jorgen D'Hondt (VUB, *secretary*)
Prof. Dr. Steven Lowette (VUB, *promotor*)
Prof. Dr. Dominique Maes (VUB)
Dr. Krijn De Vries (VUB)
Prof. Dr. Albert De Roeck (CERN, UA)
Prof. Dr. Fabio Maltoni (UCL)

© 2017 Isabelle De Bruyn

All rights reserved. No parts of this book may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the author.

Samenvatting

Alle materie die we rondom ons zien, is opgebouwd uit atomen, bestaande uit negatief geladen elektronen die rond een positief geladen kern cirkelen. Hoewel de elektronen – voor zover onze huidige kennis strekt – elementaire deeltjes zijn, bevat de kern van een atoom protonen en neutronen, die op hun beurt bestaan uit up en down quarks. Al deze deeltjes en hun onderlinge wisselwerking worden beschreven in een theorie die gekend staat als het Standaard Model van de Deeltjesfysica. Desondanks kan deze uiterst succesvolle theorie niet alle waargenomen fenomenen verklaren. Verschillende kosmologische waarnemingen tonen namelijk aan dat de gekende materie die door het Standaard Model beschreven wordt maar 15% van de totale materie in het universum beslaat. De overige materie kan waargenomen worden door middel van zwaartekrachteffecten, maar is niet zichtbaar via waarnemingen gebaseerd op licht van eender welke golflengte. Dit wijst erop dat deze zogenaamde donkere materie ongeladen is. Verder is er over deze materie bitter weinig geweten, maar bestaan er talloze theoretische modellen die de oorsprong ervan proberen te verklaren.

Afhankelijk van de aard van de donkere materie deeltjes, kunnen deze geproduceerd worden in hoogenergetische botsingen die plaatsvinden in deeltjesversnellers. Vele modellen nemen aan dat de donkere materie deeltjes zwak met de gewone materie interageren, via een nieuwe, onbekende kracht. Dit maakt het mogelijk om deze deeltjes te produceren door deeltjes van de gekende materie aan een hoge snelheid tegen elkaar te laten botsen.

Deze thesis behandelt twee zoektochten naar donkere materie, die uitgevoerd werden aan het Compact Muon Solenoid (CMS) experiment dat zich aan de Large Hadron Collider (LHC) in het CERN bevindt. De LHC is momenteel de werelds grootste en krachtigste deeltjesversneller en laat protonen aan een hoge frequentie tegen elkaar botsen, met een massamiddelpuntsenergie van 13 TeV. De CMS detector is veelzijdig en wordt zowel voor precieze testen van het Standaard Model als voor onderzoek naar nieuwe fysica gebruikt.

In de eerste analyse worden de donkere materie deeltjes verondersteld de CMS detector ongezien te verlaten, aangezien ze neutraal zijn en zwak interageren. Wanneer ze samen met andere deeltjes geproduceerd worden, kunnen deze evenwel waargenomen worden door een onevenwicht in de gemeten energien, wat leidt tot ontbrekende energie in de detector. Deze methode wordt in de zogenaamde monojet analyse gebruikt door te zoeken naar een combinatie van ontbrekende energie en gecollimeerde bundels van deeltjes, zogenaamde jets. Het werk in deze thesis heeft tot een nauwkeurigere voorspelling van de achtergrond geleid en heeft het bekomen resultaat aanzienlijk verbeterd. Er werd geen nieuwe fysica waargenomen boven op de voorspelde achtergrond, en nieuwe, strengere beperkingen werden op deze manier aan de beschouwde modellen opgelegd.

In het tweede deel van deze thesis wordt een enigszins ongewoon model bestudeerd. In dit geval wordt er gezocht naar sterk interagerende donkere materie deeltjes, in tegenstelling tot zwak interagerende deeltjes. Deze donkere materie deeltjes zouden in paren geproduceerd worden via een nieuw krachtdragend deeltje, dat in gelijke mate met de gekende materie interageert als protonen en neutronen. Het signaal dat deze deeltjes in de detector achterlaten lijkt bijgevolg sterk op dat van neutronen, aangezien deze ook ongeladen zijn. Er wordt dus gezocht naar een paar neutrale jets, zogenaamde *trackless jets*, die gemakkelijk onderscheiden kunnen worden van de achtergrond bestaande uit geladen jets. Het resultaat van dit onderzoek is volledig

compatibel met de voorspelde achtergrond, waardoor het beschouwde model uitgesloten wordt.

De twee onderzochte scenario's vullen elkaar aan, daar de ontbrekende energie die in de monojet analyse gebruikt wordt in neutrale jets omgevormd kan worden wanneer de interactiewaarschijnlijkheid van de deeltjes groot genoeg wordt. Hoewel er geen teken van nieuwe fysica waargenomen werd, leiden deze zoektochten tot het uitsluiten van een aantal donkere materie modellen.

Summary

All matter we know and see around us is made up of atoms, which consist of negatively charged electrons revolving around a positively charged nucleus. While the electrons are – as far as we currently know – fundamental particles, the nucleus contains protons and neutrons, which are in turn composed of up and down quarks. The theoretical framework that describes all these fundamental particles and their interactions is called the Standard Model of Particle Physics. While it is an extremely successful theory, multiple unresolved questions and observations cannot be explained by the Standard Model. Cosmological observations, for example, indicate that the known matter described by the Standard Model only contributes 15% of all the matter in the universe. The remaining matter is observed through gravitational interactions, but is not visible in observations of light at any wavelength, implying it is electrically neutral. Only very little is known about this so-called dark matter, and many theoretical models exist to explain its origin.

Depending on their exact nature, dark matter particles might be produced in high-energy collisions at particle colliders. Many models assume that the dark matter particles interact weakly with ordinary matter, through a new force, making it possible to produce them in the collision of two Standard Model particles.

This thesis covers two searches for dark matter performed at the Compact Muon Solenoid (CMS) experiment at the CERN Large Hadron Collider. This particle accelerator is currently the largest in the world, and provides proton-proton collisions with a record centre-of-mass energy of 13 TeV at a high collision rate. The CMS detector is a multi-purpose particle detector, used for various precision measurements of the Standard Model and many searches for new physics.

In the first analysis, the dark matter particles are expected to leave the CMS detector undetected as they are neutral and weakly interacting. When they are produced in association with other particles, they can however be observed due to an imbalance of energies measured in the detector, called missing energy. This technique is used in the first dark matter search described in this thesis, called the monojet analysis, where the missing energy is balanced with one or more collimated sprays of particles emerging from the collision, so-called jets. The work in this thesis refined the background prediction and thus increased the sensitivity of the search. No significant excess above the predicted background was observed, setting new, stronger limits on several dark matter models, and excluding a larger part of the available parameter space.

As no observation was made in this first analysis, a more unusual model is studied as well. Instead of looking for weakly interacting massive particles, strongly interacting candidates were considered. These dark matter candidates would be produced in pairs through a new mediating particle, which has a probability to interact with matter that is similar to protons or neutrons. As a result, these particles will leave a signal in the detector that is similar to neutrons, which are electrically neutral as well. The investigated signature is therefore a pair of neutral or so-called trackless jets, which can efficiently be differentiated from the background consisting of charged jets. The result of this search is compatible with the predicted background, and again a part of parameter space was excluded.

To conclude, the two searches covered in this thesis are very complementary, as the missing transverse energy signature used in the monojet search can transform into a trackless jets signature when the interaction probability becomes large enough. Although no sign of new physics was observed, these searches have led to the exclusion of more dark matter scenarios.

Table of Contents

1	Introduction	1
2	Dark Matter Scenarios Beyond the Standard Model	5
2.1	The Standard Model of Particle Physics	5
2.1.1	Elementary particles and their interactions	5
2.1.2	The theoretical framework of the Standard Model	6
2.1.3	Unanswered questions of the Standard Model	8
2.2	Dark matter	10
2.2.1	Observational evidence	10
2.2.2	Dark matter models	12
2.2.3	Detection of dark matter	16
2.2.3.1	Direct detection experiments	16
2.2.3.2	Indirect detection experiments	18
2.2.3.3	Collider experiments	19
2.2.4	From EFTs to simplified models	20
2.3	Strongly Interacting Massive Particles	22
2.3.1	The SIMP simplified model	22
2.3.2	Experimental constraints	23
3	The LHC and the CMS Detector	27
3.1	The Large Hadron Collider at CERN	27
3.1.1	The LHC injector chain	27
3.1.2	The Large Hadron Collider	28
3.1.3	The experiments at the LHC	30
3.2	The CMS detector	30
3.2.1	The tracker	31
3.2.1.1	The pixel tracker	31
3.2.1.2	The strip tracker	32
3.2.2	The electromagnetic calorimeter	33
3.2.3	The hadronic calorimeter	35
3.2.4	The muon system	35
3.2.5	Trigger and data acquisition	36
3.2.6	CMS performance in Run 2	38
3.2.6.1	Pre-amplifier saturation in the APV25 chip	38
4	Event Simulation and Reconstruction	43
4.1	Event generation	43
4.1.1	Generation of the monojet signals	46
4.1.2	Generation of the SIMP signal	46
4.2	Detector simulation	47
4.3	Event reconstruction	47
4.3.1	Track and vertex reconstruction	47

4.3.2	Electron and isolated photon reconstruction	49
4.3.3	Electron and photon identification	50
4.3.4	Muon reconstruction	50
4.3.5	Muon identification	51
4.3.6	Particle flow	52
4.3.7	Jet reconstruction	52
4.3.8	Identification of b-jets	54
4.3.9	Reconstruction of tau leptons	54
4.3.10	Missing transverse momentum reconstruction	55
4.4	Simulation and reconstruction of the SIMP signal	55
5	The Monojet Analysis	59
5.1	Physics object reconstruction	59
5.1.1	Jets	59
5.1.2	Missing transverse momentum and hadronic recoil	60
5.1.3	Leptons	60
5.1.4	Photons	60
5.2	Trigger selection	60
5.3	Event selection	61
5.4	Background estimation	62
5.4.1	The Z and W background estimation	62
5.4.2	The QCD background estimation	68
5.4.3	Simulation-based background estimation	68
5.5	Systematic uncertainties	69
5.6	Results	70
5.7	Improvements going from the 2015 to 2016 analysis	74
5.8	Limit setting	78
5.9	Interpretation	78
6	Search for SIMPs using Trackless Jets	81
6.1	Physics object reconstruction	81
6.1.1	Jets	81
6.1.2	Photons	82
6.1.3	Primary vertex	82
6.2	Trigger selection	83
6.3	Event selection	85
6.4	Background estimation	86
6.4.1	Photon + jets	86
6.4.2	QCD multijets	88
6.5	Systematic uncertainties	92
6.6	Results	93
6.7	SIMP model interpretation	95
7	Conclusions & Outlook	97
References		110
List of Acronyms		114

1

Introduction

There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened.

– Douglas Adams, *The Hitchhiker’s Guide to the Galaxy*

Nevertheless, humankind is still trying to understand the most fundamental aspects of our universe, by studying the fundamental particles matter is made of and the interactions between them. The story so far has been summarised in a theory called the Standard Model of Particle Physics. This theory has been extensively tested and has already predicted many experimental observations. However, it cannot explain the full story, and some pieces remain missing. Gravity, for example, is not incorporated in the Standard Model. Similarly, it cannot explain the observed neutrino masses or the matter-antimatter asymmetry.

Another mystery stems from a series of cosmological observations made during the last century. These observations are based on gravitational effects, such as measurements of the rotation curves of galaxies [1] and gravitational lensing [2], and on the analysis of the Cosmic Microwave Background (CMB) [3–5]. The collected evidence shows that there is matter in the universe, which is not visible from measurements at any wavelength of the electromagnetic spectrum. This so-called dark matter was found to constitute about 85% of the matter in the universe, which means that the ordinary matter described by the Standard Model only accounts for 15% of all matter. So far, only very little is known about the dark matter, as it does not interact through any of the forces included in the Standard Model, and has only been observed through gravitational interaction at large scales. Many theoretical models therefore exist, that try to model this unknown type of matter. These theories generally assume that this form of matter is composed of particles, just as the known matter, and describe a new type of interaction through which these dark matter particles interact with the Standard Model particles.

If the dark matter indeed interacts with ordinary matter through a new force, mediated by a new particle, it can be searched for, and many existing theories can be tested. A myriad of experiments are currently looking for dark matter, and can be divided into three categories. Firstly, direct detection experiments take advantage of the dark matter particles that should be present in a halo permeating our galaxy and try to measure the recoil of nuclei generated by dark matter particles passing through the Earth and scattering off the ordinary matter. The detectors used for this type of experiment are mostly located underground and are well shielded from radiation, though a few are airborne or space experiments. Indirect detection experiments on the other hand look for particles or radiation coming from the annihilation of dark matter particles

in dense regions such as the galactic centre. These searches are studying gamma rays, neutrinos, electrons and positrons, or radio emissions. Finally, dark matter particles could potentially also be produced and detected at particle colliders. One of the direct detection experiments observed evidence pointing to the existence dark matter particles, but so far no conclusive observations have been made.

At collider experiments, such as ATLAS and CMS, dark matter candidates are often looked for by focusing on missing energy. Indeed, if the dark dark matter is assumed to interact weakly with the ordinary matter, it will be able to leave the detector unnoticed. However, these so-called weakly interacting massive particles (WIMPs), can be detected when they recoil against another object. Some examples of such collider searches are the monophoton, monojet, monolepton, and mono-Higgs analyses, categorised based on the signature in the detector. Additionally, different signatures are obtained when the dark matter is for example produced in a cascade of decays. Also resonances in e.g. the dijet mass spectrum are looked for, as this could indicate the existence of a new dark matter mediator. In general, more and more analyses are adding dark matter interpretations to their results. This thesis describes two searches for dark matter performed using data from high-energy proton-proton collisions produced at a centre-of-mass energy of 13 TeV and recorded with the CMS detector. The first analysis is the so-called the monojet analysis, which investigates the existence of WIMPs as dark matter candidates. Conversely, the second search looks for dark matter in the form of strongly interacting massive particles (SIMPs).

In Chapter 2, an overview of the Standard Model is given, as well as a short description of a few of its shortcomings. Furthermore, a summary of the existing evidence for dark matter, together with a concise review of popular dark models and a brief description of the operational or developing dark matter experiments are given. As this thesis covers dark matter searches that are performed using the CMS detector, located at one of the collision points of the LHC at CERN, more details concerning this accelerator and particle detector are summarised in Chapter 3. In Chapter 4, the procedure to reconstruct the collisions occurring inside the detector is detailed, as well as the necessary simulations of the predicted signal, which are needed in order to design a search for a particular dark matter candidate and to tune the analysis to the expected signature in the detector. The required techniques for this are described, with more details on the specific simulations needed for the searches covered in this thesis. The two complementary dark matter searches are described in Chapters 5 and 6.

The monojet analysis, covered in Chapter 5, is one of the flagship analyses which were designed to quickly detect potential dark matter candidates, for a broad range of models. I contributed to this analysis by improving the prediction of the main background, coming from the invisible decay of Z bosons into neutrinos, produced in association with one or more jets. The used strategy for the background estimation is detailed in Section 5.4 and the resulting impact on the sensitivity of the analysis is shown in Section 5.7. In the strongly interacting massive particle (SIMP) analysis, described in Chapter 6, the dark matter candidates and the Standard Model particles interact strongly through a new force, carried by a new, light mediator. The signature therefore does not consist of missing energy, but instead trackless jets are created due to the interaction of these neutral SIMPs in the dense material of the calorimeters in the detector. First, a phenomenological study of the dark matter model was performed and published [6], and subsequently the search was carried out using data collected by CMS in 2016. This work, together with my contribution to the monojet analysis are the main topics of my PhD research. The monojet search provides new, stronger limits on WIMP dark matter candidates, while the trackless jets search rules out a new dark matter model which had not been tested at colliders yet.

Besides my work on the dark matter searches described in this thesis, I also contributed in several ways to the data taking itself, more specifically related to the CMS tracker. During the shutdown preceding Run 2, I participated in the preparations to operate the tracker at a colder coolant temperature and to the subsequent re-commissioning. Afterwards, during data taking, I continued to be involved by taking on shifts in the control room and I additionally contributed

to the development and maintenance of the tracker Data Quality Monitoring (DQM) framework, which is used to spot problems in data taking and to certify the recorded data, as mentioned in Section 3.2.6.

2

Dark Matter Scenarios Beyond the Standard Model

In modern particle physics, the fundamental structure of matter at subatomic scales is described by the Standard Model [7, 8], which has already predicted many experimental results and is today considered among the most thoroughly tested scientific theories. The Higgs boson, the last missing piece which was predicted more than 50 years ago, was recently discovered at the Large Hadron Collider (LHC) in 2012 [9, 10], thus completing this elegant theory. Although it has survived many precision tests so far, the Standard Model only describes 5% of the matter and energy in the known universe and is unable to explain many unresolved questions and observations, such as the baryon asymmetry, dark matter and dark energy, the neutrino masses, the incorporation of gravity, and the hierarchy problem.

In this chapter, a brief description of the Standard Model is first given in Section 2.1, including its shortcomings. In Section 2.2, one of the missing pieces in the Standard Model, dark matter, is discussed. The observational evidence for dark matter, as well as possible models and detection mechanisms are detailed. Finally, in Section 2.3 one of the dark matter models considered in this thesis is described in detail.

2.1 The Standard Model of Particle Physics

The Standard Model of elementary particle physics has been developed during the second half of the 20th century, and reached its current formulation in the 1970's with the combination of the electromagnetic and weak interactions into the electroweak interaction [11–13], incorporating the Brout-Englert-Higgs (BEH) mechanism [14–16], and the addition of asymptotic freedom [17, 18] into the theory of the strong interaction. It is a quantum field theory which describes the fundamental particles and their interactions, incorporating three of the four fundamental forces. While this consistent framework describes the electromagnetic force and the weak and strong nuclear interactions, the fourth interaction, gravity, has not yet been included successfully.

2.1.1 Elementary particles and their interactions

All ordinary matter we see around us is built up from atoms, which consist of negatively charged electrons circulating around the positively charged atomic nucleus, formed by protons and neutrons, which in turn consist of up and down quarks. The electrons and up and down quarks

are fundamental particles called fermions. Although all ordinary matter can be built from this so-called first generation of fermions, there are twelve fermions with different flavours in total, six quarks and six leptons, grouped in three generations with increasing mass, as shown in Figure 2.1. The electrically neutral neutrinos interact only via the weak nuclear interaction, which complicates their observation, but they can be detected in dedicated experiments.

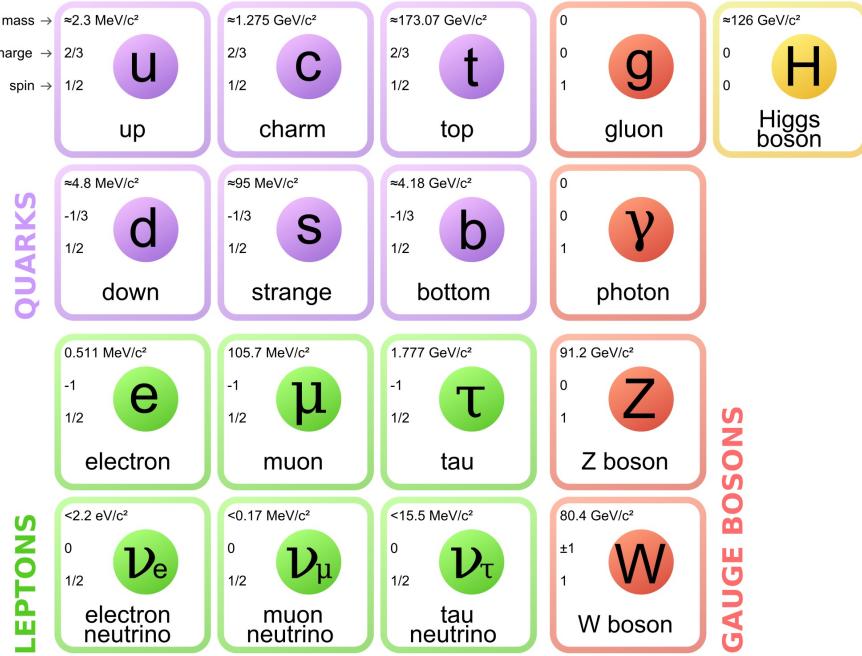


Figure 2.1: The particle content of the Standard Model, showing the fermions divided into 3 generations (columns) on the left and the bosons on the right. The electrical charges are expressed as multiples of the absolute value of the electron charge. Figure taken from [19].

A common characteristic of the fermions is their half-integer spin, in contrast to the integer spin of the force mediators, called bosons. Within the Standard Model, the mediation of the different fundamental interactions is represented by the exchange of these spin-1 gauge bosons, which are summarised in Figure 2.1. The massless photon mediates the most familiar force, the electromagnetic interaction, which is responsible for light, electromagnetic fields, and chemical reactions. The weak nuclear interaction is among other things used to describe the radioactive β decay, and is propagated by the neutral massive Z boson and two charged massive W bosons. Lastly, the strong nuclear interaction is carried by massless gluons, keeping the protons and neutrons in the atomic nuclei and holding the quark constituents together. A resulting property of the quarks is that they hadronise, i.e. they cannot exist isolated, but form bound states via the strong interaction. These bound states are referred to as hadrons, and can be made up from three quarks or a quark and an antiquark, respectively called baryons and mesons.

Finally, it is also important to note that for every fermion (f) there exists an antifermion (\bar{f}), which differs only in electric charge and handedness of spin. When matter and antimatter come into contact they annihilate, generating energy which can be transformed into other particles.

2.1.2 The theoretical framework of the Standard Model

The Standard Model goes further than merely giving an exhaustive list of elementary particles, it has a supporting theoretical framework formulated as a relativistic quantum field theory. In a quantum field theory, every particle is represented by discrete excitations of a field $\psi(x)$, where x is the space-time coordinate. The interactions and kinematics of this particle are fully determined by the action S , which is defined as the integral of the Lagrangian $\mathcal{L}(\psi(x), \partial^\mu \psi(x))$

over the space-time coordinates:

$$S = \int \mathcal{L}(\psi(x), \partial^\mu \psi(x)) d^4x. \quad (2.1)$$

The Lagrangian is a function of the field $\psi(x)$ and its first derivative $\partial^\mu \psi(x)$, where μ represents the index of the space-time coordinate. The physical behaviour of the particles is obtained by following the principle of least action $\delta S = 0$, minimizing the action.

In this framework based on the gauge invariance of the Lagrangian under the fundamental symmetries, the interactions between the fermions and bosons follow automatically. This can be illustrated with the following example for invariance under a general local gauge transformation.

As mentioned before, a fermion has a half integer spin and can thus be represented as a complex relativistic spin-1/2 field, called a Dirac spinor:

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi, \quad (2.2)$$

where γ^μ are the Dirac matrices, and the adjoint field $\bar{\psi} = \psi^\dagger \gamma^0$ is the field associated to the antifermion. The imposed local gauge invariance then requires the fermion fields, and the overall Lagrangian, to be invariant under so-called local phase transformations

$$\psi \rightarrow \psi' = U(x)\psi = e^{i\vec{\alpha}(x) \cdot \frac{\vec{\tau}}{2}} \psi \quad (2.3)$$

where $\vec{\alpha}(x)$ are the space-time dependent rotation parameters in the symmetry group represented by the Lie group generators $\vec{\tau}$. Since the derivative ∂_μ in (2.2) spoils the invariance of the Lagrangian under a local phase transformation, it is replaced with a covariant derivative

$$D_\mu = \partial_\mu - ig\frac{\vec{\tau}}{2} \vec{A}_\mu, \quad (2.4)$$

restoring the invariance. This however introduces new vector gauge fields A_μ , which interact with the fermion fields with a coupling strength g . As a result, the Dirac Lagrangian contains an additional term, which describes the interaction between the fermion fields mediated by the gauge fields A_μ , and (2.2) becomes

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + g\bar{\psi}\gamma^\mu \frac{\vec{\tau}}{2} \psi \vec{A}_\mu \quad (2.5)$$

The matrix $U(x)$ which was introduced above, was defined as a general rotation matrix of the symmetry group $SU(N)$. In order to obtain the three fundamental interactions of the Standard Model, the described procedure can be simplified using the corresponding symmetry groups as mentioned below.

Electroweak theory

The electroweak interaction describes the electromagnetic and weak interactions, which appear very different at low energies but can be merged into a single electroweak force above the electroweak energy scale. This theory is described by requiring gauge invariance under the $SU(2)_L \otimes U(1)_Y$ symmetry group. This leads to 3 gauge fields W_μ^α introduced by the $SU(2)_L$ group, and one gauge field B_μ from the $U(1)_Y$ group. Two coupling constants are introduced, g_1 and g_2 , for $U(1)_Y$ and $SU(2)_L$, respectively. The corresponding observable gauge bosons are then obtained as linear combinations of these fields:

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu, \quad (2.6)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu, \quad (2.7)$$

$$W_\mu^\pm = \sqrt{\frac{1}{2}} (W_\mu^1 \mp i W_\mu^2), \quad (2.8)$$

where A_μ , Z_μ , and W_μ^\pm are the photon, Z^0 , and W^\pm fields, respectively, and θ_W is the weak mixing angle defined as $\tan \theta_W = \frac{g_1}{g_2}$.

Quantum Chromodynamics (QCD)

The strong interaction is described by the theory of Quantum Chromodynamics and is represented by the symmetry group $SU(3)$. It describes the interaction between particles that carry a colour charge, which can be red, green, blue, or one of the three corresponding anticolours. There are eight gauge boson fields associated to this group, which are massless and known as gluons. An important aspect which is unique for this interaction is asymptotic freedom, which states that the strong coupling constant, denoted by α_s , goes to zero at high energies. Consequently, the strong force becomes stronger as the distance between the strongly interacting quarks and gluons increases. As a result, the quarks and gluons cannot exist independently and are not observed individually, but are instead confined in colour-neutral hadrons. This effect is called confinement.

At this point the resulting Lagrangian including the three fundamental forces does not contain any mass terms, and so it cannot explain the observed particle masses. Additional mass terms cannot simply be added explicitly because they would break gauge invariance. Instead, a solution to this problem is found by introducing a complex scalar doublet ϕ with a non-zero vacuum expectation value (vev) v . This breaks the electroweak symmetry and is known as the Brout-Englert-Higgs (BEH) mechanism, postulated in 1964 [14–16]. The Lagrangian of the Higgs field is

$$\begin{aligned}\mathcal{L}_H &= (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi) \\ &= (D^\mu \phi)^\dagger (D_\mu \phi) - \frac{1}{2} \mu^2 \phi^\dagger \phi + \frac{1}{4} \lambda (\phi^\dagger \phi)^2,\end{aligned}\quad (2.9)$$

where μ is a real constant representing a mass parameter and λ is a dimensionless parameter standing for the self-interaction strength. The potential V of the scalar doublet has an infinite set of minima or ground states, and by choosing a ground state and expanding the field around it, the electroweak symmetry is broken. As a result, three of the four original fields of the scalar doublet are absorbed by the massless vector fields of the weak interaction, giving mass to the W and Z bosons:

$$M_W = \frac{1}{2} v g_2 \quad M_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}. \quad (2.10)$$

From the remaining field, the H boson arises, acquiring a mass $m_H = \sqrt{2\lambda v}$.

The introduction of mass terms for the fermions also follows from the BEH mechanism, which allows to insert the following gauge-invariant term in the Lagrangian:

$$\mathcal{L}_{Yukawa} = -Y_{ij} \bar{\psi}_{L,i} \phi \psi_{R,j} + h.c. \quad (2.11)$$

with the Y_{ij} Yukawa matrices. The L and R here denote left-handed and right-handed fermions. This handedness or chirality is defined as $\psi_L = \frac{1}{2}(1 - \gamma_5)\psi$ for left-handed and $\psi_R = \frac{1}{2}(1 + \gamma_5)\psi$ for right-handed fermions. The fermion masses then arise from the Yukawa interactions describing the couplings of the fermions with the Higgs field. For massive particles, a reference frame which overtakes the spinning particle can always be found, in which case the particle will seem to move backwards, flipping its helicity¹. This is however not the case for massless particles, which travel at the speed of light. As only left-handed neutrinos and right-handed antineutrinos have been observed so far, the neutrinos are massless in the Standard Model.

2.1.3 Unanswered questions of the Standard Model

Although the Standard Model is an extremely successful theory, there are still many questions that remain unanswered, indicating that the Standard Model cannot be a complete theory of nature. A brief description of some of the main unsolved problems follows here.

¹The helicity is defined as the sign of the projection of the spin vector onto the momentum vector of a particle, left is negative and right is positive.

Grand Unified Theory

As the weak and electromagnetic interactions were successfully unified into the electroweak one, the idea of representing the three forces of the Standard Model by a single one is envisaged and studied. While this Grand Unified Theory (GUT) could be a first step towards the incorporation of gravity in the Standard Model, it cannot be achieved with the current Standard Model and requires new physics at a very high energy scale.

Baryon asymmetry

This problem refers to the imbalance of matter and antimatter in the universe. While the Big Bang should have produced an equal amount of baryonic and antibaryonic matter, this is not measured in our observable universe. It is assumed that most of the primordial matter and antimatter annihilated, but an imbalance allowed a fraction of the matter to survive. Within the Standard Model, some asymmetry in the production of matter and antimatter could be explained by the CP-violation² of the weak interaction. However, the amount of CP-violation needed to explain the baryon asymmetry is ten times higher than is observed from Standard Model measurements.

Hierarchy problem

This problem is related to the mystery as to why the Higgs boson mass is so much smaller than the Planck scale. The real physical Higgs boson mass, expressed as $m_H^2 = m_{H,\text{bare}}^2 + \mathcal{O}(\Lambda^2)$, is composed of its bare mass and quantum loop corrections. As the Higgs boson is a scalar particle, the corrections depend strongly on the cut-off scale Λ , instead of logarithmically as is the case for renormalisable interactions. At the electroweak scale, which is of the order of 100 GeV, the corrections are of the same order of magnitude as the bare mass. When going to high energy scales such as the Planck scale however, a divergence occurs. A significant fine-tuning would be needed for the bare mass and the huge quadratic radiative corrections to cancel, in order to obtain the experimentally determined mass of 125 GeV. This unnatural amount of fine-tuning is not desirable for any theory. This so-called naturalness problem can be solved by introducing new particles which interact with the Higgs boson in such a way that the correction terms coming from Standard Model particles are partially cancelled.

Neutrino masses

The Standard Model predicts that the neutrinos are massless weakly interacting particles, but observations by the Sudbury Neutrino Observatory [20] and Super-Kamiokande [21] collaborations showed the first clear evidence that the neutrinos oscillate from one flavour into another. This can only be explained if the neutrinos differ in mass, implying that at least 2 of the 3 neutrinos are not massless. As mentioned above, the Standard Model does not provide masses for the neutrinos and it should therefore be extended to explain this observation.

Dark matter and energy

This mystery arises from cosmological observations which indicate that all known matter described by the Standard Model makes up only 5% of the matter and energy in the universe. The remaining matter, called dark matter, contributes another 27%, and will be discussed in more detail in Section 2.2. In the Standard Model, neutrinos contribute to the dark matter, but their relic density is by far not enough to account for all the dark matter. The last 68% has been labelled dark energy and is believed to be responsible for the acceleration of the observed expansion of the universe, but remains even more enigmatic as no explanation can be provided by the Standard Model.

²According to Charge Parity (CP) symmetry, the laws of physics should remain identical when converting a particle into its antiparticle and mirroring the space coordinates. However, measurements of e.g. kaon-antikaon mixing show that this symmetry is violated.

2.2 Dark matter

One of the current open questions in particle physics that is not answered by the Standard Model is the existence of dark matter. Many astrophysical observations from gravitational effects (see for instance [22]) show there must be some additional matter in the universe, the so-called dark matter, next to the known matter. Despite this, its precise nature remains as of yet unknown. Countless theoretical models are being constructed in order to explain its origin, and on the experimental side dark matter is being looked for in many different ways, but no observation has been made so far.

2.2.1 Observational evidence

The first hints of dark matter were observed by F. Zwicky [23] in 1933 by studying the velocity dispersion of galaxies in the Coma cluster. The effect is not only observed for entire galaxies, but also for various luminous objects, such as stars or gas clouds, inside a galaxy. The rotation curves of galaxies have been well studied, and show clear evidence for the existence of dark matter. An example of a rotation curve is shown in Figure 2.2, exhibiting a flat behaviour of the rotational velocity at large distances, going even far beyond the edge of the visible disk. However, in Newtonian dynamics the circular velocity is expected to be

$$v(r) = \sqrt{\frac{GM(r)}{r}}, \quad (2.12)$$

where $M(r) = 4\pi \int \rho(r)r^2 dr$ with $\rho(r)$ the mass density profile. Assuming $M(r)$ to be constant, the circular velocity is expected to fall like $1/\sqrt{r}$ beyond the disk. Since the measurements show an approximately constant velocity but a dropping visible mass density, this implies the existence of a halo with $M(r) \propto r$ and $\rho(r) \propto 1/r^2$. A universal density profile seems to be suggested by the rotation curves of both low and high surface luminosity galaxies, consisting of an exponential thin stellar disk and a spherical dark matter halo with a flat core of radius r_0 and density $\rho_0 = 4.5 \times 10^{-2}(r_0/\text{kpc})^{-2/3}\text{M}_\odot\text{pc}^{-3}$ [24]³.

Another evidence for dark matter comes from the effect of gravitational lensing, allowing to determine the mass of an object regardless of the light it emits. When a distant star or quasar is aligned with a massive compact object, the bending of its light due to the gravitational field of the massive object can lead to multiple distorted, magnified, and brightened images, as illustrated in Figure 2.3. The distortion of the image can then be used to determine the potential well and thus the mass of the heavy object. Yet another way to determine the mass of a cluster of galaxies, next to gravitational lensing and the distribution of radial velocities, is by studying the profile of X-ray emission, tracing the distribution of the hot emitting gas in clusters. In general, these three methods are in reasonable agreement with each other.

Additionally, at a cosmological level, the analysis of the Cosmic Microwave Background (CMB) allows to determine the total amount of dark matter in the universe. The existence of this isotropic background radiation was already predicted in 1948, and unintentionally discovered by A. Penzias and R. Wilson in 1965 [25]. This relic radiation comes from the propagation of photons in the early universe, once they decoupled from matter. Before this, the photons were energetic enough to ionise hydrogen, creating a plasma of electrons and protons which were unable to combine into hydrogen. As the universe expanded and cooled down, the photons also cooled down enough to let the hydrogen atoms recombine, and the universe became transparent. The photons can then travel freely without scattering off the protons and electrons of the plasma, still carrying information from this surface of last scattering. The CMB is now known to be isotropic at the level of 10^{-5} and to follow the spectrum of a black body corresponding to a temperature of 2.726 K. However, small anisotropies in the CMB have first been observed by the COBE satellite [3] and more recently by WMAP [4] and Planck [5], as can be seen in

³ M_\odot denotes a solar mass, 2×10^{30} kg.

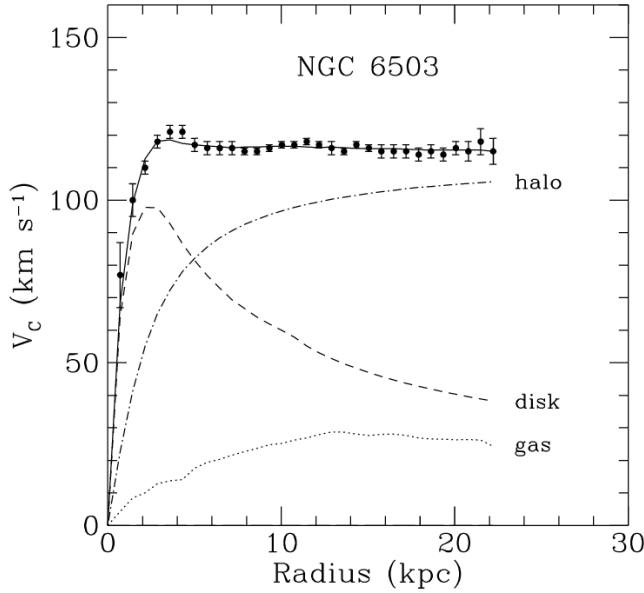


Figure 2.2: Rotation curve of NGC 6503. The dotted, dashed, and dash-dotted lines show the contributions of gas, disk, and dark matter, respectively. Figure taken from [1].

Figure 2.4. These anisotropies correspond to small thermal variations, and are usually expanded as

$$\frac{\delta T}{T}(\theta, \phi) = \sum_{l=2}^{+\infty} \sum_{m=-l}^{+l} a_{lm} Y_{lm}(\theta, \phi), \quad (2.13)$$

where $Y_{lm}(\theta, \phi)$ are spherical harmonics. The variance of a_{lm} is given by

$$C_l = \langle |a_{lm}|^2 \rangle = \frac{1}{2l+1} \sum_{m=-l}^{+l} |a_{lm}|. \quad (2.14)$$

As the temperature fluctuations appear to be Gaussian, the information contained in the CMB anisotropy maps can be condensed into the power spectrum given by the behaviour of C_l as a function of l . This is generally represented using $D_l = l(l+1)C_l/2\pi$, as illustrated in Figure 2.5.

The CMB anisotropies are caused by acoustic oscillations arising from the conflict between the gravitational pull from baryons and dark matter and the repulsive force due to the radiation pressure from the photons. One popular model to describe and interpret these observations is the Λ CDM model. CDM stands for cold dark matter, indicating that in this model the dark matter particles are moving slowly compared to the speed of light, while the Λ represents the cosmological constant, which is associated with the vacuum energy or dark energy that is used to explain the accelerating expansion of space. The Λ CDM model is compatible with a number of observations beyond the CMB fluctuations, such as the large scale structure in the distribution of galaxies, the relative abundance of light nuclei, and the accelerating expansion of the universe which is observed from the red shift of well-known spectral absorption or emission lines in the light of distant galaxies. Using the Λ CDM model to fit the power spectrum of the CMB anisotropies, the multiple peaks in the spectrum can be interpreted. The angular scale of the first peak can be used to determine the curvature of the universe. The second peak determines the reduced baryon density and the third peak can be used to retrieve information about the dark matter density. From the analysis of the Planck data the abundance of baryons and matter in the universe is determined to be

$$\Omega_b h^2 = 0.02205 \pm 0.00028 \quad \Omega_M h^2 = 0.1423 \pm 0.0029 \quad (2.15)$$

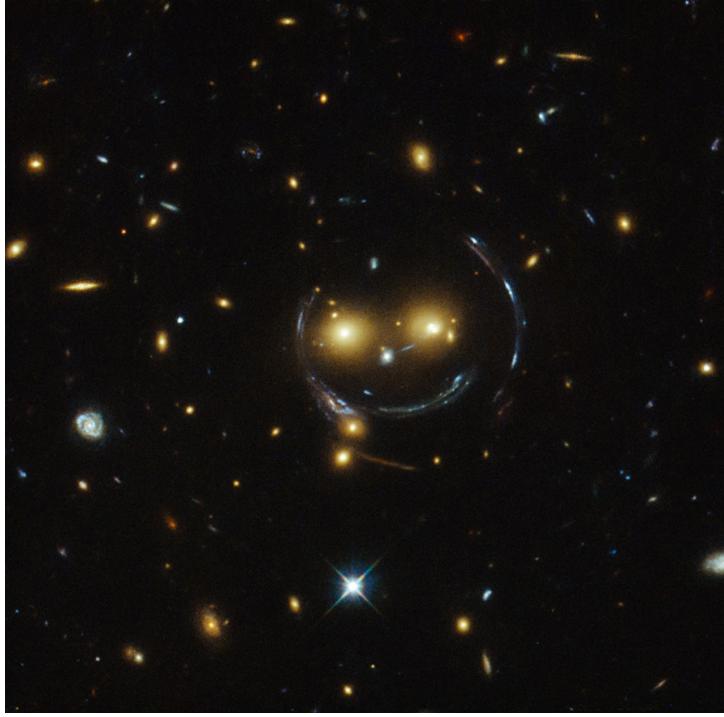


Figure 2.3: An example of gravitational lensing showing the “Cheshire Cat” image of galaxy cluster SDSS J1038+4849, taken by the Hubble Space Telescope. Figure taken from [2].

This result shows that only about 15% of the matter in the universe is made up from the ordinary known matter, and the remaining 85% is called dark matter.

Hot dark matter (HDM), in contrast to CDM, is disfavoured as it is not compatible with the observed small scale structure formation in the universe. Due to the high velocity of the HDM particles, they cannot form clumps as small as galaxies, but would correspond to the formation of larger structures that surround an entire cluster of galaxies. However, models with either a mix of hot and cold dark matter, which could explain both small and large scale structure formation, or models assuming warm dark matter are still being considered.

More evidence for dark matter was found from a great variety of data, both on subgalactic and inter-galactic scales. Without discussing them here in detail, a few examples are the velocity dispersions of spiral galaxy satellites, suggesting the existence of dark halos around spiral galaxies extending well beyond the visible disk [28], the velocity dispersion of dwarf spheroidal galaxies, implying larger mass-to-light ratios than those observed in our local neighbourhood [29], and the so-called Oort discrepancy in the disk of the Milky Way, inferring the existence of dark matter from the inconsistency between the amount of stars in the solar neighbourhood and the gravitational potential indicated by their distribution [30].

2.2.2 Dark matter models

There are two main categories of models that can explain the astrophysical observations detailed in the previous section. The models either predict candidates for a new type of matter, or propose modifications to the laws of gravity which could explain the observations at large scales. An example of theories in the latter category is Modified Newtonian Dynamics (MOND) [31–33]. While Newton’s laws of gravity have been widely tested for large accelerations, they have never been validated for objects that have an extremely low acceleration, such as stars in the outer parts of galaxies. The included modifications result in a rotation velocity which is independent of the distance to the centre of the galaxy, corresponding to a flat rotation curve. These non-relativistic models have also been extended by inserting them in relativistic theories, resulting

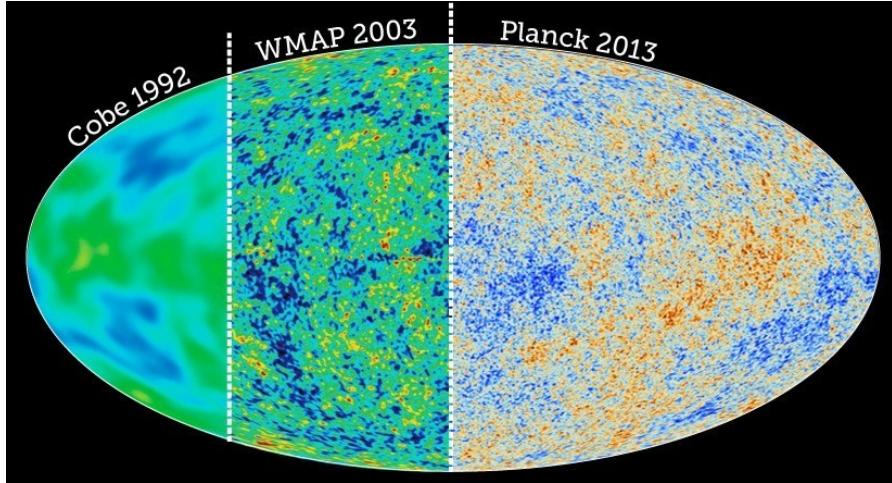


Figure 2.4: The CMB temperature fluctuations obtained from the COBE, WMAP, and Planck data. Figure taken from [26].

for example in a model called TeVeS [34]. As a result, this model can also explain observations from gravitational lensing and structure formation. However, it still faces problems to explain the observed CMB anisotropies. A more recent alternative theory is based on entropic gravity and describes emergent gravity [35].

As these theories are often not complete or are unable to explain all the astrophysical observations, the models focussing on a new type of matter to explain the observed phenomena are usually preferred. Since very little is known so far concerning the nature of dark matter, a multitude of dark matter candidates are discussed in the literature. Without attempting to be complete, a list is given and a few of the more popular candidates are briefly covered here.

Standard Model neutrinos

As mentioned before, the Standard Model could explain the existence of dark matter with the already observed neutrinos. However, it can be shown [36] that their total relic density is predicted to be

$$\Omega_\nu h^2 = \sum_{i=1}^3 \frac{m_i}{93 \text{ eV}}, \quad (2.16)$$

taking the sum over the 3 neutrino flavours. Currently, the most stringent upper bound on neutrino masses is

$$m_\nu < 2.05 \text{ eV} \quad \text{at 95% CL} \quad (2.17)$$

obtained in tritium β -decay experiments at Troitsk [37,38] and Mainz [39]. Since the mass difference between the 3 neutrinos must be very small to explain solar and atmospheric neutrino anomalies [40], this mass limit applies to the three mass eigenvalues, implying an upper bound on the total neutrino relic density of

$$\Omega_\nu h^2 \lesssim 0.07. \quad (2.18)$$

This shows that Standard Model neutrinos are not abundant enough to be the dominant component of dark matter.

Sterile neutrinos

Proposed in 1993 by Dodelson and Widrow [41], these hypothetical particles are similar to the Standard Model neutrinos, but without Standard Model weak interactions, except for mixing. The analysis of their cosmological abundance and the study of their decay products places stringent constraints on the sterile neutrinos. Light neutrinos with masses below a few keV would for example be ruled out [42].

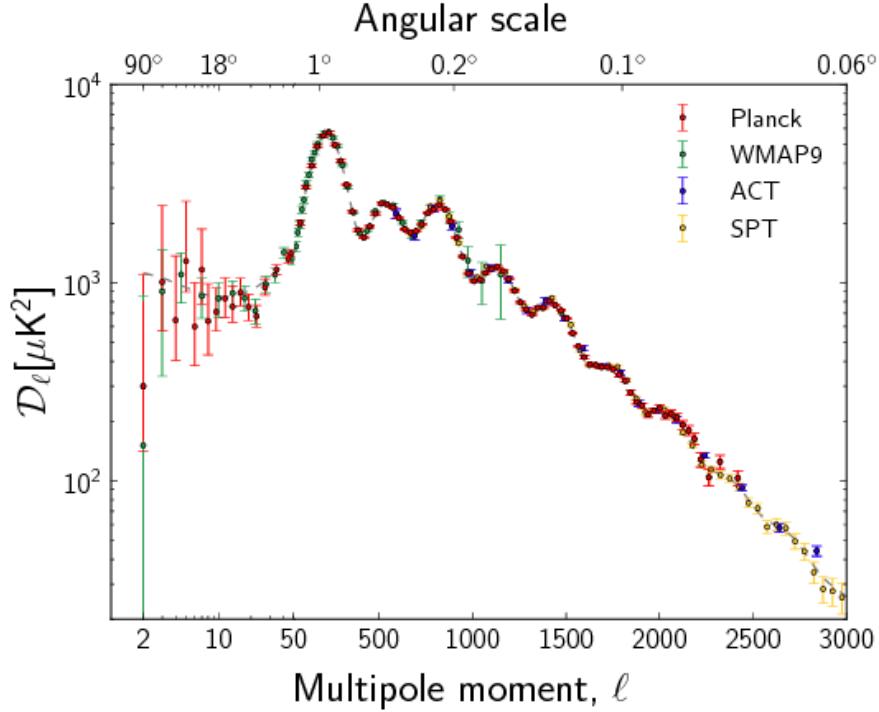


Figure 2.5: The observed power spectrum of the CMB anisotropies. The horizontal axis is logarithmic up to $l = 50$ and linear beyond. Figure taken from [27].

Axions

These particles were originally introduced to solve the problem of the apparent absence of CP-violation by the strong interaction, and have often been discussed as dark matter candidates. They are expected to interact extremely weakly with Standard Model particles. Furthermore, observations from laboratory searches, stellar cooling, and the dynamics of supernova 1987A constrain the axion mass to be very small, of the order of or below 0.01 eV [43]. More recently, axions also appear as dark matter candidates in so-called fuzzy dark matter scenarios as an intriguing alternative to cold dark matter, which is not very successful at scales of about 10 kpc or less, due to the difficulty of modelling processes such as star formation. This alternative is for example in better agreement with some of the observations that are inconsistent with simple predictions from the Λ CDM model, such as the absence of dark matter cusps at the centre of dark matter dominated galaxies or the “missing satellite” problem [44, 45]. In this type of model, the axion mass is expected to be of the order of 10^{-22} to 10^{-21} eV. More details on the motivation and the astrophysical signatures can be found in [46].

Massive primordial black holes

Primordial black holes are a hypothetical type of black holes, formed during an inhomogeneous phase of the Big Bang. First proposed by Stephen Hawking in 1971 [47], they are good dark matter candidates belonging to the class of massive compact halo objects (MACHOs). While they are stable, have non-relativistic velocities, form very early in the universe, and are nearly collision-less, tight limits on their abundance exist from various observations, excluding them as dominant dark matter candidates for most of the preferred mass range. Nevertheless, the recent detection of gravitational waves from merging black holes by the LIGO and VIRGO experiments [48] could point to the existence of primordial black holes [49], and even to a scenario where dark matter is primarily made of primordial black holes [50, 51].

SUSY candidates

Several particles in supersymmetry (SUSY) models can serve as dark matter candidate, such as gravitinos and neutralinos. Gravitinos are the superpartners of the graviton. In some SUSY models, they can be the lightest supersymmetry particle and can be stable. While they are very strongly motivated theoretically, they are very difficult to observe, as they only interact gravitationally. The neutralinos are the superpartners of the photon, Z boson, and neutral Higgs bosons. The lightest of the four can be stable and is an excellent dark matter candidate. These dark matter candidates are often called weakly interacting massive particles (WIMPs), since they are massive and interact through the weak interaction. As many SUSY models predict a new particle with the correct properties and self-annihilation cross section to obtain the correct abundance of dark matter today, a stable supersymmetric partner has long been a very plausible dark matter candidate and a lot of experimental effort has been made to detect it.

WIMPs

A prevalent assumption is that dark matter particles are a relic from the early universe, when all particles were in thermal equilibrium. At those high energies the dark matter particles and antiparticles could be formed by sufficiently energetic lighter particles, and they would annihilate back into these lighter particles as well. However, as the universe expanded and cooled down, the thermal energy of the lighter particles became insufficient to form dark matter particle-antiparticle pairs. The annihilation of the dark matter particles and antiparticles continued however, until the dark matter density decreased considerably and the interaction between the dark matter particles stopped. The number of dark matter particles would remain constant from that moment on, which is referred to as “freeze-out”. This process is illustrated in Figure 2.6.

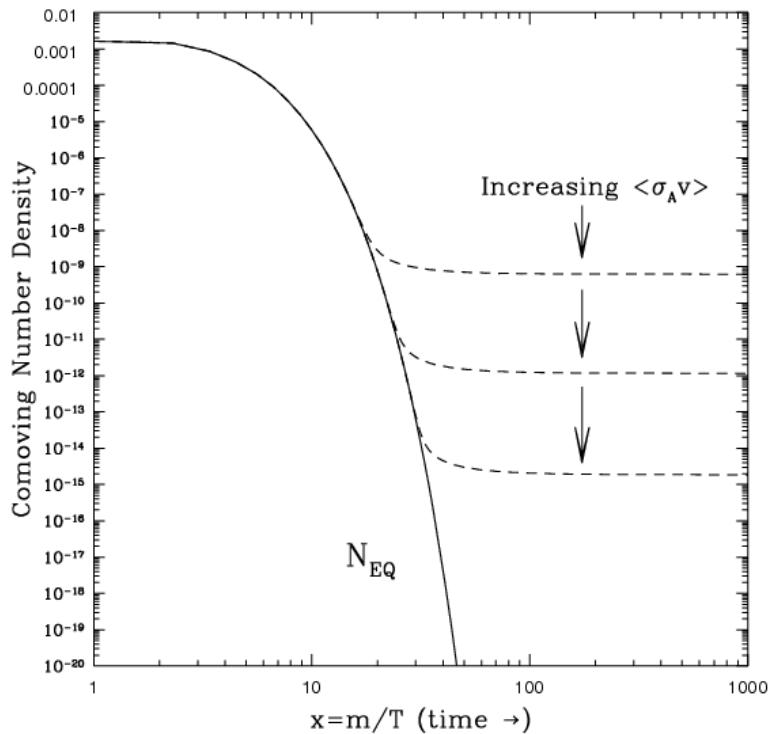


Figure 2.6: The evolution of the number density of a thermal relic in time, showing the dropping number density and the freeze-out where the density becomes constant, due to the universe expanding and cooling down. Figure taken from [52].

When the temperature in the universe is lowered below the dark matter mass, the dark matter abundance freezes out to a value which is suppressed by $e^{-\frac{m_\chi}{T}}$, and the number density of the dark matter is given by:

$$n = \left(\frac{m_\chi T}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{m_\chi c^2}{kT}}, \quad (2.19)$$

with m_χ the dark matter mass and T the temperature of the universe. However, as the universe is expanding at a rate given by the Hubble parameter H , the freeze-out will occur when the annihilation rate becomes smaller than the expansion rate of the universe. The annihilation rate of the dark matter particles is given by $n\sigma v$, with σ the annihilation cross section which depends on the dark matter model parameters, and v the relative velocity of the two colliding dark matter particles. The resulting relation is quantified by the Boltzmann equation:

$$\dot{n} + 3Hn = \langle \sigma v \rangle (n_{EQ}^2 - n^2), \quad (2.20)$$

with n_{EQ} the number density at equilibrium. Particles with a larger interaction cross section would then continue to annihilate for a longer period of time and would be less abundant.

The interaction cross section of the annihilating dark matter particles can be inferred from the current estimates of the dark matter abundance in the universe, and can in this case not be larger than the cross section of the weak interaction. According to this model, WIMPs would be the perfect candidates for dark matter. In general, they are hypothetical new elementary particles that interact gravitationally and through any other force which is as weak or weaker than the Standard Model weak interaction, and they could have been produced thermally as described in this model. Since WIMPs have a relatively large mass, they would also constitute cold dark matter, which would fit the observed large scale structure of the universe. The coincidence of WIMPs fitting so well into this model and corresponding to the current observations is known as the “WIMP miracle”.

Many more dark matter candidates are discussed in literature, such as but not limited to heavy fourth generation neutrinos [53], Kaluza-Klein states in ADD [54] or RS [55] extra dimensions models, superheavy dark matter or Wimpzillas [56], self-interacting dark matter [57], charged massive particles (CHAMPs) [58], and Q-balls [59]. More detailed reviews are given in [36, 60, 61].

2.2.3 Detection of dark matter

The detection of dark matter can be categorised in three groups, based on the diagram shown in Figure 2.7. In the case of direct detection experiments, the studied process is the scattering of dark matter off ordinary matter. Experiments searching for dark matter with the indirect approach look for particles or radiation produced in the annihilation of dark matter particles. Finally, at collider experiments, attempts are made to produce and detect dark matter particles by colliding Standard Model particles at high energies.

2.2.3.1 Direct detection experiments

This category of experiments is based on the fact that if our galaxy is filled with a static halo of dark matter particles, then many of them should pass through the Earth as it rotates around Galactic Center, and they could be detected by looking for the interaction of such particles with matter. This is for example done by recording the recoil energy of nuclei when WIMPs scatter off them. In order to determine the expected rate of events per unit detector material mass, the WIMP-nucleon scattering cross section and the density and velocity distribution of the WIMPs in the solar neighbourhood are needed.

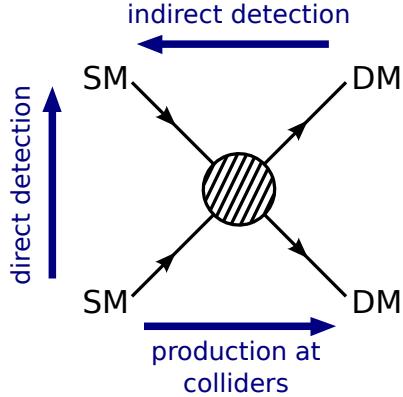


Figure 2.7: Diagram illustrating the three used methods to detect dark matter.

There are several types of scattering processes which can be classified according to two relevant characteristics: elastic or inelastic scattering and spin-dependent or spin-independent scattering. In the case of elastic scattering, the WIMP simply scatters off a nucleus as a whole, causing it to recoil. The recoil energy spectrum can then be measured by detecting the emitted scintillation light with very sensitive detectors or by measuring very small temperature changes due to crystal vibrations. Taking a Maxwell-Boltzmann velocity distribution with a characteristic velocity our galactic rotation velocity of about 270 km/s, the recoil spectrum is exponential with typical energies of $\langle E \rangle \sim 50$ keV. This range of energies is easily detectable by current experiments, which can detect recoils as low as 1-10 keV. Instead, when the WIMP scatters inelastically, it interacts with the orbital electrons of the target, exciting the electrons or ionising the target. Differently, the WIMP could also excite the target nuclei, which would then emit a photon about a nanosecond after the observed recoil. This signature has, however, to compete with the background from natural radioactivity.

The spin dependence or independence of the scattering depends on the coupling of the WIMPs to the Standard Model particles. Spin-dependent interactions result from couplings to the spin content of a nucleon, yielding cross sections that are proportional to $J(J+1)$ instead of the number of nucleons. For spin-independent interactions, the cross section instead increases considerably with the mass of the target nuclei. The sensitivity for spin-independent scattering is therefore much enhanced over the spin-dependent one in experiments which use heavy atoms.

Numerous direct detection experiments are currently operational or in development. They use one or more techniques to measure the nuclear recoil, by detecting the scintillation light, the change in temperature, or the ionisation. Some experiments also try to separate the WIMP signatures from the background by looking for an annual modulation in the rate, which arises due to the Earth's movement around the Sun. This effect causes the Earth to have a relative velocity with respect to the galaxy's reference frame, given by

$$v_E = 270 \text{ km/s} (1.05 + 0.07 \cos(2\pi(t - t_m))), \quad (2.21)$$

where the time is in units of years and t_m is approximately the beginning of June. As a result, a small variation of about 7% in the WIMP flux can be measured in the direct detection rate.

Currently, there is some tension between the results obtained by the different experiments, as some observations can be interpreted as dark matter signals, while other experiments are ruling out those models. The DAMA experiment for example observes an annual modulation in the event rate, pointing to the existence of WIMPs scattering elastically off the sodium and iodine nuclei in the detector [62]. Other experiments, such as SuperCDMS [63], EDELWEISS-III [64], CRESST-II [65], XENON100 [66], have seen no evidence for dark matter so far and placed limits on many dark matter models, creating a tension with the observed signal at DAMA. For WIMP masses above a few GeV, the strongest limit of direct detection experiments for spin-independent interactions is currently given by LUX [67]. For a spin-dependent WIMP-proton

cross section, the most stringent limit is set by the PICO experiment [68], while the PandaX experiment places the strongest limit on the WIMP-neutron cross section [69]. An overview of the existing limits and signal observations is given in Figure 2.8, showing the mentioned experiments, and a more complete review of the existing direct detection results is given in [70].

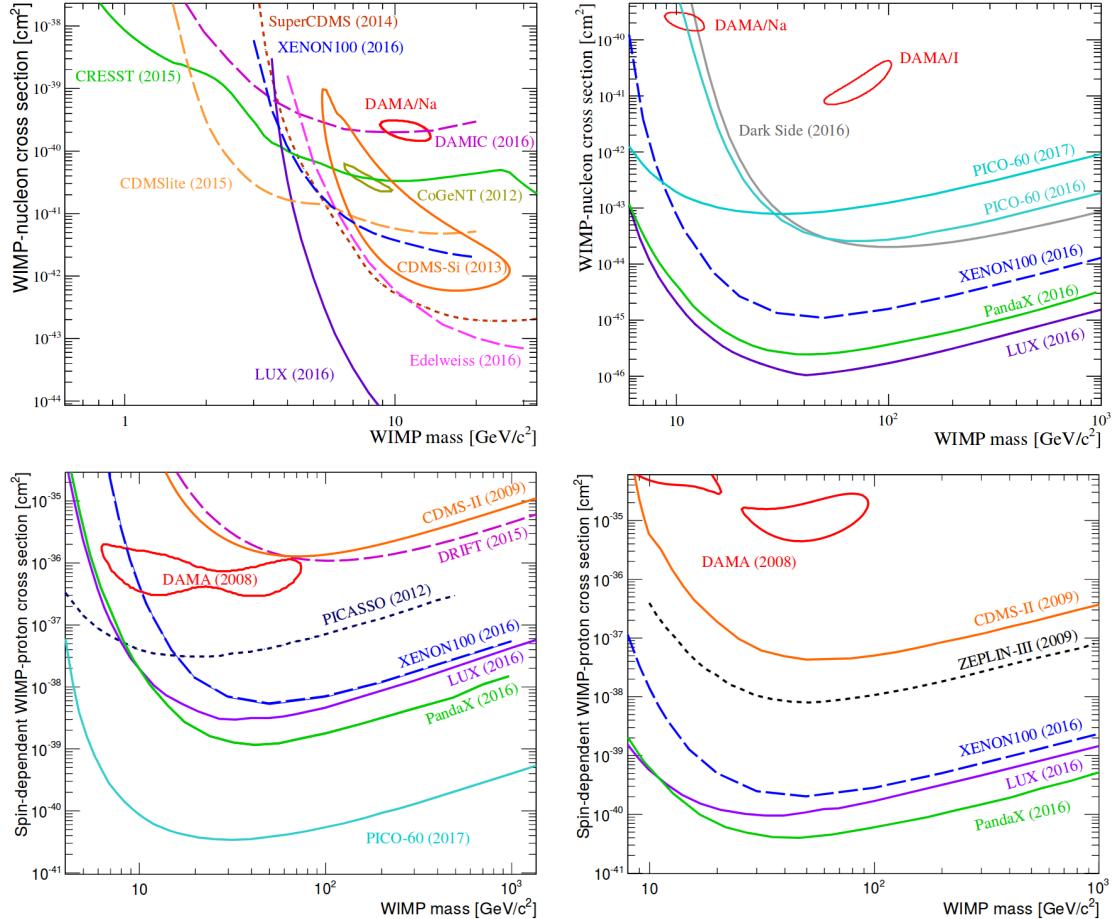


Figure 2.8: Overview of the current limits for spin-independent WIMP-nucleon interactions at low (top left) and high (top right) WIMP masses, spin-dependent WIMP-proton interactions (bottom left), and spin-dependent WIMP-neutron interactions (bottom right, the y axis title wrongly states WIMP-proton cross section). The observed signals from DAMA, CoGeNT, and CDMS-Si are shown as well. Figure taken from [70].

2.2.3.2 Indirect detection experiments

The indirect detection of dark matter is performed by looking for radiation produced in dark matter annihilations. A reasonable place to look at would then be in regions with large dark matter densities and thus larger annihilation rates, which will result in a higher flux of the studied radiation. Some examples are dense regions of the galactic halo such as the galactic centre, or objects like the Sun or the Earth, which could also capture dark matter particles through scattering with nucleons in their core. In the latter case, only neutrinos would be able to escape those dense objects. Other annihilation products include gamma rays, positrons, and antiprotons.

In order to observe gamma rays directly, the detectors must be placed in space, as photons of the relevant energy range (GeV to TeV) interact with matter via e^+e^- pair production and cannot traverse more than a surface density of about 38 g cm^{-2} . The gamma rays will not reach the ground-based telescopes as the Earth's atmosphere is 1030 g cm^{-2} thick. Neverthe-

less, efforts are being made to observe gamma rays indirectly via ground-based experiments as well, by detecting the secondary particles and the Cherenkov light produced by their passage through the Earth's atmosphere. In the energy range between approximately 100 MeV and 100 GeV, gamma ray telescopes on satellites such as the Fermi Large Area Telescope [71] are being used. Above 100 GeV, the ground-based Imaging Air Cherenkov Telescopes such as HESS [72], MAGIC [73], and VERITAS [74] become more adequate.

Neutrinos can also be produced in the annihilation of dark matter particles, but they are considerably more difficult to detect than gamma rays due to their weak interaction with ordinary matter. They are not easily absorbed, which makes it possible to observe them with underground, low-background experiments. Very energetic neutrinos, in the GeV-TeV range, are most easily observed by detecting the Cherenkov light from muons produced through charge current interactions of the neutrinos inside of or close to the detector volume. Two very large neutrino detectors are ANTARES in the Mediterranean Sea [75] and IceCube at the South Pole [76].

Additionally, evidence for dark matter annihilations can also be found by studying the spectra of cosmic positrons and antiprotons. Contrary to neutrinos and gamma rays, these charged particles do not point to their source, as their trajectory is modified by the presence of galactic magnetic fields. Currently, the main detector for positrons and antiprotons is AMS [77], which is operating on the International Space Station. Until 2016, PAMELA [78] was also active on board of the Resurs-DK1 satellite.

Finally, radio emissions from the galactic halo, and in particular from the galactic centre, can also provide evidence for dark matter annihilation. Electrons and protons produced in dark matter annihilations will emit synchrotron radiation at radio wavelengths as they move through galactic magnetic fields. This type of searches is performed with radio telescopes and belongs to the realm of classical astronomy.

2.2.3.3 Collider experiments

Since dark matter particles are usually assumed to be neutral and to interact only weakly with ordinary matter, they are expected to pass through the detectors at colliders without leaving a signal, similar to neutrinos. These particles can however still be searched for at colliders as well, when they are produced in association with other visible particles which are detected as jets or charged leptons. The dark matter particles are then observed as missing energy, as they create an imbalance in the net momentum in the transverse plane perpendicular to the colliding beams, which should be zero. One of these flagship analyses is the monojet analysis, looking for dark matter produced together with one or more jets [79, 80]. Similarly, many more searches are performed at the CMS and ATLAS experiments at the LHC by looking for signatures containing missing energy. Recent summaries are given in [81] and [82].

Additionally, other signatures without missing energy can also be used to search for dark matter. If the dark matter particle is produced in a cascade of decays for example, different signatures can be obtained, such as displaced vertices [83], disappearing tracks [84], and displaced lepton-jets [85]. Furthermore, in dijet searches [86–88], resonances in the mass spectrum are being looked for, as this could point to the existence of a new dark matter mediator. If the dark matter particles couple to quarks via a dark matter mediator, this mediator can either decay to a pair of dark matter particles or a pair of Standard model quarks which can be observed as a pair of jets. Finally, for some particular types of dark matter candidates, such as strongly interacting massive particles (SIMPs) [89] or heavy stable charged particles (HSCPs) [90, 91], more unusual signatures are expected. This is currently a developing area of dark matter research, and more and more searches looking for new signatures are appearing.

In Figures 2.9 and 2.10, recent limits from dark matter searches at the CMS experiment are compared to the direct detection results, for spin-dependent and spin-independent interactions, respectively.

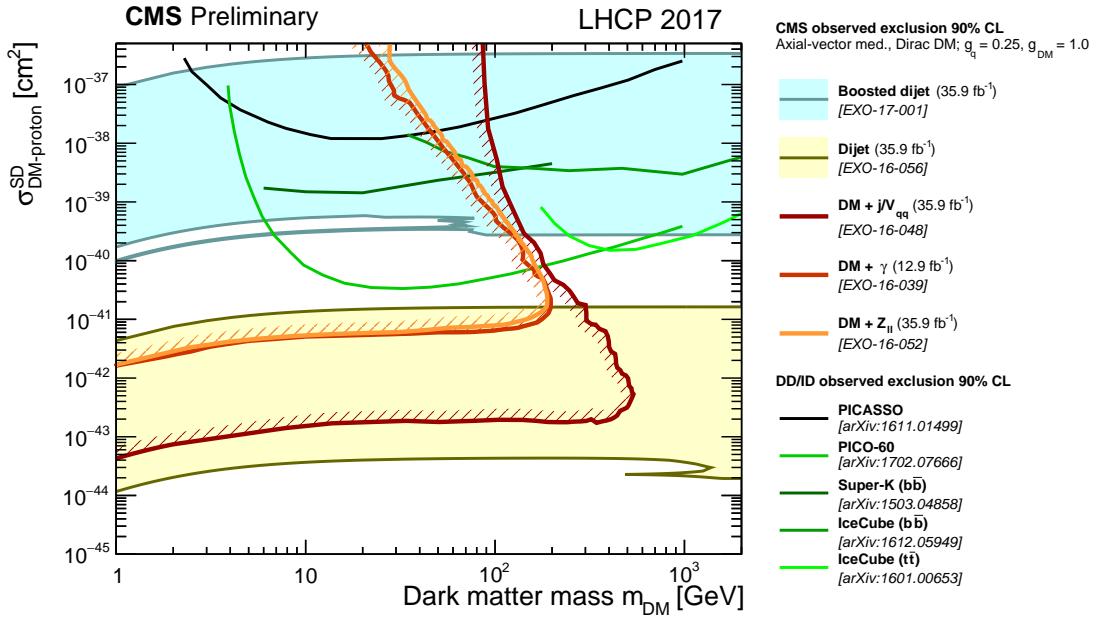


Figure 2.9: A comparison of CMS results to direct detection experiments in the $m_{\text{DM}} - \sigma_{\text{SD}}$ plane. The limits are shown at 90% CL. The shown CMS contours are for an axial-vector mediator with Dirac dark matter and couplings $g_q = 0.25$ and $g_{\text{DM}} = 1.0$. The spin-dependent exclusion contours are compared with limits from the PICASSO and PICO experiments, the IceCube limit for the $t\bar{t}$ and $b\bar{b}$ annihilation channels, and the Super-Kamiokande limit for the $b\bar{b}$ annihilation channel. It should be noted that the CMS limits do not include a constraint on the relic density and also the absolute exclusion of the different CMS searches as well as their relative importance will strongly depend on the chosen coupling and model scenario. Therefore, the shown CMS exclusion regions in this plot are not applicable to other choices of coupling values or models. Figure taken from [92].

2.2.4 From EFTs to simplified models

In order to efficiently look for dark matter at colliders, effective field theories (EFTs) have been used extensively to model the dark matter signal. The EFT models assume the dark matter production can be described as a contact interaction defined by an effective mass scale and coupling structure. This contact interaction is for example illustrated in Figure 2.11 for the monojet final state where the dark matter pair is produced in association with an initial state radiation jet. The resulting signal models can then be classified by coupling structure, and the effective scale Λ can be extracted for a specific model, defining both the coupling strength and the scale of the theory. An EFT is characterized by a total of 3 parameters, the dark matter mass, the EFT scale, and the EFT coupling structure. However, this approach has several limitations [93–95]. First, it implicitly assumes that the dark matter production happens through a heavy mediator, which is not resonantly enhanced at the LHC. Additionally, for low enough effective scales, the EFT breaks down. Finally, the incompleteness of the EFT makes a comparison with direct detection experiments difficult or inconsistent. Due to the limitations of EFTs, there has been a trend in the past few years to instead use simplified models which allow for a fair comparison to low energy underground direct detection experiments. In a simplified model the effective scale is then replaced by a physical mediator. The resulting models contain six parameters that can be scanned to search for dark matter, namely the coupling structure, the dark matter mass, the mediator scale, the couplings to the Standard Model and the dark matter, and, if the mediator also couples to other particles than the quarks and dark matter particles, the mediator width. This transition has been overseen by the joint ATLAS/CMS dark matter forum [96] by establishing a

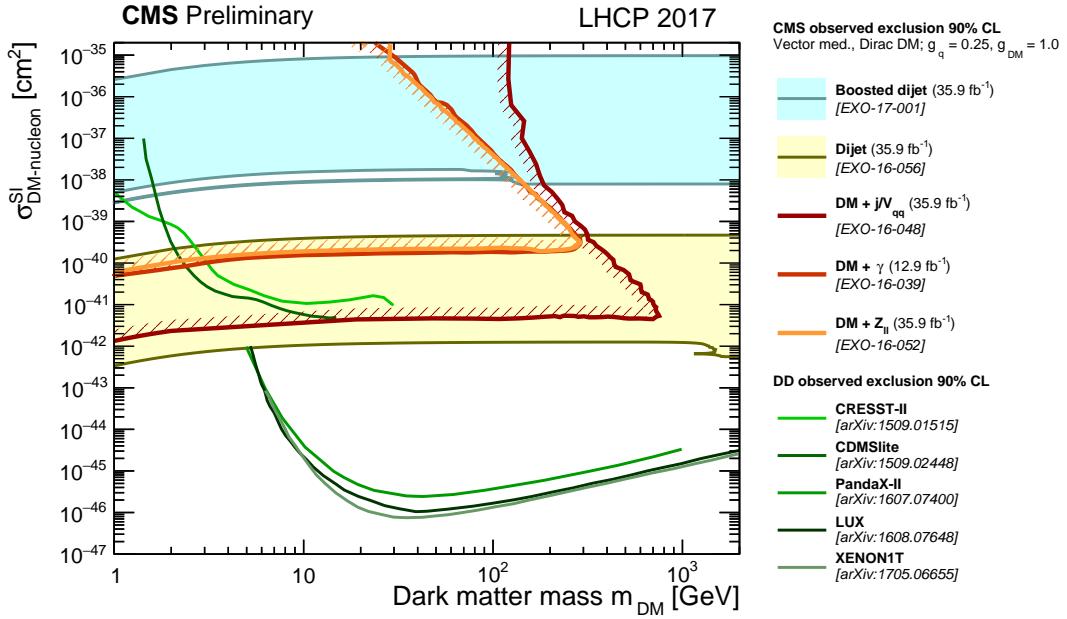


Figure 2.10: A comparison of CMS results to direct detection experiments in the $m_{\text{DM}} - \sigma_{\text{SI}}$ plane. The limits are shown at 90% CL. The shown CMS contours are for a vector mediator with Dirac dark matter and couplings $g_q = 0.25$ and $g_{\text{DM}} = 1.0$. The spin-independent exclusion contours are compared with the XENON1T 2017, LUX 2016, PandaX-II 2016, CDMSlite 2015 and CRESST-II 2015 limits, which constitutes the strongest documented constraints in the shown mass range. It should be noted that the CMS limits do not include a constraint on the relic density and also the absolute exclusion of the different CMS searches as well as their relative importance will strongly depend on the chosen coupling and model scenario. Therefore, the shown CMS exclusion regions in this plot are not applicable to other choices of coupling values or models. Figure taken from [92].

well defined set of benchmark models to enable the combination of different channels and the recasting of dark matter models against direct and indirect detection searches [97].

In the two dark matter searches covered in this thesis, the results have been interpreted in terms of simplified models. The monojet search described in Chapter 5 includes several simplified models recommended by the dark matter forum. Four types of mediators are considered, i.e. a vector, axial-vector, scalar, and pseudoscalar mediator. In the case of a scalar or pseudoscalar coupling, the production mode is dominated by gluon fusion. As illustrated in the right diagram of Figure 2.12, the scalar is produced through a t or b quark loop. A Yukawa coupling is assumed for the coupling of the mediator to Standard Model particles, proportional to the mass of the particle. For a vector or axial-vector mediator, the production happens through the fusion of two quarks into a heavy mediator, similarly to the Z and W boson production. The coupling to quarks and potentially leptons is taken to be unity, and universal for all flavours. For all mediator types, the coupling to the dark matter particles is assumed to be unity. In addition, the minimal width assumption is made, implying that the mediator couples to all Standard Model and the dark matter particle and no extra particles are introduced. If such particles would be present, the width would increase and the sensitivity of the analysis would be reduced. A scan is then performed over the mass of the dark matter candidate and the mass of the mediator.

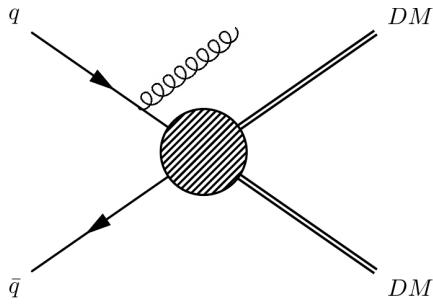


Figure 2.11: Illustration of EFT dark matter production in the monojet final state.

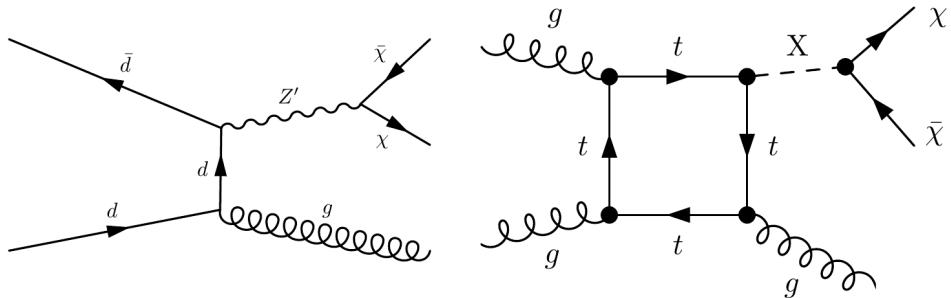


Figure 2.12: The vector (left) and scalar (right) production diagrams in the monojet final state.

2.3 Strongly Interacting Massive Particles

As no observation of dark matter has been made so far, despite many searches probing the more popular models described in the previous section, many scenarios now venture beyond minimal models or give up basic assumptions for the WIMP. In the following model, which is studied in this thesis, the interaction cross section of the dark matter with normal matter is so high that the particles are no longer WIMPs, but so-called SIMPs. This model is not part of the models recommended by the dark matter forum, but can also be motivated by the long lasting interest for self-interacting dark matter (SIDM)⁴ particles with a large cross section [57], which could help to explain observations that present a challenge for the cold dark matter scenarios, such as the missing satellites or core-cusp problems [98–101]. While it is possible to create models with a strongly interacting hidden sector that is weakly coupled to the Standard Model particles, SIDM particles that interact rather strongly with the known matter particles can be considered as well.

2.3.1 The SIMP simplified model

In this simplified model, the dark matter particles χ can be produced at the LHC in pairs, through a new strong interaction with a new mediator ϕ , as illustrated in Figure 2.13. These SIMPs are neutral and stable, and are generated off-shell as the mediator is very light, of the order of the pion mass: $m_\phi = 140$ MeV. We only consider the case of fermionic candidates, since the bosonic form is ruled out by constraints coming from neutron stars and black holes, as is described in Section 2.3.2. Both the cases with a scalar or a vector mediator can be studied, and

⁴Incidentally, self-interacting and strongly interacting share the same abbreviation, such that SIDM can also stand for strongly interacting dark matter and SIMP for self-interacting massive particles in the literature.

the corresponding interaction Lagrangian is

$$\mathcal{L}_{\text{int}} = \begin{cases} -g_\chi \phi \bar{\chi} \chi - g_q \phi \bar{q} q & \text{(scalar mediator)} \\ -\tilde{g}_\chi \phi_\mu \bar{\chi} \gamma^\mu \chi - \tilde{g}_q \phi_\mu \bar{q} \gamma^\mu q & \text{(vector mediator)} \end{cases} \quad (2.22)$$

with $g_\chi g_q, \tilde{g}_\chi \tilde{g}_q < 0$ to avoid the formation of bound states. For simplicity we assume that the SIMPs have a universal coupling to quarks, although a flavour dependent coupling could be preferred, as light SIMPs with a significant coupling to b or c quarks are probably constrained by B and D meson phenomenology. SIMPs lighter than about 5 GeV could for example appear in the decay of b or c quarks, and would be constrained by limits on the invisible decay of B and D mesons.

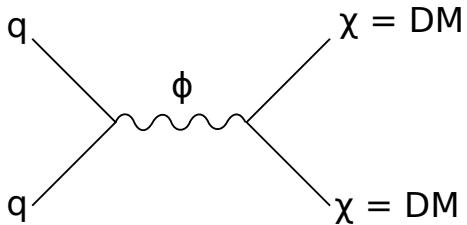


Figure 2.13: Feynman diagram showing the production of a SIMP pair, through a new low-mass mediator.

Introducing a new strong interaction between quarks can however modify nuclear potentials. In order to keep the impact small, the mediator is assumed to not modify nuclear potentials by more than $\mathcal{O}(10\%)$, such that $g_{\chi N} \lesssim 0.3 g_{\pi NN} \sim 3$ for a mediator with the mass of a pion, where $g_{\pi NN} \sim 13$ is the effective pseudoscalar pion-nucleon coupling [102]. The quoted values are however not very precise, as a large spread of values can be found in the literature for meson-nucleon effective couplings, sometimes differing by a factor 2 or more (see e.g. [103] for comparison). This shows the difficulty of dealing with strong interactions in the framework of effective field theories, which arise because contributions from many mesons need to be taken into account, each of them with a different coupling. No constraints on modified strong interactions at low energies seem to exist in literature so far, however searches at fixed-target experiments do place constraints on the existence of strongly interacting stable neutral particles.

In summary, the model has 4 free parameters: the two couplings, the mass of the mediator m_ϕ , and the mass of the SIMP m_χ . At the LHC, only the product of the couplings appears, while astrophysical observations constrain both the dark matter self-interaction and the interaction with the Standard Model.

2.3.2 Experimental constraints

Naively, one would expect such an unusual model with strong interactions not to be viable, as various types of experiments and observations set constraints on SIMPs as dark matter candidates. However, some of these limitations can be avoided by the assumptions in the model described above. The relevant existing measurements are described below, showing there is still a part of phase space which remained unexplored so far.

Bound states

Searches for heavy isotopes, in particular heavy water, constrain the formation of bound states between SIMPs and nucleons, ruling out particles with a mass below 10 TeV for the scenario with SIMPs as dominant contribution to dark matter. This constraint is evaded by assuming a purely repulsive SIMP-nucleon interaction with opposite sign couplings,

as is specified in the Lagrangian (2.22). In the vector mediator case, vector mediators would however couple to the dark matter antiparticles with an opposite charge. This is avoided if no dark matter antiparticles are around, i.e. if the abundance of dark matter is asymmetric. A reason for having asymmetric SIMPs is that if they are the dominant source of dark matter, then the dark matter abundance is set by either an asymmetry or through a non-thermal mechanism. In the case of a symmetric SIMP candidate, the dark matter abundance is determined by thermal freeze-out, and it can only be a sub-dominant component. Additional constraints also exist on the dark matter self-interacting strength from halo shapes and merging galaxies such as the Bullet cluster [104, 105].

Earth heating

A second argument for an asymmetric abundance of SIMPs comes from experiments measuring the heat emitted from the Earth's core. For the typical SIMPs cross sections, the dark matter particles can be captured by the Earth and accumulate in its core over time. Annihilating SIMPs would then provide a substantial source of heat and could modify the Earth's heat flow. This can be measured by detectors in deep underground shafts [106] and rules out the scenario with symmetric SIMPs.

Neutron stars and black holes

In the asymmetric scenario, light scalar dark matter particles can however be collected in the cores of neutron stars and cause them to collapse into black holes. Bosonic dark matter candidates are therefore excluded, and we consider only fermionic candidates as mentioned previously.

Direct detection searches

Many bounds on the SIMP parameter space come from the direct detection searches as well. Underground experiments, such as CDMS and XENON, place strong constraints at smaller cross sections, about 5 orders of magnitude below the SIMP cross section, as can be seen from Figure 2.14. At the higher cross sections considered here, the SIMPs are stopped by the Earth's atmosphere, and they cannot reach the underground detectors. At higher altitudes however, space or airborne experiments such as RSS [107], a balloon-based experiment with a silicon semiconductor detector, and XQC [108], a sounding rocket experiment, exclude SIMPs in some regions of phase space. More details on these constraints can be found in [106], where they have been extensively reviewed.

Nucleosynthesis and cosmic rays

There are also bounds from primordial nucleosynthesis and cosmic rays, reviewed in [109] and [110]. The protons in cosmic rays can scatter off dark matter particles and create neutral pions, which decay to photons and could be detected in gamma ray telescopes. Although limits have been placed on dark matter-nucleon interactions [110], these constraints depend on many assumptions and adopt a form of the dark matter density near the galactic core. Since the considered model describes a nonstandard form of dark matter with a relatively strong interaction with baryons, these densities may be considerably different.

CMB and large scale structure

Observations of the CMB anisotropies and the large scale structure power spectrum, including from the Lyman- α data [111, 112] additionally also place strong constraints on interactions between dark matter and baryons.

Fixed-target experiments

Finally, a relatively old fixed-target experiment led in 1976 at FNAL with a beam of neutral particles produced by 300 GeV protons hitting a beryllium target was used to look for massive, strongly interacting, neutral particles [113]. The mass of the particles was determined using their flight time and their kinetic energy which was measured in a

calorimeter. Neutral particles with a mass larger than 2 GeV were searched for, in order to discriminate the candidates from the background of neutrons and lighter hadronic states, up to $m_\chi \lesssim \sqrt{E}/2 \approx 12$ GeV, limited by the beam energy of $E = 300$ GeV. Single particle production was considered, but the results apply to pair production as well when they are translated into the case where 2 neutral particles are boosted and fly away in the same direction. The search showed no significant excess above the expected background and limits were placed on the invariant production cross section per nucleon versus the neutral particle interaction cross section. As an example, for an interaction cross section of 1 mb, a limit on the total production cross section of about $2.5 \times 10^{-35} \text{ cm}^2 = 25 \text{ pb}$ is found, and this limit is reported by the Particle Data Group [114]. Comparing the considered SIMP model to this result by simulating the pair production at $\sqrt{s} = 25$ GeV, one can conclude that SIMPs between 2 and about 6 GeV are already excluded by this experiment [6].

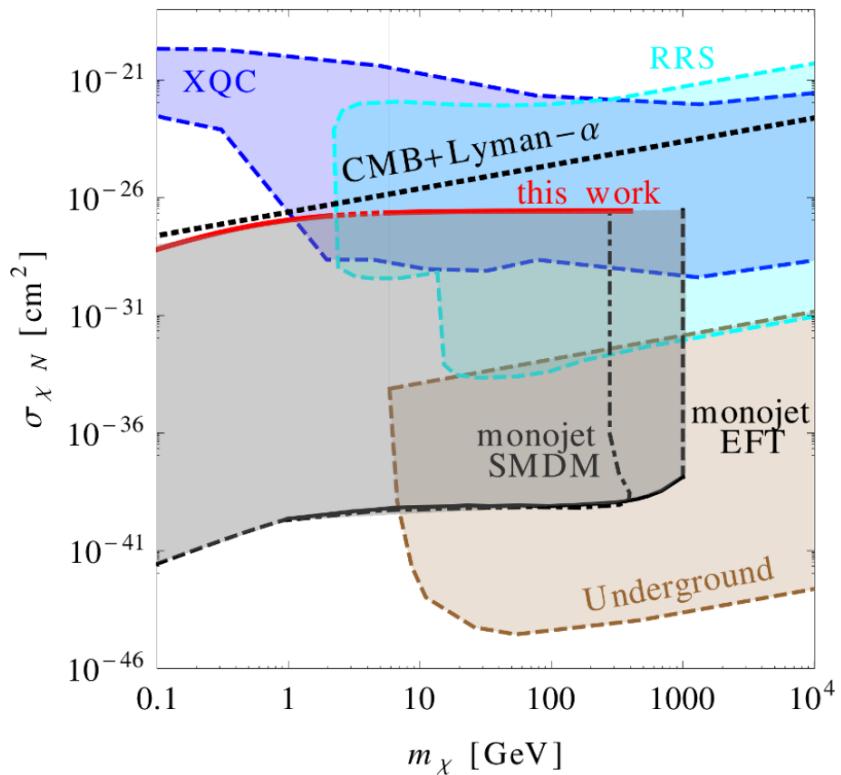


Figure 2.14: Summary plot showing the SIMP model (red) in comparison with the most important applicable constraints, coming from the LHC monojet analyses (black), the atmospheric XQC and RRS experiments (blue), underground experiments (brown), and the CMB observations and Lyman- α data (black dashed line). Figure taken from [6].

3

The LHC and the CMS Detector

In order to investigate the currently unsolved mysteries of particle physics, such as the existence of dark matter, many experiments can be conducted, among other things at particle colliders. The largest particle accelerator in the world is the LHC, located at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. At this accelerator, protons are being accelerated at energies up to 6.5 TeV, giving rise to a record centre-of-mass energy of 13 TeV in the proton collisions. Using data from the collisions generated at the interaction points along the accelerator ring, the Standard Model can be tested in many ways and searches for particles beyond the Standard Model can be performed.

In Section 3.1 more details are given about the LHC and the 4 main experiments situated at the interaction points. In particular, the general-purpose Compact Muon Solenoid (CMS) detector is described in Section 3.2.

3.1 The Large Hadron Collider at CERN

The LHC was built in the already existing Large Electron Positron (LEP) collider tunnel, which was excavated in the 1980's and has a circumference of 27.6 km. Contrary to the LEP collider, the LHC accelerates particles of the same charge, namely protons or lead ions. Much higher luminosities can therefore be reached, since only particles are used and the generation of antiparticles is not needed. This was the limiting factor at the Tevatron, where protons and antiprotons were used. Additionally, at the probed energies the colliding particles are not the protons or ions, but their constituents, which carry a varying fraction of the total momentum. This makes the LHC an ideal instrument for exploration at higher energies, as the collisions naturally cover a wide energy range.

3.1.1 The LHC injector chain

The protons (or lead ions) can not directly be injected in the LHC, but need to be accelerated gradually in several pre-accelerators, as illustrated in Figure 3.1. For the proton beams, the LHC injection chain starts at a bottle of hydrogen, where protons are stripped from the hydrogen atoms and accelerated up to 50 MeV by a linear accelerator (LINAC2). The protons are then transferred to a chain of circular accelerators, starting with the Proton Synchrotron Booster (PSB) which accelerates them to an energy of 1.4 GeV. Next, the protons go through the Proton Synchrotron (PS) and are delivered to the Super Proton Synchrotron (SPS) at an energy of 26 GeV. Finally, the protons are injected in the LHC in opposite direction with an energy of 450 GeV.

The lead ions are first accelerated in a different linear accelerator, LINAC3, before being injected in the Low Energy Ion Ring (LEIR) at an energy of 4.5 MeV per nucleon. Here the ions are accelerated to an energy of 72 MeV per nucleon, and they then follow the same path as the protons through the PS, where they are accelerated to 5.9 GeV and stripped from the last of their electrons, and the SPS, where they are accelerated to 177 GeV. The record centre-of-mass energy for heavy ion collisions at the LHC so far has been 5.02 TeV and 8.16 TeV, for lead-lead (Pb-Pb) and proton-lead (p-Pb) collisions in 2015 and 2016, respectively.

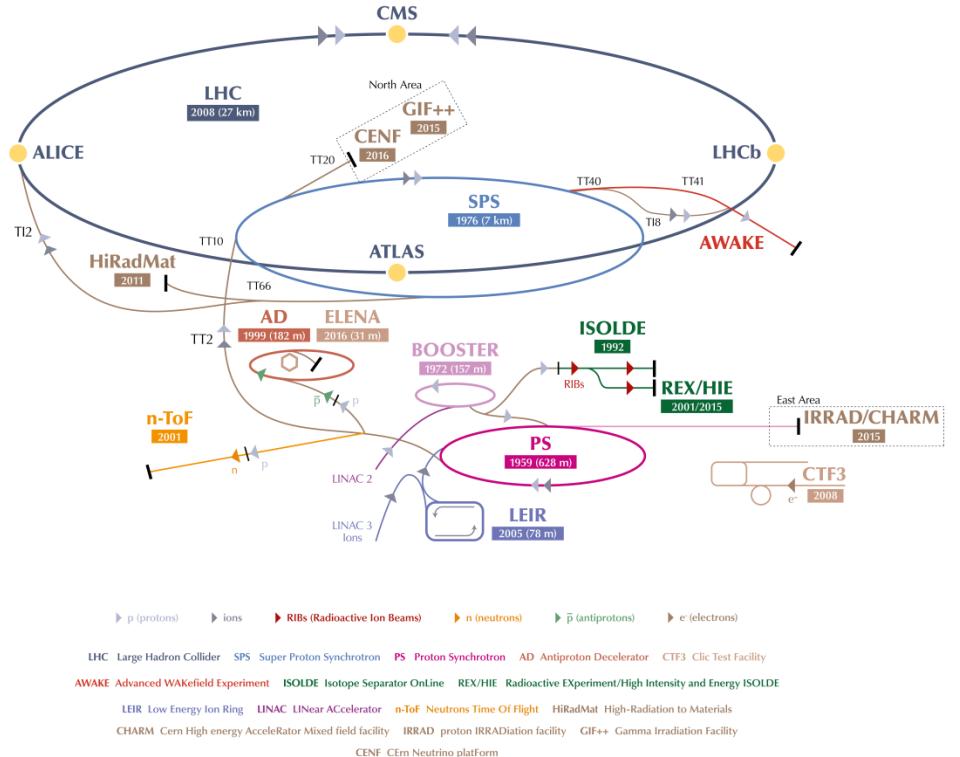


Figure 3.1: Schematic view of the various linear and circular accelerators of the CERN accelerator complex, including the LHC injection chain. Figure taken from [115].

3.1.2 The Large Hadron Collider

The most relevant specifications for a particle physics accelerator are the maximum energy and the luminosity that can be reached. High energy is necessary in order to be able to create new heavy particles, which are for example predicted in many theories beyond the Standard Model.

The protons are kept on the correct orbit by the 1232 LHC dipole magnets. These magnets are cooled down to 1.9 K with liquid Helium and supplied with a current of 12 kA to reach the design field of 8.33 T. This limits the maximum beam momentum of the accelerator to

$$p = B/\rho = 8.33 \text{ T}/2804 \text{ m} = 7 \text{ TeV}/c, \quad (3.1)$$

with ρ the bending radius of the tunnel. The protons are accelerated up to the desired energy by radio-frequency (RF) cavities, which produce an oscillating electric field.

A high event rate or luminosity \mathcal{L} is equally important, to obtain a sufficiently high number of collisions. For a process with cross section σ , this rate is

$$\frac{dN}{dt} = \mathcal{L}\sigma. \quad (3.2)$$

In order to achieve the high design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ in the LHC, the protons are concentrated in bunches which are focused by strong quadrupole magnets around the interaction regions. The RF frequency is 400 MHz, which corresponds to a wavelength of 75 cm and confines the protons into buckets of 2.5 ns. The LHC circumference would then allow for 35640 buckets in total, which would result in bunch crossings spaced by only 37.5 cm. A more realistic spacing is 25 ns, corresponding to one bunch per 10 RF buckets. Additionally, 225 ns gaps are present between the bunch trains, corresponding to the rise time of the injection kicker magnets. One gap of 3 μs is necessary as well to allow clean beam dumps. These requirements limit the number of bunches to a maximum of 2808.

After almost 25 years of design and construction, the LHC was completed in 2008 and the commissioning of the machine started. However, during a powering test on 19 September of the same year an electric arc developed inside a bus bar which led to a large release of helium and a pressure wave that caused extensive mechanical damage to the affected LHC sector. This incident delayed the first collisions, with one bunch per beam and at a beam energy of 900 GeV, until late 2009. During 2010 and 2011 a centre-of-mass energy of 7 TeV was used for the collisions, which was then increased to 8 TeV in 2012. The instantaneous luminosity was also increased, starting from $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ in 2010 to more than $6 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ in 2012. During the 3 years of data-taking in Run 1, data corresponding to an integrated luminosity of 45.0 pb^{-1} , 6.1 fb^{-1} , and 23.3 fb^{-1} , respectively, were delivered. After Run 1, a long shutdown (LS1) of 2 years followed, which was used to correct the problems that were discovered in the aftermath of the incident at the startup in 2008, and to upgrade and consolidate the experiments located on the LHC ring.

In 2015, the LHC restarted operations with Run 2, at an even higher centre-of-mass energy of 13 TeV. During 2016 the design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ was exceeded and a total of 41 fb^{-1} of data were delivered. A comparison of the delivered integrated luminosity per year is shown in Figure 3.2.

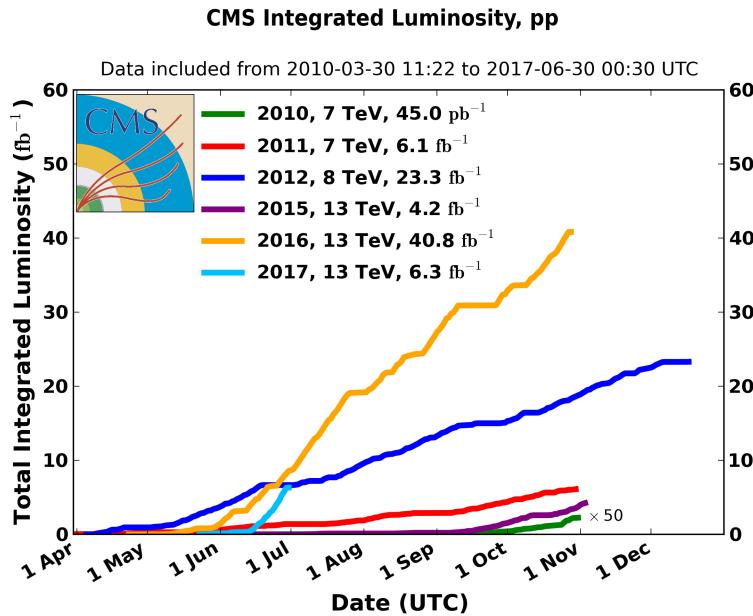


Figure 3.2: Overview of the integrated luminosity delivered to the CMS detector during Run 1 (2010 to 2012) and Run 2 (2015 to 2017).

3.1.3 The experiments at the LHC

There are four interaction points (IPs) where the proton or lead ion beams of the LHC can collide, and around each of these points large particle detectors were built in order to record the generated collisions. The ATLAS and CMS detectors, located at IP1 and IP5, are both high luminosity general-purpose detectors and consist of several layers surrounding the IP in an onion-like structure to avoid particles escaping detection. These detectors can cover a wide range of high energy physics, from precision measurements of the Standard Model to searches beyond the Standard Model. At IP2 the ALICE detector is specialized in heavy ion collisions with low instantaneous luminosities, around $10^{27}\text{cm}^{-2}\text{s}^{-1}$. With this detector information is gathered about the quark-gluon plasma, a state of matter that exists at extremely high temperatures and densities where quarks and gluons are no longer confined in hadrons. The fourth main detector, LHCb, is located at IP8 and requires instantaneous luminosities of the order of a few $10^{32}\text{cm}^{-2}\text{s}^{-1}$. Using this detector b quarks are being studied, focusing among other things on the matter-antimatter asymmetry in the universe.

3.2 The CMS detector

The searches described in this thesis were conducted using data collected with the CMS detector, a general-purpose detector located on the LHC ring. It consists of the typical components of a particle physics detector, namely a tracker, an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), a solenoidal magnet, and muon detectors. One peculiar aspect is however that both calorimeters are situated inside the superconducting magnet. This design was chosen in order to improve the energy resolution by reducing the amount of material in front of the calorimeters. The overall detector has a length of 21.6 m, a diameter of 14.6 m and a total weight of 12500 t. A schematic overview is shown in Figure 3.3.

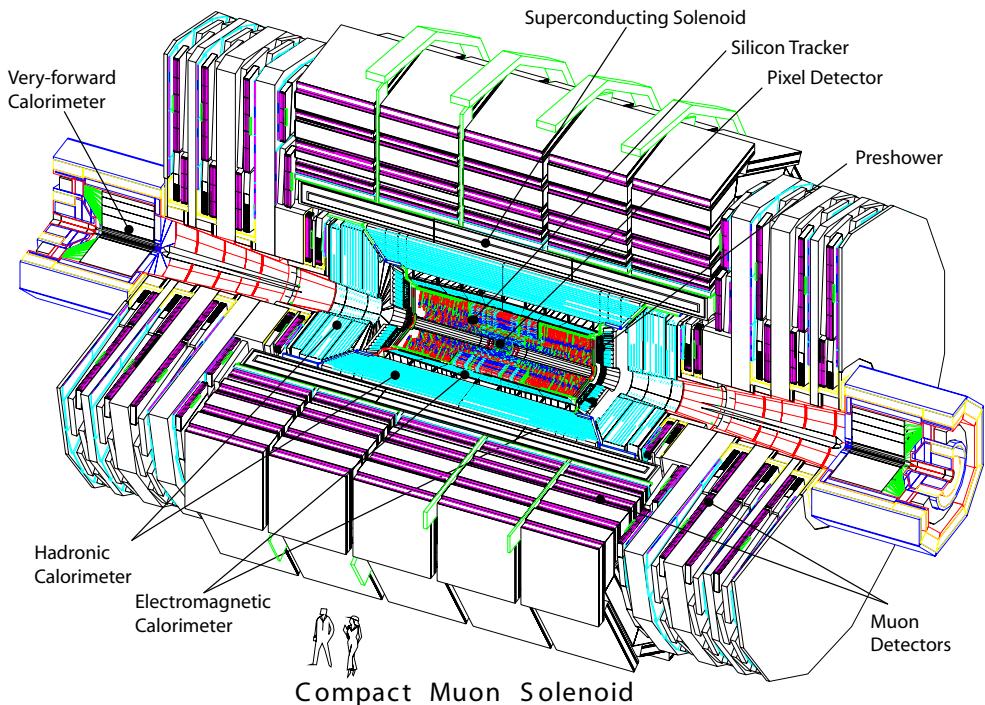


Figure 3.3: The CMS detector, consisting of the pixel and strip tracker, the electromagnetic calorimeter (ECAL) with preshower, the hadronic calorimeter (HCAL) with its forward component, and the muon systems. Figure taken from [116].

The CMS coordinate system places the origin at the nominal collision point. The x axis is perpendicular to the beam and points towards the centre of the LHC ring, the y axis is vertical and pointing upwards, and the z axis is defined anticlockwise along the beam direction. The azimuthal angle ϕ is then defined in the xy plane, relative to the x axis and the polar angle θ is measured with respect to the z axis. In general, the polar angle is converted into the pseudorapidity

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \quad (3.3)$$

for convenience, since differences in pseudorapidity are invariant under Lorentz boosts along the z axis. A pseudorapidity of 0 corresponds to the direction perpendicular to the beam ($\theta = \pi/2$), and an infinite pseudorapidity corresponds to the direction parallel to the beam ($\theta = 0$).

Due to the conservation of momentum before and after the collision, the momenta of the particles in the final state of a collision should be balanced in the transverse plane. Another variable that is therefore often used in particle physics is the transverse momentum of a particle, defined as

$$p_T = p \cdot \sin \theta. \quad (3.4)$$

3.2.1 The tracker

The innermost part of the CMS detector, closest to the IP, is the tracking system, which is designed to provide a precise measurement of the trajectories of charged particles. This all-silicon detector is divided into a pixel and a strip detector, with a layout as shown in Figure 3.4. The inner part, consisting of pixel modules, provides very precise 3D hits, which are important for vertex reconstruction and track seeding. This allows to have a precise measurement of secondary vertices and track impact parameters, necessary for the efficient identification of e.g. heavy flavour particles. As the hit occupancy is lower in the outer part of the detector, a larger cell size can be afforded, and silicon strips are used instead of pixels. This strip detector provides a large lever arm and a link to the calorimeters and the muon system. The tracker covers a pseudorapidity range $|\eta| < 2.5$.

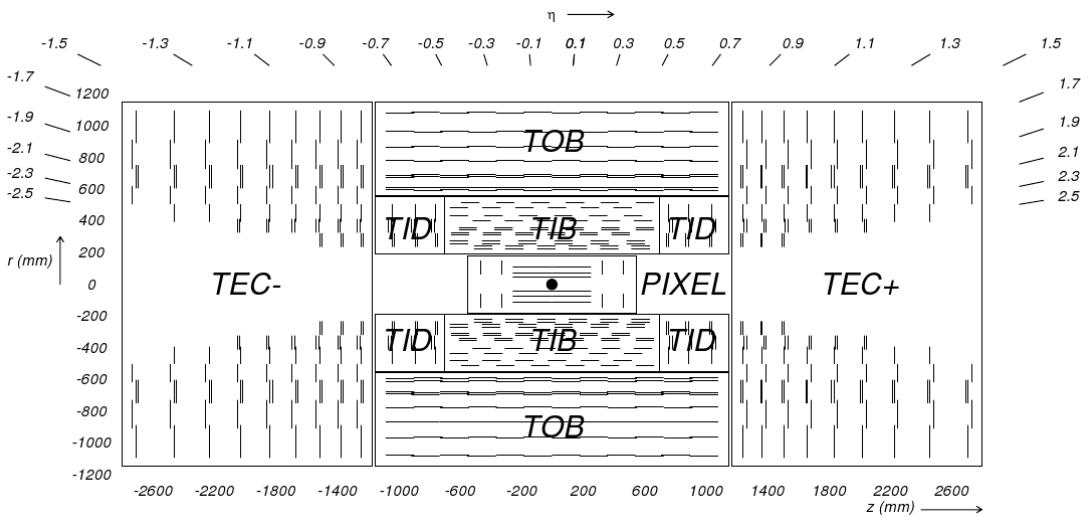


Figure 3.4: A transverse view of the pixel and strip tracker detectors. Figure taken from [116].

3.2.1.1 The pixel tracker

The pixel tracker was replaced during the extended technical stop in 2016 and 2017 [117], as a part of the CMS Phase 1 upgrades. As the data used for this thesis were recorded before that,

only the so-called Phase 0 detector is described here.

For the pixel modules n+ pixels on n-substrate are used, allowing the sensors to also work in under-depletion after type inversion. The 1440 modules are arranged in several cylindrical layers and disks, as illustrated in Figure 3.4. The barrel, consisting of 3 pixel layers surrounding the beam pipe at radii of 4.4, 7.3 and 10.2 cm, is complemented by the forward pixel detector, composed of 2 endcap disks on each side extending from 6 to 15 cm in radius. The barrel and the forward parts contain respectively 48 million and 18 million pixels with a size of $100 \times 150 \mu\text{m}^2$, covering a total area of 1.06 m^2 .

In the barrel, the magnetic field of CMS is perpendicular to the drift of the electrons to the collecting pixels, which results in a Lorentz drift. This drift leads to a spread of the charge over several pixels. Since the read-out of the modules is analogue, an improved spatial resolution can therefore be achieved with charge interpolation. In the forward pixel detector the drift of the electrons would be parallel to the magnetic field so in order to profit from the Lorentz angle, the modules are tilted by 20° in a turbine-like arrangement, as can be seen in Figure 3.5. A spatial resolution of $10 \mu\text{m}$ ($30 \mu\text{m}$) can be achieved in the local directions x (y) of the module, respectively. In the barrel x is the longitudinal direction perpendicular to the beam and y is the longitudinal direction parallel to the beam.

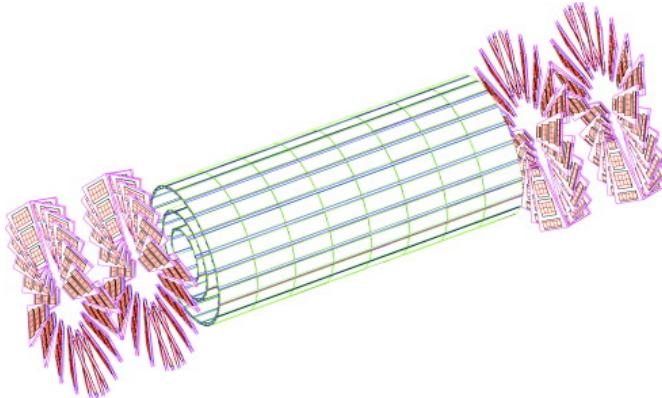


Figure 3.5: A 3D view of the barrel and forward pixel detector. Figure taken from [118].

The signals from the pixel sensors are read out by custom read-out chips (ROCs), which amplify and store the signals, and already apply zero-suppression on-detector. The data rate from the detector to the Front End Drivers (FEDs) is therefore not constant for every event. Additionally, if there are too many hits on a pixel module for a given event, they can not all be stored on the finite buffer of the ROC. Consequently, as the instantaneous luminosity increases the pixel modules start to show a “dynamic inefficiency” which is most pronounced in the first layer, closest to the beampipe. This was one of the main motivations for the Phase 1 upgrade of the pixel detector.

3.2.1.2 The strip tracker

The outer part of the tracker consists of 15 148 strip modules, which are distributed among multiple barrel layers and endcap disks and make up a total active area of 198 m^2 . The inner part is composed of 4 Tracker Inner Barrel (TIB) layers with 3 Tracker Inner Disks (TID) on each side. Surrounding these are 6 Tracker Outer Barrel (TOB) layers and the 2 Tracker EndCaps (TEC), which are composed of 9 disks. This geometric arrangement is shown in Figure 3.4, with double lines to indicate back-to-back modules. These so-called double-sided modules are mounted with a stereo angle of 100 mrad to improve the 3D point resolution by providing a measurement of the z and r coordinate in the barrel and disks, respectively. The choice of strip pitches is driven by the two particle separation capability and two-hit resolution, and ranges

from 80 μm to 205 μm . The length of the strips varies from 63 mm to 117 mm, minimizing the occupancy and noise levels.

In the TOB and the 3 outermost rings of the TEC two silicon sensors are daisy chained, while single sensors are used in the inner part. This is done to limit the number of read-out channels, since the area that had to be instrumented is larger in the outer region. The larger cell size can be afforded due to the lower occupancy in the outer part. However, the noise of the sensors also increases with strip length, so thicker silicon sensors, 500 μm compared to 320 μm in the inner part, are used in order to collect more signal per traversing particle.

The strip sensors are single sided p-on-n type silicon. The signals from the sensors are amplified, shaped, and stored by 4 or 6 custom APV25 chips per module. When the trigger has made a positive decision, the analogue signals from two APV25 chips are multiplexed and sent to the FED boards in the service cavern via optical fibres, where they are converted to digital signals. The FEDs then perform pedestal and common mode subtraction as well as cluster finding. Additionally, the data is sparsified in these off-detector electronics, before being sent to the CMS central data acquisition (DAQ). Due to charge sharing, this analogue read-out scheme also results in an improved spatial resolution of 15 to 40 μm , depending on the position of the modules and the strip pitch.

3.2.2 The electromagnetic calorimeter

Surrounding the tracker, the CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of photons and electrons. It is composed of 75 848 lead tungstate (PbWO_4) crystals arranged in a cylindrical barrel and 2 endcaps. The barrel crystals measure $22 \times 22 \text{ mm}^2$ at the front face of crystal, and $26 \times 26 \text{ mm}^2$ at the rear face, which corresponds to approximately 0.0174×0.0174 in $\eta\phi$. The length of the crystal is 230 mm, corresponding to 25.8 radiation lengths. In the endcaps, the crystals have a rear face cross section of $30 \times 30 \text{ mm}^2$, front face cross section of $28.62 \times 28.62 \text{ mm}^2$, and a length of 220 mm, corresponding to 24.7 radiation lengths. The crystals are arranged such that they are off-pointing with respect to the centre of the detector, and instead point at a focus 1.3 m beyond the interaction point, in order to avoid cracks between the crystals to be aligned with the trajectories of the particles.

The high density material was chosen due to its short radiation length and small Molière radius, resulting in a small lateral spread of the electromagnetic shower generated by an incoming photon or electron. This allows for a fine granularity, a better shower separation, and a compact calorimeter. Additionally, this scintillating material has a fast response, as about 80% of the light is emitted during the first 25 ns. The scintillation light is collected by photodetectors, digitized, and read out.

The layout of the ECAL is shown in Figure 3.6, with the barrel (EB) extending up to $|\eta| < 1.470$ and the endcaps (EE) on each side covering the range $1.479 < |\eta| < 3.0$. A preshower detector (ES) is positioned in front of the endcap crystals, covering the pseudorapidity range between $|\eta| = 1.653$ and $|\eta| = 2.6$. This detector consists of a layer of lead which initiates an electromagnetic shower from incoming photons or electrons, and a layer of silicon sensors which measures the deposited energy. The main goal of this 20 cm thick detector is to discriminate between photons and neutral pions.

The energy resolution of calorimeters can be parametrized by the following stochastic (S), noise (N), and constant (C) terms:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (3.5)$$

The stochastic term represents contributions from the shower containment, the number of photoelectrons and the fluctuations in the gain process. The noise term takes into account all noise components, such as electronics and digitization noise. Finally, the constant term characterizes among others energy leakage from the back of the calorimeter crystals and non-uniformities of

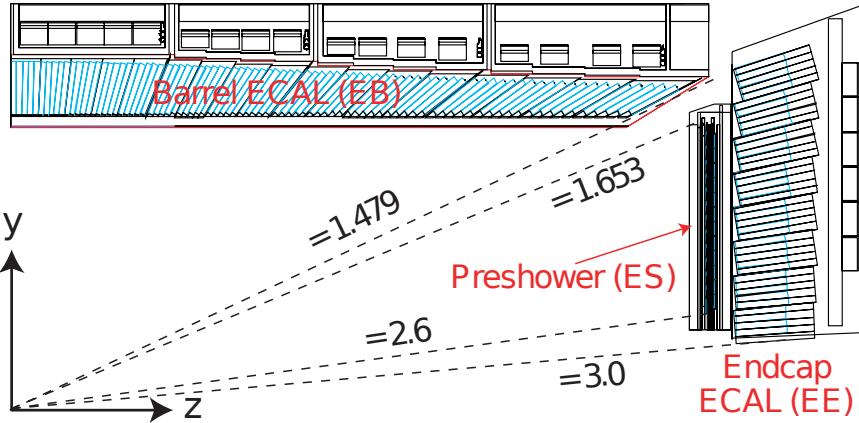


Figure 3.6: A transverse view parallel to the beamline showing one quarter of the ECAL, with its barrel (EB), endcap (EE), and preshower (ES) detectors. Figure taken from [118].

the longitudinal light collection. The latter term dominates the energy resolution for high-energy electron and photon showers. Figure 3.7 shows the energy dependence of this resolution for incident electrons as measured in a beam test, as well as the determined stochastic, noise, and constant terms obtained by fitting equation 3.5 to the data.

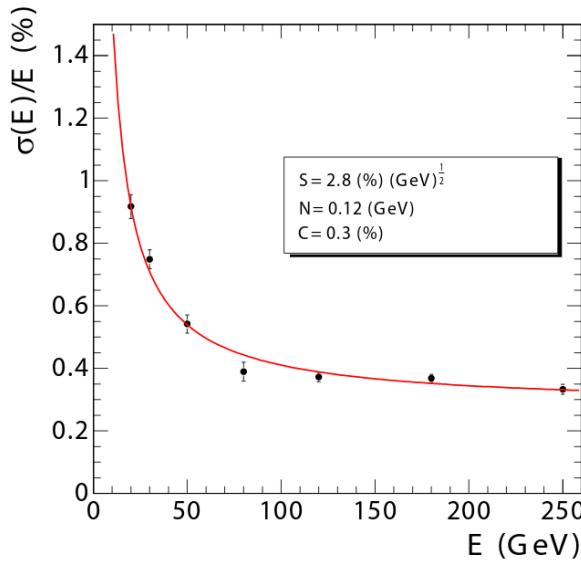


Figure 3.7: The ECAL energy resolution as a function of the electron energy, measured from a beam test. The stochastic (S), noise (N), and constant (C) are given as well. Figure taken from [116].

A more recent measurement of the energy resolution was performed using electrons from Z boson decays in collision data. In the central region, up to $|\eta| < 0.8$, it was measured to be better than 2%. Outside of this region, in the more forward direction, the energy resolution is 2-5% [119]. The reconstruction of the electrons and photons will be discussed in Section 4.3.2.

3.2.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) surrounds the ECAL with the aim to measure the energy of charged and neutral hadrons. The missing transverse energy can then be inferred from this measurement together with the measured energies in the ECAL, in order to identify neutrinos or exotic particles. The HCAL consists of brass absorber plates interleaved with plastic scintillator tiles.

Figure 3.8 shows a longitudinal quarter view of the different HCAL components. A cylindrical barrel (HB) covers the region up to $|\eta| < 1.4$ and is complemented by endcaps (HE) on each side, extending the pseudorapidity range to $|\eta| < 3.0$. In the central region, the stopping power of the ECAL and HCAL barrel is not sufficient to contain the entire hadron showers of very energetic jets. The HCAL was therefore extended outside the solenoid with an outer calorimeter (HO), which uses the magnet coil as absorber and consists of scintillators. Two layers are positioned at $\eta = 0$, where the absorber depth is minimal, and only 1 layer is used for the 2 rings on each side of the central ring. Finally, a forward calorimeter (HF) is positioned at 11.2 m from the IP covering $3.0 < |\eta| < 5.2$. Unlike the other HCAL components, this detector consists of iron and quartz fibres, which are radiation-hard. This is an important aspect for the choice of active material since it is exposed to very large particle fluxes. A signal is generated by the charged shower particles which generate Cherenkov light.

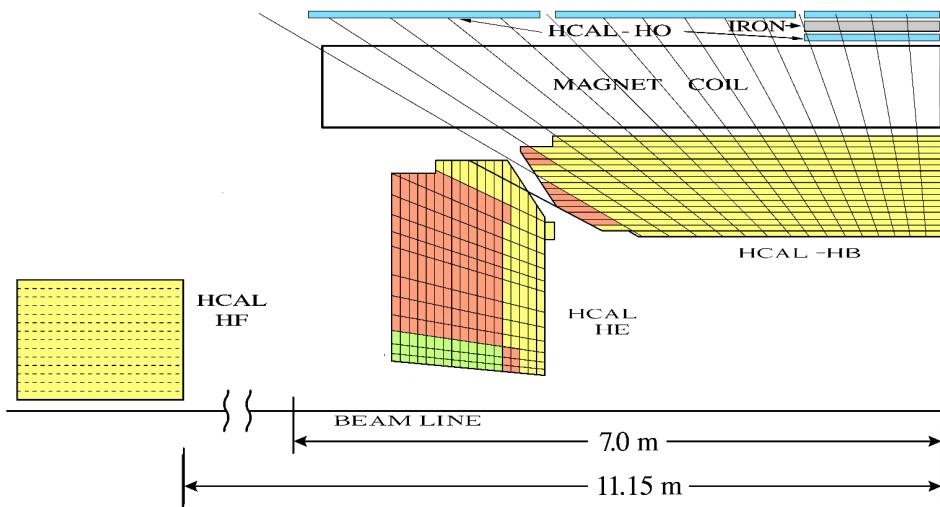


Figure 3.8: A quarter view of the hadronic calorimeter (HCAL), parallel to the beamline. The barrel (HB), endcap (HE), outer (HO), and forward (HF) detectors are indicated. Figure taken from [116].

The optical signals from the scintillators in the HB and HE are converted to electrical signals by multichannel hybrid photodiodes, while silicon photomultipliers (SiPMs) are used in the HO. In the HF, the Cherenkov light emitted in the quartz fibres is detected by standard photomultiplier tubes (PMTs), since the magnetic field is much smaller in this region.

The expected transverse energy resolution for jets reconstructed from calorimeter information is shown in Figure 3.9 for various pseudorapidity regions: barrel jets ($|\eta| < 1.4$), endcap jets ($1.4 < |\eta| < 3.0$), and very forward jets ($3.0 < |\eta| < 5.0$). Details about the reconstruction of jets from calorimeter and tracking information will be given in Section 4.3.7.

3.2.4 The muon system

The outermost detector, located entirely on the outside of the solenoid, is a dedicated muon detection system. The purpose of this subsystem is muon identification, momentum measurement,

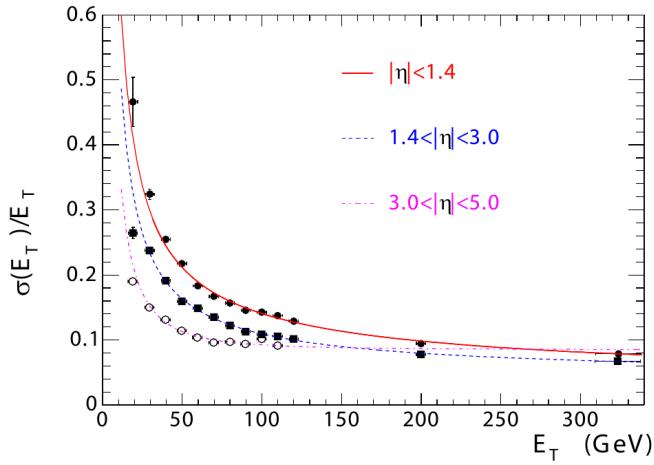


Figure 3.9: The jet transverse energy resolution as a function of the jet transverse energy, for barrel jets ($|\eta| < 1.4$), endcap jets ($1.4 < |\eta| < 3.0$), and very forward jets ($3.0 < |\eta| < 5.0$). Figure taken from [116].

and triggering. As illustrated in Figure 3.10, the layers of muon chambers are embedded in the iron yoke constraining the magnetic field lines. The strong magnetic field completely saturates the return yoke with a field of about 2 T, in opposite direction with respect to the field inside the magnet.

Three different types of gas-ionisation detectors are used. In the barrel, Drift Tubes (DTs) are installed, covering the pseudorapidity range up to $|\eta| < 1.2$, while Cathode Strip Chambers (CSCs) are used in the endcap region ($0.9 < |\eta| < 2.4$). The DTs are designed for the low muon rates that are expected in the barrel and thus have a slower response time than the CSCs, which are used at larger pseudorapidities due to the higher flux and the larger and non-uniform magnetic field. Resistive Plate Chambers (RPCs) complement the DT and CSC systems in the pseudorapidity region up to $|\eta| < 1.8$. They have a good time resolution but a worse spatial resolution than the DTs or CSCs, and provide a fast response which is much shorter than the 25 ns between 2 consecutive bunch crossings, making them very well suited to trigger on muons.

The offline muon reconstruction efficiency in simulated events containing one muon is typically between 95% and 99%, except for the regions between 2 DT wheels ($|\eta| = 0.25$ and $|\eta| = 0.8$) and the transition region between the DTs and CSCs ($|\eta| = 1.2$), where the efficiency drops to 92%. The reconstruction of muons using the information from the tracker and the muon detectors will be detailed in Section 4.3.4. For central muons with a small momentum, the offline momentum resolution of the standalone muon system is about 9%. At momenta around 1 TeV, the resolution varies from 15% to 40%, depending on the pseudorapidity. As demonstrated in Figure 3.11, performing a global momentum fit using the tracker as well improves the resolution by an order of magnitude at low muon momenta. At high momenta the resolution of the full system is about 5%.

3.2.5 Trigger and data acquisition

Collisions are provided by the LHC at high interaction rates, with an interval of 25 ns between bunch crossings. This corresponds to a frequency of 40 MHz. Additionally, multiple collisions occur at the same time, depending on the luminosity. Since it is impossible to store and process the large amount of data produced in the collisions at this high rate, a severe rate reduction is needed. This rate reduction is performed by the trigger system, which decides whether to store or reject an event. Since this decision process is constrained in time, the computing time is optimized by rejecting uninteresting events as quickly as possible. The rate is reduced to 1 kHz

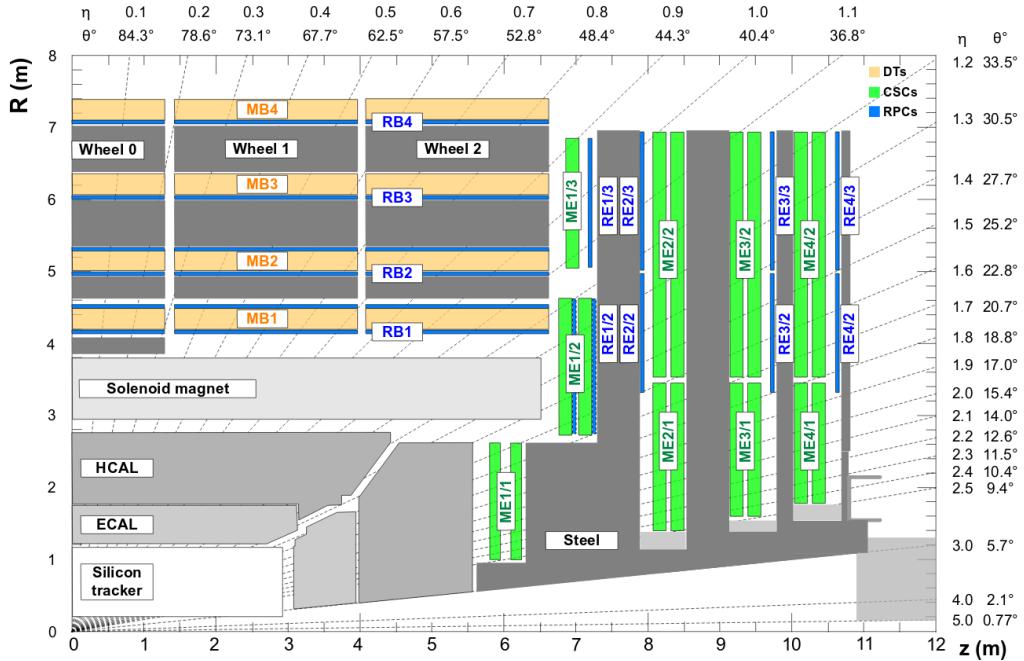


Figure 3.10: A transverse view of one quarter of CMS showing the position of the 3 types of muon detectors. The Drift Tube (DT) are located in the barrel, the Cathode Strip Chamber (CSC) in the endcaps, and the Resistive Plate Chamber (RPC) in both regions up to $|\eta| < 1.8$. Figure taken from [120].

in two steps by the Level-1 (L1) Trigger and the High-Level Trigger (HLT).

The L1 Trigger decision is based on information from the calorimeters and muon systems, following the structure illustrated in Figure 3.12. At the lowest level, the Local Triggers are based on energy deposits in calorimeter towers and track segments or hit patterns in the muon system. Regional triggers, indicated as Calo Trigger Layer 1 and Muon Track-Finder Layer in the figure, then combine this information and use pattern logic to determine trigger objects such as jet or muon candidates in separated spatial regions. The candidates are ranked based on their energy or momentum and quality, reflecting the level of confidence assigned to the L1 parameter measurements. Finally, the Calo Trigger Layer 2 and the Global Muon Trigger (GMT) determine the highest-rank calorimeter and muon objects across the whole detector and transfer them to the Global Trigger, which makes the final decision to accept or reject an event. Following this procedure, the L1 Trigger thresholds are tuned to reduce the event rate and keep it below 100 kHz. The L1 Trigger is composed of custom electronics located partially on the detectors, and partially in the underground service cavern. The L1 decision needs to be made and distributed to the detector front-end electronics within $3.8 \mu\text{s}$ [121].

The readout of the data proceeds as illustrated in Figure 3.13. When an event is accepted by the L1 Trigger, the data from about 740 FEDs is read out by the Readout Units (RUs). For so-called *legacy* systems, i.e. systems which are using VME-based hardware from the initial installation, the FEDs are read out by custom Front-End-Readout-Link (FRL) cards, while for systems that changed their readout architecture from the VME standard to the newer μ TCA standard during or after LS1 they are read out via the newer Front-End-Readout-Optical-Link (FEROL) cards. The event fragments are then sent over the event-builder switch to the Builder Units (BUs), which assemble the events. Next, the events are distributed to the Filter Units (FUs) by a large switch network.

The HLT software system is implemented in this filter farm, which uses more than 15 000 CPU cores for the final event selection. In this second step, the HLT reduces the event rate

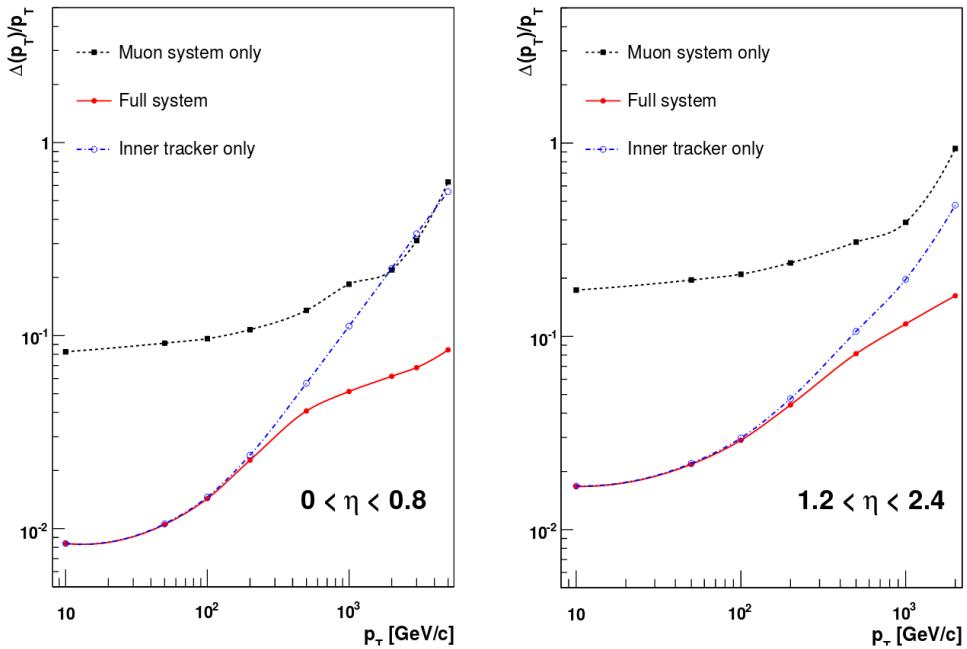


Figure 3.11: The muon transverse momentum resolution as a function of transverse momentum for low (left) and (high) pseudorapidities. The resolution is shown for the muon system and the tracker separately, and for the full system. Figure taken from [116].

further to 1 kHz. The complete read-out data, including information from the pixel and strip tracker, are available for this step. New objects can therefore be reconstructed such as e.g. tau leptons and b-jets, as is done in the offline software, but speed-optimized.

3.2.6 CMS performance in Run 2

The number of collisions recorded at the experiments will differ from the amount delivered by the LHC. Data loss can be caused by e.g. problems with a particular subdetector, data acquisition, the trigger rate, or the infrastructure. During Run 2, CMS achieved a data taking efficiency of 89% and 92% in 2015 and 2016, respectively. The comparison between the delivered and recorded cumulative integrated luminosity in 2016 is shown in Figure 3.14. Subsequently, the recorded data is certified by the offline Data Quality Monitoring (DQM), to ensure that the data are suited for physics analysis. The efficiency of this certification process with respect to the amount of recorded data was 80% in 2015 and 96% in 2016. This is shown in Figure 3.14 as well for 2016.

The detector was in good condition during data taking in 2015 and 2016. Figure 3.15 shows the active fraction of channels per subdetector, and the evolution from 2015 to 2016. In 2016, all subdetectors have an active fraction higher than 97.5%, except for the strip tracker, which has the largest number of channels and has an active fraction well above 95%. For this subdetector, channels that are not included or masked are accounted for, as well as channels that are identified as out-of-sync or irresponsible by the back end electronics, or as too noisy or inefficient from further analysis.

3.2.6.1 Pre-amplifier saturation in the APV25 chip

During Run 2, the instantaneous luminosity delivered by the LHC increased continuously, and even exceeded the design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ in 2016. As the luminosity increased, a dynamic inefficiency appeared in the strip tracker, which was most noticeable in the first layer

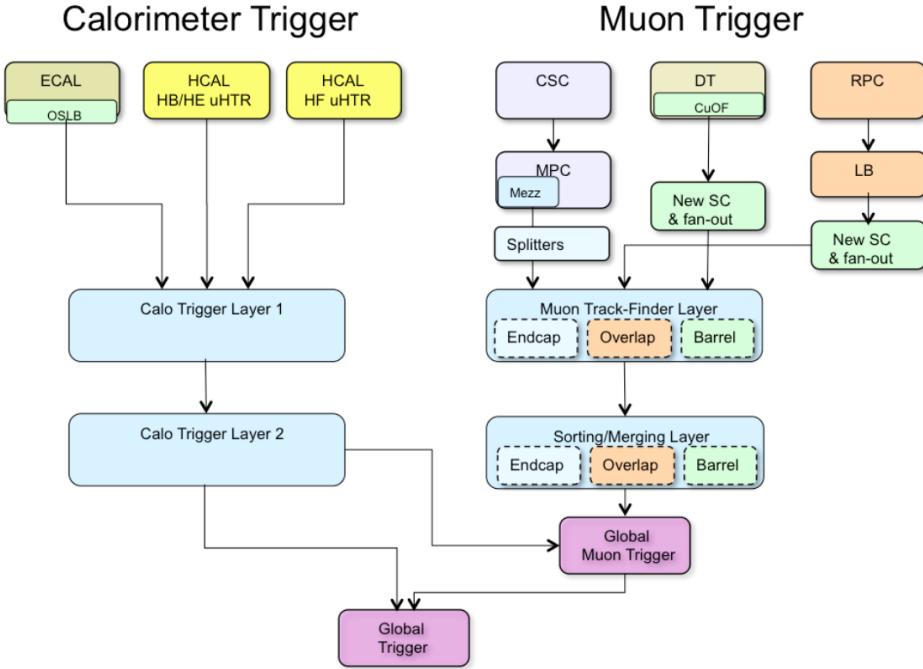


Figure 3.12: Schematic overview of the L1 Trigger. Figure taken from [121].

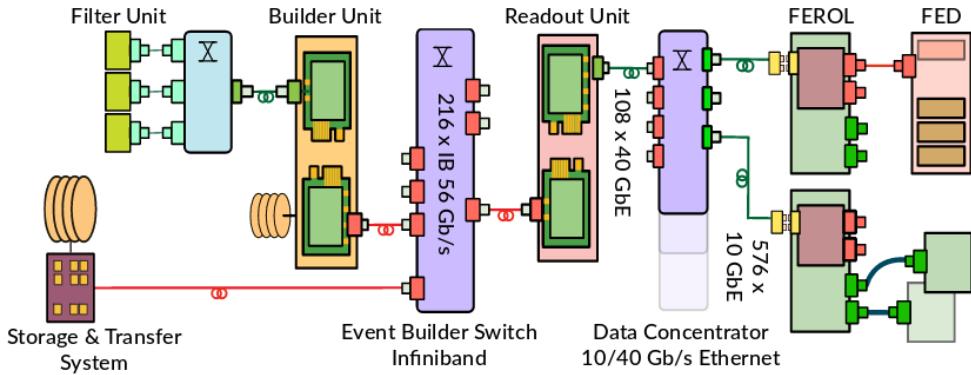


Figure 3.13: Schematic of the data acquisition (DAQ) system. Figure taken from [122].

of the TOB. The symptoms were a change in the signal-to-noise ratio and loss of hits. As can be seen from Figure 3.16, the most probable value (MPV) of the signal-to-noise ratio is shifted towards lower values and the low tail increased as well. The loss of hits is clearly visible in Figure 3.17, showing the change in number of hits per track for increasing instantaneous luminosities. The run periods indicated in the plot refer to a subset of the data taken over the course of the year. Run period boundaries are typically defined by changes in the LHC running conditions, changes to the detector configuration or calibration, or other parameters. The number of hits decreases for later run periods such as D and F, as the instantaneous luminosity increases. This loss of hits results in less and shorter tracks.

The origin of this inefficiency was eventually tracked down to saturation effects in the pre-amplifier of the APV25 chip. The pre-amplifier decay time changes significantly with temperature. As the operating temperature of the strip tracker was lowered from $+4^{\circ}\text{C}$ to -15°C coolant temperature during LS1, the decay time was no longer sufficient to cope with the high luminosities. The dynamic inefficiency was cured in August 2016 by changing the pre-amplifier drain

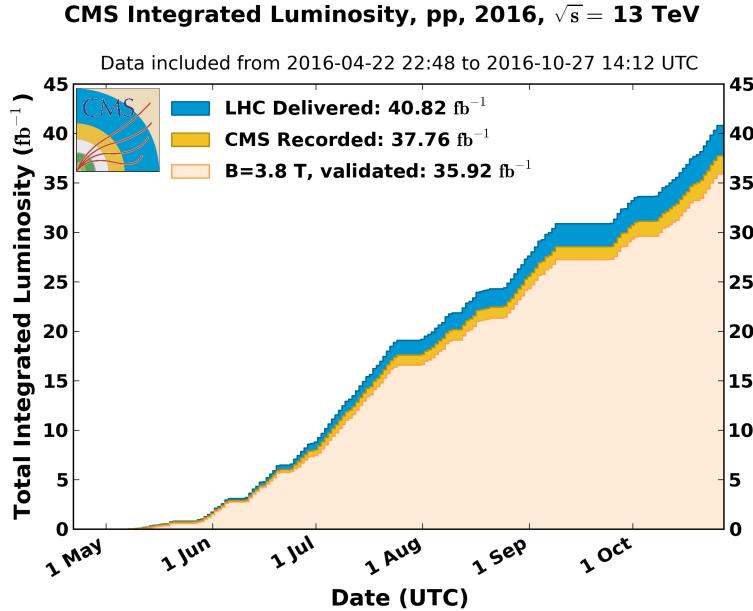


Figure 3.14: The cumulative distribution of the instantaneous luminosity delivered by the LHC (blue), recorded by CMS (yellow), and certified as good for physics analysis (light orange) in 2016. Figure taken from [123].

speed. This led among other things to the recovery of the muon efficiency, which showed a large drop for the highest luminosities before the change and an essentially flat behaviour afterwards, as demonstrated in Figure 3.18. This is a crucial effect, for example for the analysis described in Chapter 6, where the difference between the two drain speed settings is clearly visible.

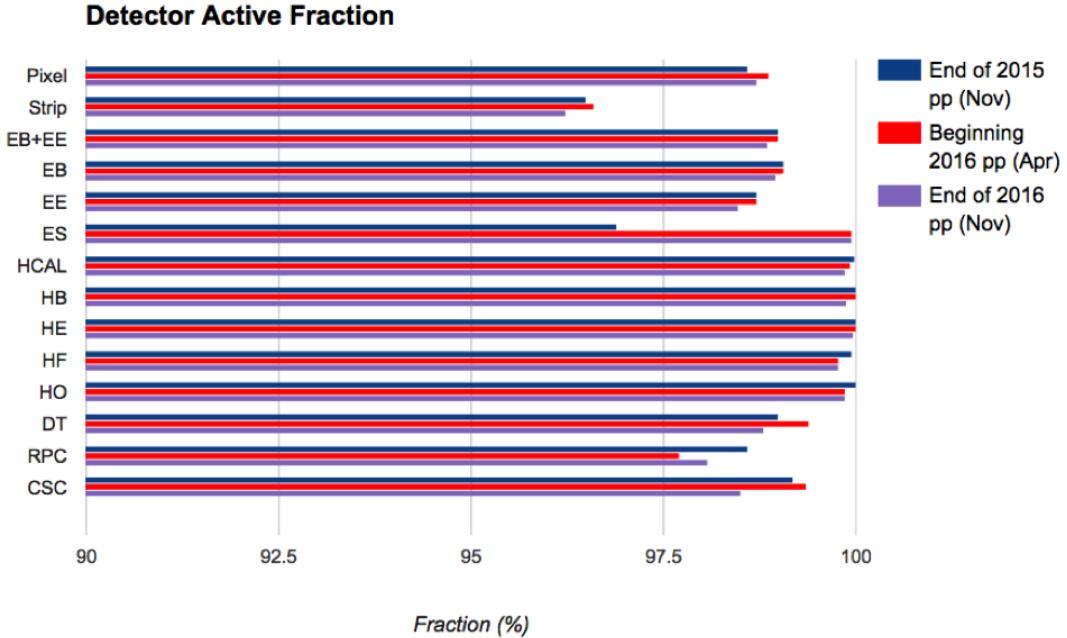


Figure 3.15: Fraction of active channels per subdetector, compared for the end of data taking in 2015, and the beginning and end of data taking in 2016. Figure taken from [124].

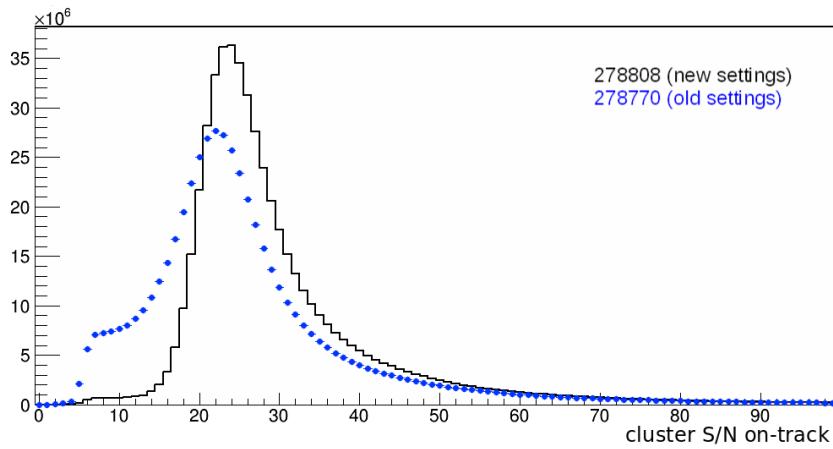


Figure 3.16: The signal-to-noise ratio for clusters on reconstructed tracks in the first layer of the TOB for two different runs in summer 2016. The first run (blue) is affected by the saturation effects in the APV25 pre-amplifier, resulting in a downward shift of the peak and an increase of the distribution at low values. In the second run (black), after the change of pre-amplifier drain speed, the distribution has its standard shape again. Figure taken from [125].

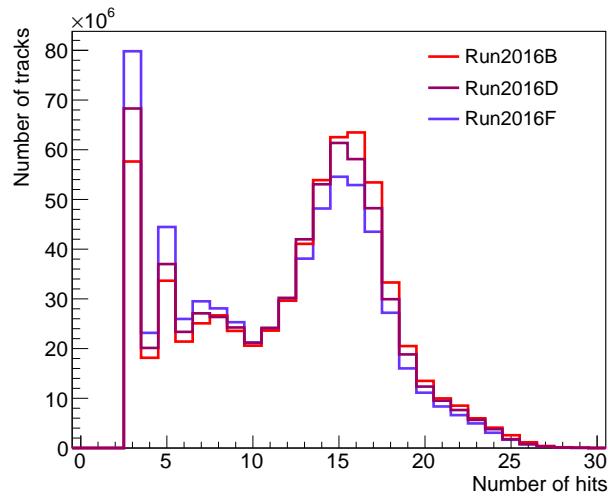


Figure 3.17: The number of hits per track for run periods B, D, and F, showing the effect of the increasing instantaneous luminosity.

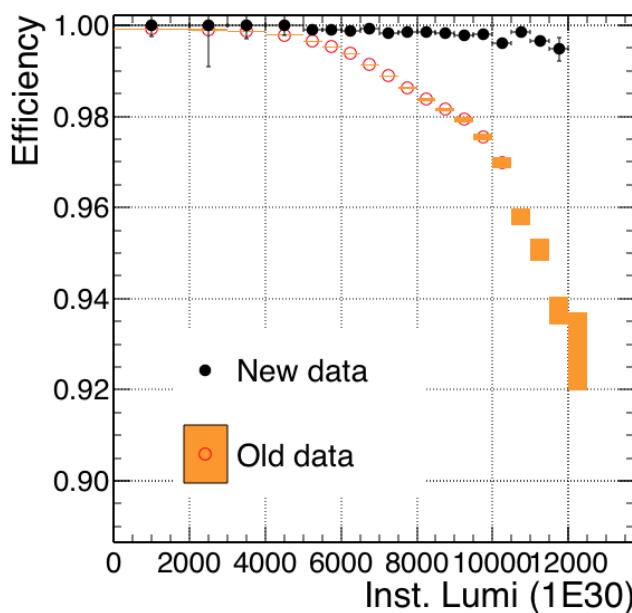


Figure 3.18: The muon efficiency as a function of the instantaneous luminosity for before (orange) and after (black) the change of pre-amplifier drain speed which cured the dynamic inefficiency. Figure taken from [125].

4

Event Simulation and Reconstruction

In order to use the recorded data, the obtained signals coming from various parts of the detector must be reconstructed to be able to identify the particles in the event. Additionally, to compare the experimental results with theory, events are generated and the resulting signals in the detector are simulated, as detailed in Sections 4.1 and 4.2, respectively. The event reconstruction is detailed in Section 4.3. Finally, some details about the simulation of SIMPs are given in Section 4.4.

4.1 Event generation

The event structure at the LHC is complicated by the composite nature of the colliding protons, as demonstrated in Figure 4.1. This sketch shows the hard interaction in red, with a tree-like structure surrounding it, representing the ensuing shower. In this hard scattering, the quark or gluon constituents of the protons, called partons, will interact according to a so-called parton distribution function (PDF), which is determined by the parton's momentum fraction and the momentum transfer. Due to their colour charge, the partons involved in the hard interaction will induce parton showers consisting of a cascade of radiation from QCD processes. This is shown in blue for the incoming partons and in red for the outgoing partons. The produced partons will also hadronize due to colour confinement, as illustrated in green, with hadron decays in dark green and radiated photons in yellow. Finally, the purple interaction represents a second interaction between the proton remnants. Next to these multiple parton interactions, additional activity in the event can also come from pileup. All these aspects must be taken into account when generating events, as detailed below.

Hard scattering

In the hard interaction, two partons of the colliding protons, will interact with a certain probability at a given momentum transfer. This is parametrized by the PDFs $f(x, Q^2)$, where x is the proton's momentum fraction and Q^2 is the momentum transfer scale. Experimentally determined PDFs are available from various groups, including e.g. CTEQ [127], MRST/MSTW [128], and NNPDF [129]. An example of such PDFs obtained by the NNPDF group is shown in Figure 4.2. This figure shows that the valence quarks in the proton, namely the up and down quarks it is made of, have a larger momentum fraction than the remaining sea quarks, which appear as virtual quark-antiquark pairs forming from gluons and annihilating again. The PDFs are then convoluted with the matrix element of the hard scattering, which is the process of interest where the two colliding partons create

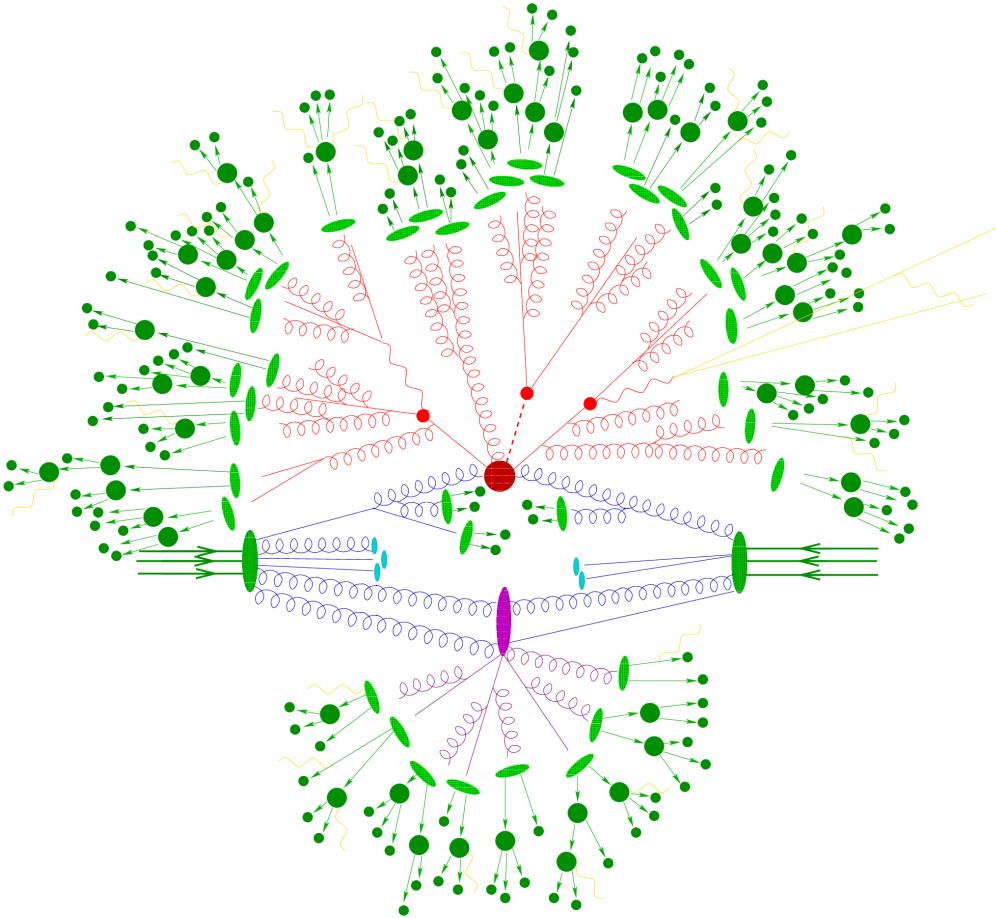


Figure 4.1: Illustration of an event showing the hard scattering, parton shower, hadronisation, and underlying event. Figure taken from [126].

high-energetic final state particles. This is done using an event generator, such as `MADGRAPH5_aMC@NLO` [130] and `POWHEG` [131]. With `MADGRAPH5_aMC@NLO` the matrix element can for many processes be calculated at tree-level or leading order (LO), and since the addition of `aMC@NLO` at next-to-leading order (NLO) as well. This generator was used to produce most of the background processes for the Monojet analysis detailed in Chapter 5 and for the SIMP signal used in Chapter 6. `POWHEG` is able to generate events using NLO computations, but only for a relatively limited number of physics processes. This generator was used to produce the monojet signal samples and the background processes from single-top production. Since NLO calculations are more time-consuming, one can instead use the less precise method of scaling a LO cross section to the NLO level by using a so-called k-factor, defined as the ratio of the NLO and LO cross sections. However, these k-factors often need to be determined as a function of the relevant kinematic variables as they depend on the kinematic phase space and the probed energy scale.

Parton showering

Since the colliding partons have a colour charge, the hard scattering will be accompanied by a cascade of radiation from QCD processes. The partons will for example radiate soft gluons or split into two collinear partons. This radiation can originate from the incoming partons, which is referred to as initial state radiation (ISR), or the outgoing partons in the final state, the so-called final state radiation (FSR). The perturbative evolution of the cascade can be modelled using the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi)

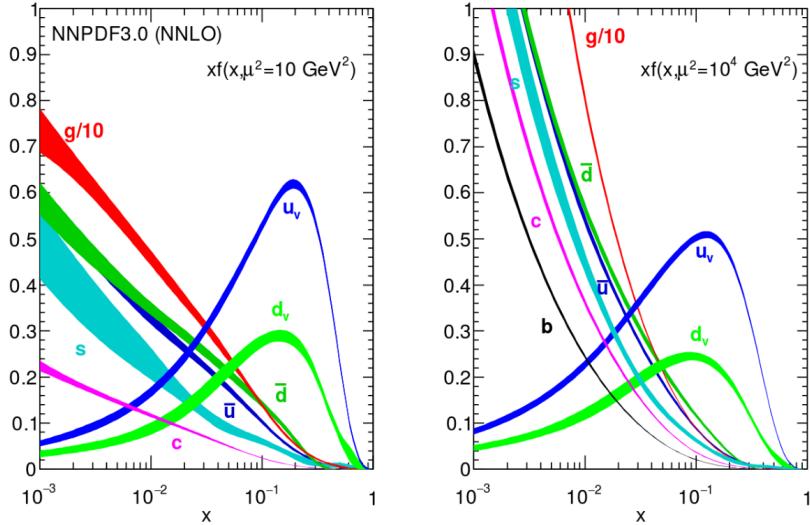


Figure 4.2: The parton distribution functions multiplied by the momentum fraction x at energy scales $Q^2 = 10 \text{ GeV}^2$ (left) and $Q^2 = 10000 \text{ GeV}^2$ (right), obtained in the NNLO NNPDF 3.0 global analysis. Figures taken from [129].

equations [132–134]. These equations describe the time evolution of the probability of a ‘mother’ parton to split into ‘daughter’ partons at an energy scale Q^2 . The momentum of the mother is then divided among the daughter partons, which can in turn split into other partons at a lower Q^2 scale. The cascade continues down to an energy scale Λ_{QCD} where the strong interaction becomes non-perturbative. The resulting number of jets can vary depending on the modelled process.

Hadronisation

The next step after the showering is the hadronisation of the coloured particles produced in the parton shower, transforming them into colour-neutral hadrons. Since this happens at low energy scales where the perturbative approach of QCD is not valid, phenomenological models have to be used. For most of the processes considered in this thesis, the showering and hadronisation is done with PYTHIA 8 [135], using a standard set of parameters which were tuned to reproduce the experimental data. In PYTHIA, the string Lund model [136] is used, based on string fragmentation. This model starts from the idea of a string connecting a quark q and an antiquark \bar{q} , following the assumption of linear confinement. As the two quarks move away from each other, the string stretches and the potential energy stored in the string increases. The increase in potential energy is assumed to be proportional to the distance between the quarks. When the energy becomes sufficient to produce a new pair of quarks $q'\bar{q}'$ with mass m , the string breaks and the original quark pair is split into two new pairs, $q\bar{q}'$ and $q'\bar{q}$. If the invariant mass of the new strings is large enough, the same process is repeated, leading to a new break-up. This procedure continues until only colour-neutral hadrons with an on-shell mass remain.

Additional activity in the event

In addition to ISR and FSR, also beam remnants and multiple parton interactions give rise to additional activity in the event, referred to as the underlying event. After the partons participating in the hard scattering are extracted, the remainder of the protons have a non-zero colour charge. The creation of additional hadrons during the hadronisation is therefore possible. Multiple parton interactions represent additional interactions which can take place between other incoming partons. As the probability for an additional hard interaction to occur is rather small, the activity from multiple parton interaction is typically much

less energetic than the hard interaction, producing mostly low energetic hadrons. Finally, additional collisions between other protons in the same bunch crossing, or from a previous or future bunch crossing, respectively referred to as in-time and out-of-time pileup, add extra activity in the event. The pileup distribution is for example shown in Figure 4.3 for QCD dijet events recorded in 2016, and is compared to simulated QCD events. This shows that there were about 20 collisions per bunch crossing on average. Typically, the simulation does not completely agree with the data and needs to be reweighted in order to match the data.

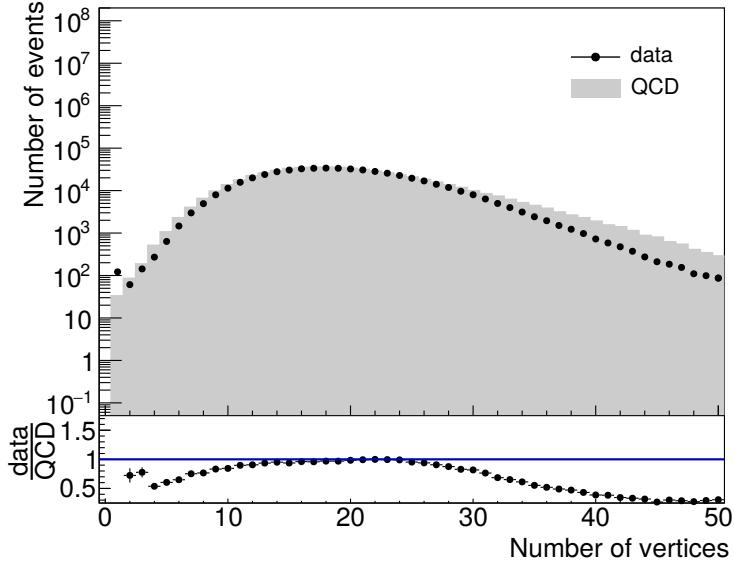


Figure 4.3: The distribution of the number of vertices for QCD dijet events recorded in 2016 compared to simulated QCD events.

4.1.1 Generation of the monojet signals

In the monojet analysis, the simplified models described in Section 2.2.4 are considered. The used signal samples were generated with POWHEG, which can generate NLO vector and axial-vector mediator production and LO scalar and pseudoscalar production. The samples were also produced at LO with MCFM [137] as a cross check. The scanned mediator masses are $m_\phi = 10, 20, 50, 100, 200, 300, 500, 1000, 2000, 10\,000 \text{ GeV}$, for dark matter masses of $m_\chi = 1, 10, 50, 100, 150, 500, 1000 \text{ GeV}$, with $m_\chi \leq m_\phi$.

4.1.2 Generation of the SIMP signal

For the generation of the SIMP signal, the model Lagrangian given in equation 2.22 is implemented in FEYNRULES 2.0 [138]. The matrix element is then calculated at LO and events are generated using MADGRAPH 5. The subsequent parton shower and hadronisation is done with PYTHIA 8, using tune CUEP8M1. Several samples were produced, with SIMP masses $m_\chi = 1, 10, 100, 200, 400, 700, 1000 \text{ GeV}$. The corresponding production cross sections are given in Table 4.1.

m_χ [GeV]	$\sigma_{\bar{\chi}\chi}$ [pb]
1	4.46
10	4.40
100	2.55
200	0.790
400	0.0743
700	0.00485
1000	0.000571

Table 4.1: Production cross section for each SIMP mass, after $|\eta_\chi| < 2.5$ and $p_T^\chi > 200$ GeV generator level cuts.

4.2 Detector simulation

After being generated, the collision events are passed on to the CMS detector simulation, which is based on the GEANT 4 [139] simulation toolkit. This toolkit provides a description of the interaction between particles and the detector material, including effects such as bremsstrahlung of charged particles, photon conversions, energy loss of charged particles by ionization, and the showering of electrons, photons and hadrons in the calorimeters due to interaction with the material. The CMS simulation package contains the geometry of the detector with all the sensitive layers designed to detect the traversing particles, as well as the dead material regions consisting of e.g. support structures, cables and cooling pipes. A precise map of the magnetic field is also included in order to simulate the curvature of the charged particles correctly.

Next, the impact of the detector, coming from the electronic response produced by the hits in the active detector material, the digitization, the data transmission, and any reconstruction performed in the electronics such as zero-suppression or cluster reconstruction, is simulated. In this way, an event content similar to the output of the real detector is obtained. At this point the effect of pileup is also included by adding detector hits of generated proton-proton interactions on top of the hits resulting from the main interaction. Most of the simulated event samples used in this thesis are processed using this detector simulation. However, the interaction of new particles that can arise from specific theory models is not always readily described in GEANT. This is the case for the signal samples used in the analysis described in Chapter 6, so an additional step was needed in order to simulate strongly interacting massive particles (SIMPs) in the CMS detector, described in Section 4.4.

4.3 Event reconstruction

Once the detector response has been simulated, the obtained events can be reconstructed. The same method is applied for these simulated events and for data coming from the detector. First, the reconstruction of tracks is performed, with a specific track reconstruction for electrons and muons. Furthermore, the calorimeter deposits, generated by electrons, photons, and hadrons, are grouped into clusters. Additionally, the reconstruction is further improved by using the so-called particle flow (PF) algorithm described in Section 4.3.6. This algorithm greatly improves the performance for jet and hadronic τ decay reconstruction, missing transverse momentum determination, as well as electron and muon identification. Finally, the obtained particle flow (PF) candidates are clustered into jets, and the missing transverse energy can be derived.

4.3.1 Track and vertex reconstruction

The tracks of charged particles going through the CMS tracker are reconstructed with an iterative tracking approach. This is used to cope with the high occupancy of hits and consequently high combinatorics. Additionally, the first iterations search for tracks with less possible combinations,

such as tracks with many pixel hits or a high momentum. After every iteration, the hits associated with the found track are removed to reduce the combinatorics. Each iteration consists of four steps:

1. **Seed generation.** In this first step hits are combined into seeds for the subsequent track finding. In the initial iterations pixel triplets are used, then pixel pairs together with a constraint on the origin of the trajectory based on the assumption that it originated near the beam spot, in order to take gaps or non-working modules into account. Next, mixed pixel/strip triplets are taken, and finally strip-only seeds are used. These additional iterations improve the acceptance in p_T and in displacement with respect to the primary vertex.
2. **Track finding.** The seeds are used as starting point for a Kalman filter algorithm. This method extrapolates the seed trajectory outward to the next layer, taking into account the presence of the magnetic field, potential energy loss and multiple scattering. If compatible hits are found in the next layer, the parameters of the trajectory are updated. This process continues until the outermost layer of the tracking system. Using this method, a given seed can generate multiple tracks, or different tracks can share hits. A trajectory cleaner therefore determines the fraction of hits the tracks have in common and discards the track with the lowest number of hits when there are too many shared hits. If both tracks have the same number of hits, the track with the largest χ^2 value is removed.
3. **Track fitting.** The track parameters are then refitted using a Kalman filter and smoother, taking all hits determined in the track finding step into account. This is done in order to find the optimal track parameters.
4. **Track selection.** Finally, the tracks are selected based on quality requirements, such as the number of layers that have hits, the χ^2/dof , and the distance to a primary vertex. This greatly reduces the fraction of reconstructed tracks that are fake.

The performance of the track reconstruction is excellent, and a high track-finding efficiency is obtained [140] while keeping the rate of fake tracks negligible. The highest tracking efficiency is obtained for muons, which traverse the full detector volume and have an improved momentum resolution at high p_T due to tracking information from the muon detectors giving a long lever arm. For isolated muons with p_T between 1 and 100 GeV the tracking efficiency is higher than 99% for the entire η coverage of the tracker, as can be seen from the left plot in Figure 4.4. The p_T resolution is about 2-3% for a muon with $p_T = 100$ GeV up to $|\eta| < 1.6$, but worsens for higher pseudorapidities. Different types of particles interact differently with the detector material. Charged hadrons, for example, are also subject to elastic and inelastic nuclear interactions and have a tracking efficiency of 80-95% depending on pseudorapidity and transverse momentum, as shown in the right plot of Figure 4.4.

Finally, the primary vertices are reconstructed from the tracks. Since the collisions happen between bunches of protons, at the luminosities reached at the LHC multiple protons will be colliding at the same time. The particles generated in these pileup interactions are all detected simultaneously and form a challenge to disentangle them from the particles coming from the to be studied interaction.

The reconstruction is performed as detailed in Section 9.4.1 of [141], by first clustering the tracks based on their z coordinate at the point of closest approach to the beam line. The vertex position is then estimated using an adaptive vertex fit [142] with as input the collection of tracks that is expected to come from the same interaction. Next, the tracks originating from the same vertex are clustered into jets with the anti- k_T algorithm described in Section 4.3.7 using a distance parameter of 0.4.

The vertex with the highest $\sum p_T^2$ is chosen as primary vertex, where the sum runs over these reconstructed jets, as well as the remaining tracks associated to the vertex and the missing transverse momentum at this vertex, in order to take into account neutral particles. While in events with jets many tens of high-momentum tracks can usually be associated to a primary

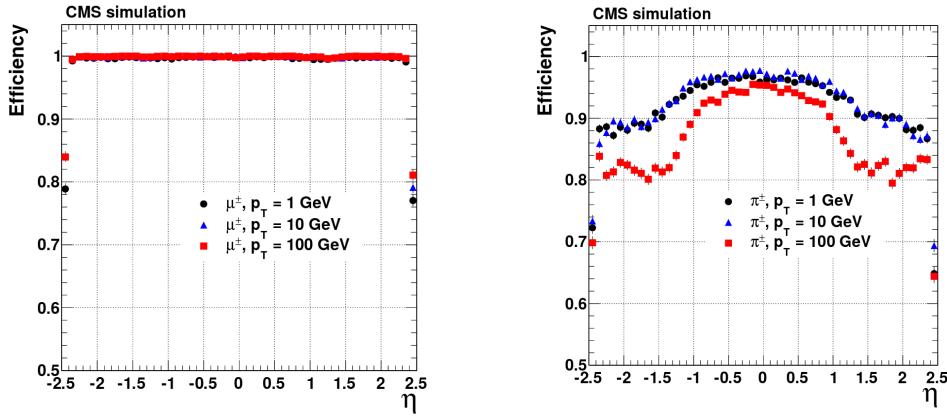


Figure 4.4: The muon efficiency (left) and pion efficiency (right) as a function of pseudorapidity, for multiple transverse momenta. Figures taken from [140]

vertex, thus making primary vertex finding almost fully efficient and pure, in the case of a pair of neutral jets, produced for example by SIMPs, this is not the case any more. The underlying event and potentially initial state QCD radiation can still provide some tracks, but in extreme cases a wrong vertex is chosen, arising from a hard pileup collision.

4.3.2 Electron and isolated photon reconstruction

Electrons are reconstructed using information from both the tracker and the calorimeters. Due to the large amount of material present in the tracker, electrons will emit bremsstrahlung photons, and photons will often convert into e^+e^- pairs, which can again radiate bremsstrahlung photons.

For electrons, a Gaussian-sum filter (GSF) [143] candidate is taken as starting point. This GSF candidate is obtained using 2 different methods to reconstruct the electron track from the hits in the tracker, which should gather all radiated energy from the electron. First, the ECAL-based approach is used, grouping ECAL clusters into superclusters. These superclusters collect the energy of the electron and the bremsstrahlung photons in a small η window and a large ϕ window, taking the bending of the electron track in the magnetic field into account. The supercluster energy and position is then used to estimate the position of the corresponding hits in the tracker layers. Subsequently, the tracker-based approach is used to find electrons missed by the ECAL-based method. In this case, the tracks from the iterative tracking with transverse momentum larger than 2 GeV are used. Additional requirements are placed on the number of hits and the χ^2 of the fit, and the specific electron tracking is performed, using a GSF fit, which is more adapted to electrons than the Kalman filter used in the iterative tracking, as it describes the energy loss in each tracker layer. The electron seeds obtained with both methods are merged and used as input for the full electron tracking. The obtained electron tracks are then linked to ECAL clusters by the PF algorithm, as described in Section 4.3.6. In the case of isolated photons, a candidate is seeded from an ECAL supercluster with transverse energy larger than 10 GeV which is not linked to a GSF track.

The total energy of the accumulated ECAL clusters is corrected for the energy that was lost in the process of reconstruction, using analytical functions of the energy and pseudorapidity. The applied corrections can be as large as 25%, at low transverse momentum and at $|\eta| = 1.5$, where the material density in the tracker is largest. The energy of the electron is then obtained from a combination of the corrected energy and the momentum of the GSF track, while the direction of the electron is taken from the GSF track. For photons, the corrected energy and the direction of the supercluster are used.

4.3.3 Electron and photon identification

In general, the electron and photon candidates must satisfy identification criteria to be retained. In the case of electrons two methods for identification are available: a cut-based identification or a multivariate boosted decision tree (BDT) combining fourteen variables including the amount of energy radiated and the ratio between the energies gathered in HCAL and ECAL. In the monojet analysis described in Chapter 5, the former is used. In this method, four different working points are defined, denoted as “tight”, “medium”, “loose”, and “veto”, with varying signal efficiency and background rejection. For the electron veto the loose selection is used, while a tight identification is required on one electron to select the events in the dielectron and single electron control regions.

variable	loose		tight	
	barrel	endcaps	barrel	endcaps
full 5x5 $\sigma_{i\eta i\eta}$	< 0.0114	< 0.0352	< 0.0101	< 0.0279
$ \Delta\eta_{in} $	< 0.0152	< 0.0113	< 0.00926	< 0.00724
$ \Delta\phi_{in} $	< 0.216	< 0.237	< 0.0336	< 0.0918
H/E	< 0.181	< 0.116	< 0.0597	< 0.0615
relative isolation	< 0.126	< 0.144	< 0.0354	< 0.0646
$ 1/E - 1/p [\text{GeV}^{-1}]$	< 0.207	< 0.174	< 0.012	< 0.00999
$ d_{xy}(\text{vtx}) [\text{cm}]$	< 0.0564	< 0.222	< 0.0111	< 0.0351
$ d_z(\text{vtx}) [\text{cm}]$	< 0.472	< 0.921	< 0.0466	< 0.417
expected inner missing hits	≤ 2	≤ 3	≤ 2	≤ 1
pass conversion veto	yes	yes	yes	yes

Table 4.2: Loose and tight electron identification criteria. The isolation is computed in a cone of $\Delta R < 0.3$ around the electron.

Similarly, for the photons, both a cut-based identification and a multivariate analysis can be used. Both for the monojet and the SIMP analysis, described in Chapters 5 and 6, the cut-based identification is used. Three standard working points are provided, denoted as “loose”, “medium”, and “tight”, with an average efficiency of 70%, 80%, and 90%, respectively. In both analyses, the loose identification is used for the photon veto. In the SIMP analysis, the event is only rejected when the identified photon is within a cone of $\Delta R < 0.1$ of one of the two leading jets. Additionally, the photon veto is extended to reject events containing jets with a large photon energy fraction and unidentified photons. This happens for instance when there is a photon conversion in the tracker. The event is therefore rejected when the jet photon energy fraction is larger than 0.8, the photon is not identified by the loose criteria, and a conversion is matched to the photon within $\Delta R < 0.2$ and has $p_{T,\text{conv}}/p_{T,\gamma} > 0.3$. Lastly, the two jets are also required to have a neutral electromagnetic energy fraction lower than 0.9, corresponding to one of the standard tight jet identification requirements mentioned in Section 4.3.7. Finally, the monojet analysis uses a photon + jets control region as well. These events are selected by applying the tight photon identification.

4.3.4 Muon reconstruction

Muon tracking is performed using 2 complementary approaches. The first method starts from standalone muons, which are reconstructed from hits in the muon detectors only using pattern recognition. The standalone muons are then matched to tracks in the tracker, and the hits are combined to form a global muon track. This global muon fit improves the momentum resolution compared to the tracker-only fit at muon momenta larger than 200 GeV.

For momenta below 10 GeV, muons often fail the global muon conditions which require the muon to penetrate through more than one muon detector plane, due to the large multiple

variable	loose		tight
	barrel	endcaps	barrel
full 5x5 $\sigma_{inj\eta}$	< 0.0102	< 0.0274	< 0.0102
H/E	< 0.05	< 0.05	< 0.05
charged hadron isolation	< 3.32	< 1.97	< 1.37
neutral hadron isolation	$< 1.92 + 0.014 \times p_T$ $+ 1.9 \times 10^{-5} \times p_T^2$	$< 11.86 + 0.0139 \times p_T$ $+ 2.5 \times 10^{-5} \times p_T^2$	$< 1.06 + 0.014 \times p_T$ $+ 1.9 \times 10^{-5} \times p_T^2$
photon isolation conversion safe	$< 0.81 + 0.0053 \times p_T$	$< 0.83 + 0.0034 \times p_T$	$< 0.28 + 0.0053 \times p_T$
electron veto	yes	yes	yes

Table 4.3: Loose and tight photon identification criteria. The isolation is computed in a cone of $\Delta R < 0.3$ around the photon.

scattering in the return yoke. In this case, tracker-only muon reconstruction is more efficient since it only requires one muon segment. Each track in the tracker with a transverse momentum larger than 0.5 GeV and a total momentum larger than 2.5 GeV is therefore extrapolated to the muon system and if at least one matching track segment is found, it is retained as muon candidate.

Within the geometrical acceptance of the muon system about 99% of the muons are reconstructed, either as global muon or as tracker muon and frequently as both. Global and tracker muons that share the same track inside the tracker are merged into a single candidate. Muons that are only reconstructed as standalone muons have a worse momentum resolution compared to the global and tracker muons. These standalone muons are however only considered in the further reconstruction when the fit is of high quality and is associated with a large number of hits in the muon system.

Charged hadrons can be misreconstructed as muons if e.g. a part of the hadron shower reaches the muon system. In order to improve the muon identification, the PF muon identification algorithm described in Section 4.3.6 also matches energy deposits in the ECAL and HCAL with the muon track.

4.3.5 Muon identification

When using muons for physics analysis, some identification criteria are generally applied in order to ensure the quality of the muons. There are several standard levels of identification, denoted as “tight”, “medium”, and “loose”, which provide a trade-off between muon identification efficiency and misidentification. In general, the tight and loose identification are the most widely used identification criteria.

The loose identification only requires the muons to be either global or tracker-only muons, and to be identified by the PF algorithm as a muon. As a result, it is highly efficient for both prompt muons and muons from quark decays. In analyses with prompt muons, this identification is therefore often complemented by an impact parameter cut, associating the muon to the primary vertex.

For the tight identification, the muon is required to be a global muon and to be identified by the PF algorithm as a muon. The normalized χ^2 of the global muon track fit should be smaller than 10 to suppress hadronic punch through and muons from hadron decays in flight. To further suppress these contributions at least one muon chamber hit should be included in the global muon track fit and muon segments should be found in at least two muon stations. Cosmic muons and tracks from pileup are suppressed by requiring the tracker track to have $|d_{xy}| < 2$ mm and $|d_z| < 5$ mm, with d_{xy} the traverse impact parameter and d_z the longitudinal distance with respect to the primary vertex. Finally, at least one pixel hit is required, as well as hits in at least 5 tracker layers, in order to ensure a good p_T measurement.

In the case of the monojet analysis described in Chapter 5, the loose muon identification is used to select muons for the muon veto. An additional isolation cut of 0.2 is applied in order to reject muons inside jets. The isolation value is computed as the sum of the transverse momenta of all charged hadrons associated to the primary vertex, neutral hadrons, and photons in a cone of $\Delta R < 0.4$ around the muon, relative to the p_T of the muon. The tight muon identification is used as well, to select events in the dimuon and single muon control regions.

4.3.6 Particle flow

The particle flow (PF) algorithm [144] reconstructs so-called particle flow candidates by combining information from all different CMS subdetectors, linking different elements, such as tracks in the tracker, calorimeter clusters, and muon tracks. A global picture of the event is thus formed, where each particle is uniquely identified. The obtained collection of particle candidates is subsequently used to reconstruct jets and tau leptons, and to determine the missing transverse momentum.

In a first step, the PF algorithm identifies charged particle tracks, as defined in Section 4.3.1 for all tracks, and in Sections 4.3.2 and 4.3.4 for electron and muon tracks in particular. At the same time, the calorimeter clusters are reconstructed with a clustering algorithm designed specifically for the PF event reconstruction. In this algorithm, cluster seeds are first identified as local energy maxima with respect to the four or eight closest cells, if the energy deposited in the cell is above a given seed threshold. The clusters are then formed by accumulating neighbouring cells with an energy above a given cell threshold, suppressing noise.

The PF elements in the different subdetectors are then connected by a link algorithm which avoids any double counting. The link algorithm produces blocks of associated elements, quantifying the quality of the link by defining a geometrical distance between the elements. When an element is linked to multiple other elements, only the link with the shortest distance is kept. More precisely, a link between a track in the tracker and a calorimeter cluster is made by extrapolating it from the last hit in the tracker to the calorimeters. The distance between the position of the extrapolated track and the cluster in the (η, ϕ) plane is then used to define the link distance. At the interaction points between the track and the tracker layers, tangents to the GSF tracks are extrapolated to the ECAL in order to collect the energy of photons radiated by electron bremsstrahlung. A dedicated conversion finder was also developed to identify bremsstrahlung and prompt photon conversions into e^+e^- pairs. Links between calorimeter clusters are established outside of the tracker acceptance, or between the preshower and ECAL clusters in the preshower acceptance. In this case the link distance is also defined as the distance between the position of the clusters. Charged particle tracks can also be linked by a common secondary vertex. Finally, the PF muon identification algorithm associates the muon tracks to the muon energy deposits in the ECAL and HCAL, to improve the muon identification performance.

In a next step, the PF blocks obtained by linking the multiple PF elements, are classified as muons, electrons, or isolated photons. The corresponding elements are then excluded from further consideration. Once electrons, muons, and isolated photons have been identified, the remaining elements are identified as charged hadrons, neutral hadrons, or photons produced in jets. Within the tracker acceptance, the ECAL clusters not linked to any track are classified as photons, while the clusters in the HCAL without a matched track are labelled as neutral hadrons. Outside of the tracker acceptance, charged and neutral hadrons can not be distinguished. ECAL clusters linked to an HCAL cluster are then assumed to arise from the same hadron shower, and the estimated energy for these particles is the sum of the energy deposited in the ECAL and the HCAL. The ECAL clusters that are not linked to an HCAL cluster are classified as photons.

4.3.7 Jet reconstruction

Jets are reconstructed with the anti- k_T algorithm [145], which clusters either the generated particles from event simulation, or the particles reconstructed by the PF algorithm (PF jets), or the

energy deposits in the calorimeters (Calo jets). This procedure takes into account the transverse momentum p_T , also called k_T , of the particles and the distance between particles, defined as

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \quad (4.1)$$

The strategy consists of the following steps:

1. For every pair of particles i and j , a distance d_{ij} defined as

$$d_{ij} = \min \left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2} \right) \frac{\Delta R_{ij}^2}{R^2} \quad (4.2)$$

is calculated.

2. For every particle i , a distance d_{iB} to the beam pipe is calculated with

$$d_{iB} = 1/p_{Ti}^2. \quad (4.3)$$

3. The minimum of d_{ij} and d_{iB} is then determined.

4. If it is d_{ij} , particles i and j are recombined into a new particle by adding the four-momenta of the particles. If it is d_{iB} , particle i is declared to be a jet and it is removed from the list of particles.

5. This is repeated until no particles remain.

In this clustering algorithm, the parameter R determines what is called a jet. If a particle i has no other particles within a distance R , d_{iB} will be smaller than d_{ij} and the particle will become a jet. A consequence of this is that an arbitrarily soft particle can become a jet, and therefore a minimum transverse momentum for a jet to be of interest is defined.

The anti- k_T algorithm favours clustering around hard particles, and the jets then grow outward from this seed. This gives rise to circular jets, with a cone size that is proportional to R . Since it still involves a combination of energy and angle in the distance measure, this is a collinear-safe growth, meaning that the jet will not change when one of the particles of the jet is split collinearly. This algorithm is also infrared-safe, i.e. the same set of jets is obtained when soft particles are emitted.

A reliable determination of the jet energy is not straightforward, since many effects can distort the energy estimation, such as the calorimeter response, the limited particle reconstruction efficiency, the underlying event, the pileup, and the charged particles bending out of the jet cone due to the strong magnetic field. The pileup is mitigated by applying charged hadron subtraction (CHS), which consists of removing charged hadrons associated with vertices other than the primary vertex from the list of PF candidates. Additionally, the jet energy is corrected using a factorised approach, as illustrated in Figure 4.5, with the following steps:

- **Pileup correction (L1).** The first level of jet energy corrections is applied event-by-event and jet-by-jet, and is determined from simulation. It is dependent on the pseudorapidity and transverse momentum of the jet, the average p_T density in the event, and the effective jet area. This effective area is determined by injecting a large number of very soft particles in the event before the jet clustering. The spread of the soft particles in each jet then defines the jet area. When these corrections are applied on data, residual corrections are also applied to take into account the difference between the simulated events and the data.
- **Relative η and absolute p_T corrections (L2L3).** These corrections are also obtained from simulations and correct for the non-uniform response of the calorimeters in η and p_T . They are determined by comparing the reconstructed p_T to the one obtained from the jets built from the generated particles. These corrections have the largest impact on the jet energy.

- **Residual η and p_T corrections (L2L3Residual).** Since the L2L3 correction is derived from simulation, additional residual corrections are needed in order to correct for the remaining small differences between the jet response in data and simulation. These corrections are typically of the order of a few percent.

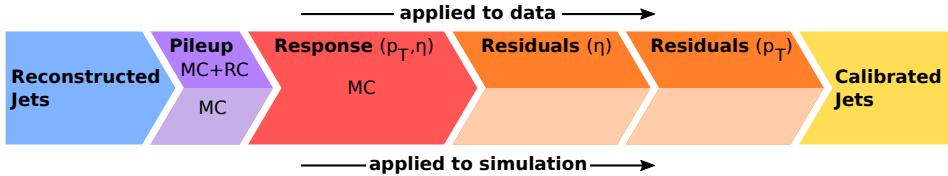


Figure 4.5: Graphical overview of the factorised approach used at CMS to apply jet energy corrections.

Optionally, a set of identification criteria are applied on the PF jets. A jet is required to consist of at least two particles. For jets in the region $|\eta| < 2.7$, the fraction of energy coming from either neutral hadrons or photons should not exceed 99%. Additionally, for jets restricted to the tracker acceptance ($|\eta| < 2.4$), there should at least be some energy deposited in the HCAL, the jet should contain 1 or more charged constituent, and the fraction of energy corresponding to electrons or photons should not exceed 99%.

Moreover, jets from pileup can be identified as well. This pileup jet identification relies on the topology of the jet shape which is used to disentangle jets coming from the overlap of multiple interaction and real hard jets, the object multiplicity, and the compatibility of the tracks in the jets with the primary vertex. This last property of pileup jets can evidently only be exploited for jets within the tracker acceptance.

Both jet and pileup jet identification are used in the monojet analysis described in Chapter 5. In the trackless jets analysis described in Chapter 6 however, the full jet ID is not applied, since the requirements on e.g. the neutral hadronic energy fraction and the charged multiplicity would reject the signal events. Instead, only the requirement on the jet neutral electromagnetic energy fraction is applied, in order to reject photons.

4.3.8 Identification of b-jets

For the identification of jets originating from b quarks, the long lifetime of the B hadrons arising from the hadronisation of b quarks is exploited. The B hadrons will therefore decay at a position that is displaced with respect to the primary interaction vertex. The b-jets can then be identified by looking for the presence of displaced tracks from which a secondary vertex may be reconstructed. Additionally, the B hadrons have a probability of 20% to decay into muons or electrons. Consequently, the presence of these charged leptons can be used as well for b-jet identification techniques.

Within the CMS collaboration, two different algorithms are being used during Run 2, namely the Jet Probability and the Combined Secondary Vertex taggers [146]. In this thesis, the latter is used to identify b-jets, combining the information from displaced tracks and secondary vertices in a multivariate technique. Jets are then identified or “tagged” as b-jets by applying a cut on the discriminator output. Three standard operating points are defined, denoted as “loose”, “medium”, and “tight”, corresponding to a misidentification probability of 10%, 1%, and 0.1% for light jets with $p_T > 30$ GeV, respectively. In the monojet analysis described in Chapter 5, the loose working point is used to veto b-jets.

4.3.9 Reconstruction of tau leptons

Tau leptons can decay into either a charged lepton and two neutrinos, or a few hadrons and one neutrino. The hadronic decays of the tau lepton can be separated from quark or gluon jets by

analysing the decay products. With the PF algorithm, it is possible to resolve the particles originating from the tau decay and to determine its isolation. Hadronic tau decays are reconstructed using the hadrons-plus-strips (HPS) algorithm [147] by using those particles as input. The jet constituent particles are combined into candidates compatible with one of the main hadronic tau decay modes, $\tau^- \rightarrow h^-\nu_\tau$, $\tau^- \rightarrow h^-\pi^0\nu_\tau$, $\tau^- \rightarrow h^-\pi^0\pi^0\nu_\tau$, and $\tau^- \rightarrow h^-h^+h^-\nu_\tau$, where h represents a charged hadron. This PF reconstruction of the tau decay products has significantly improved the reconstruction and identification of the tau leptons compared to the previous method which only took the energy deposits in the calorimeters into account.

4.3.10 Missing transverse momentum reconstruction

While most particles produced in the collisions can be reconstructed from the hits and energy deposits in the detector, some collision products might not leave energy deposits in tracker, calorimeters or muon system. This makes an accurate reconstruction of this type of particles impossible, and an alternative method is used, based on indirect observations. As the detector is hermetically closed such that all other particles in the event can be detected, the missing transverse momentum can be determined. This momentum then corresponds to all undetected particles in the event, and can be calculated from the vectorial sum of the transverse momenta of all the observed final state particles:

$$\vec{E}_T^{miss} = - \sum \vec{p}_T, \quad (4.4)$$

where the sum runs over all reconstructed PF particles. The variable that is generally used in particle physics analyses is the norm of the missing transverse momentum:

$$E_T^{miss} = |\vec{E}_T^{miss}|. \quad (4.5)$$

A notable example of particles leaving no hits or energy deposits behind are neutrinos, as they are neutral and weakly interacting and will therefore traverse the entire detector unhindered. Other hypothetical neutral weakly interacting particles, which are being searched for in many physics analyses, would escape the detector without producing hits as well.

4.4 Simulation and reconstruction of the SIMP signal

The SIMP events generated as described in Section 4.1.2 are then simulated in the CMS detector using GEANT. However, the SIMPs are not included in the simulation, as these new particles are unknown in GEANT and their interaction with matter has not been implemented yet. In order to simulate the new dark matter candidates in the CMS detector two new approaches were implemented.

In the first approach, the SIMPs were incorporated by adding an additional step to the standard reconstruction described in Section 4.3. In this additional step the SIMPs are directly converted to neutral PF candidates and merged with the rest of the PF candidates. Additionally, the new PF candidates are smeared with jet energy resolution (JER) distributions obtained from a sample produced using neutrons instead of SIMPs. Neutrons were chosen because of their resemblance to the SIMPs as single neutral particles generating a hadronic shower.

In order to produce this neutron sample used to derive the JER, the same additional custom step is applied, but in this case the neutrons will also be correctly reconstructed by the standard reconstruction. As a result, the neutrons will be present in the standard reconstruction and will be added a second time converting them into neutral PF candidates. The standard reconstructed PF candidates that are matched to the generated neutrons are therefore removed before injecting the converted generated neutrons to the collection of PF candidates. The JER distributions are then derived by comparing the resulting uncorrected PF jets with the corresponding neutrons in sample produced with the standard reconstruction using the full GEANT simulation. The

resolution is computed in bins of η and p_T , and an example is shown in Figure 4.6 for central neutrons with low and high transverse momentum.

After applying this smearing, the jets are processed with the standard sequence of CHS, jet clustering, L1FastJet, and L2/L3 corrections described in Section 4.3.7.

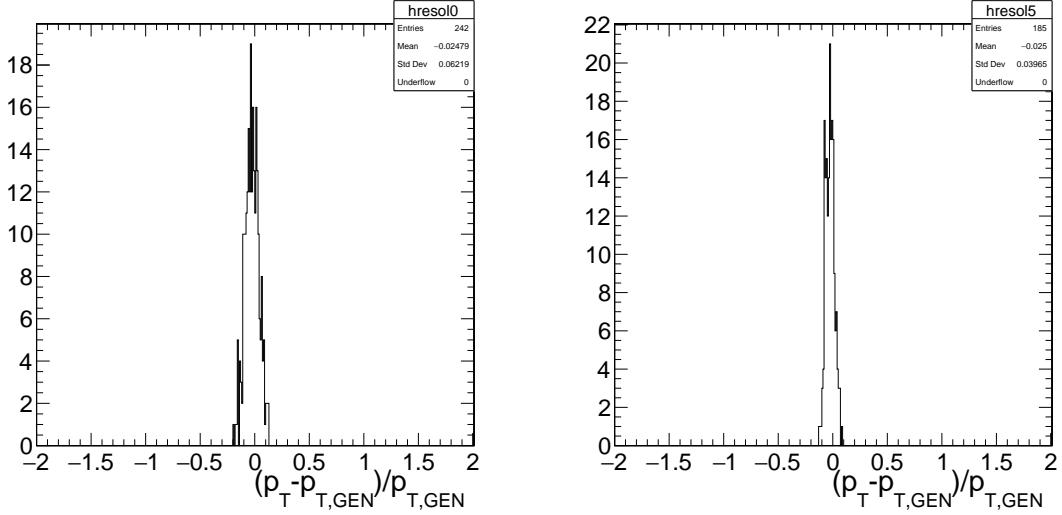


Figure 4.6: The jet energy resolution of neutrons with $0 < |\eta| < 0.5$ and $200 \text{ GeV} < p_T < 300 \text{ GeV}$ (left) or $700 \text{ GeV} < p_T < 800 \text{ GeV}$ (right).

In order to validate this method, the custom and standard neutron samples are used to compare the two leading generator-level jets to the new jets from the custom sample, denoted as P2PF jets, and the PF jets from the standard sample. This is illustrated in Figure 4.7, where the p_T of the generated neutrons is shown on the horizontal axis, and the p_T of the reconstructed jet is shown on the vertical axis. The left plot shows the standard neutron sample produced with the full GEANT simulation, while the right plot shows the custom neutron sample where the neutron was directly converted into a neutral PF candidate. The JER distributions are also compared in Figure 4.8 and fitted with a Crystal Ball function, showing compatible parameters.

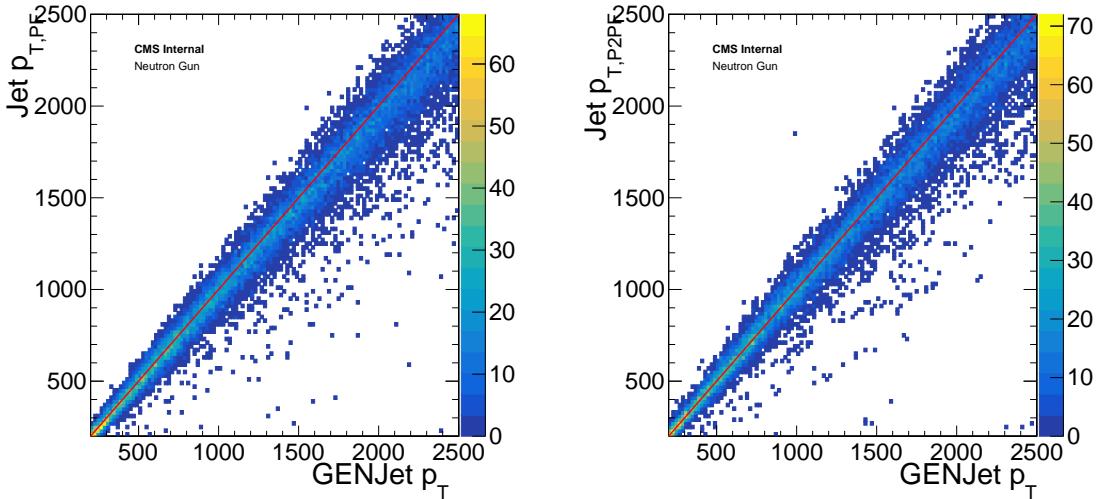


Figure 4.7: Comparison of the transverse momentum of the generator-level jets to the PF jets (left) and P2PF jets (right) in the region $0 < |\eta| < 0.5$ with jet energy resolution smearing, using a neutron sample.

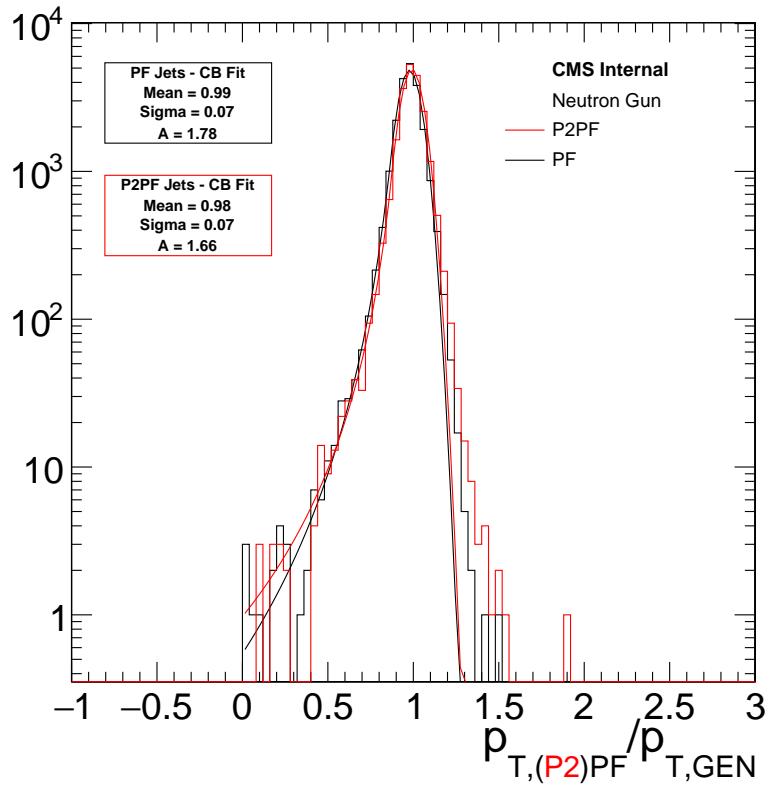


Figure 4.8: The jet energy resolution of the corrected P2PF jets (red) and PF jets (black), fitted with a Crystal Ball function.

This demonstrates that the procedure, where the JER distributions derived from a neutron sample are used to smear the PF candidates from generator-level SIMPs, can sufficiently accurately simulate SIMPs in a realistic detector, assuming SIMPs are neutron-like. However, since this procedure directly converts the generated SIMPs into PF candidates, the SIMPs do not interact in the Tracker and the resulting jets have a very steeply falling charged hadron energy fraction (CHF) distribution. This gives an optimistic image, which translates in a maximal signal efficiency.

The SIMP signal simulation and reconstruction was therefore further improved by moving to the second approach. In this method, the generated SIMP particles are not converted into neutral PF candidates, but they are instead replaced by neutrons, keeping the SIMP kinematics. The standard reconstruction and full GEANT simulation can then be applied, since the neutrons are correctly recognized and simulated. In this case, interactions will happen inside the Tracker as well and the resulting jets will contain a larger CHF, as is shown in Figure 4.9. Eventually, this second method will be used in the trackless jets analysis described in Chapter 6.

This method gives a good approximation of a SIMP signal, since the shower generated by the SIMP is in principle contained inside the calorimeters, as the model described in Section 2.3 is constructed so that for a specific choice of couplings the SIMPs may be detected as regular hadrons. Although the considered SIMP-nucleon interaction is repulsive, this does not differ considerably from known attractive interactions at the probed high energies. The incoming SIMP hits a nucleon at rest in the calorimeter, breaking it up, and because of the large incoming momentum, there is a boost forward into the calorimeter and the shower starts. The cross section would therefore be identical for a repulsive or attractive interaction and the effect on the shower is negligible since the scattering angle is very small due to the momentum boost. Furthermore, the higher the momentum of the SIMP, the shorter the distance it travels to deposit its characteristic momentum. With the considered couplings, the depth containing a SIMP with 500 GeV

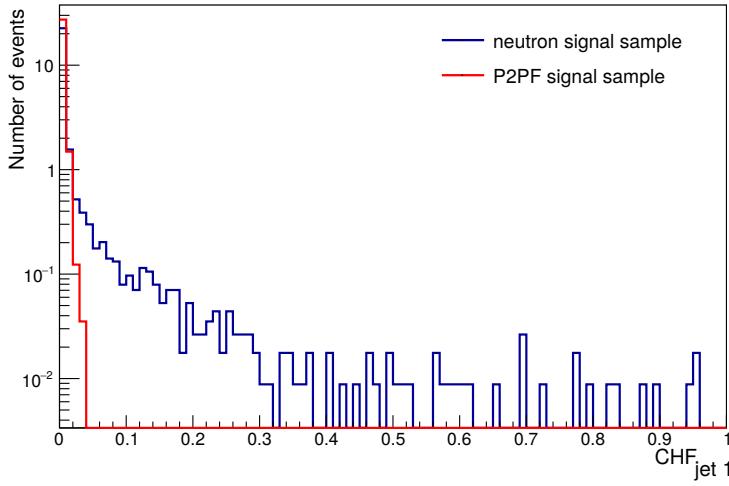


Figure 4.9: CHF distribution of the leading jet, for a signal sample produced with the first approach (red) and the corresponding sample produced using the second method (blue).

momentum is below 1 m, within the calorimeter [148]. Most of the energy will therefore be deposited in the first interaction with the material. Given the expected forward energy flow in the calorimeter shower, and the shower containment achieved by the choice of couplings in the simplified model, the shower induced by the SIMP interaction can to first order be modelled by the interaction of a high-momentum neutral hadron, like a neutron.

5

The Monojet Analysis

As has been described in Chapter 2, there are many searches for dark matter both at particle accelerators and elsewhere. At the LHC, one very promising channel is the so-called monojet search, where the detection of dark matter is done by looking for missing momentum in association with one or more jets. The dark matter particles are expected to pass through the detector without leaving any signal since they are neutral and only interact very weakly. They can however be detected indirectly as missing momentum when they recoil off one or more jets coming from initial state radiation. While the Run 2 analysis is described here, this analysis was already performed using the data collected during Run 1 [149] as well.

First, the used physics objects are described in Section 5.1. Next, the selection of the events using the trigger is described in Section 5.2, and the subsequent event selection performed with fully reconstructed events is detailed in Section 5.3. The estimation of the background and the included systematic uncertainties are described Section 5.4 and 5.5, respectively. In Section 5.6, the obtained results are shown. The improvements achieved by going from the analysis strategy used in 2015 to the 2016 version are detailed in Section 5.7. Finally, the results are interpreted in terms of the considered dark matter models in Section 5.9.

5.1 Physics object reconstruction

While jets and missing transverse momentum are evidently important objects in this analysis, other physics objects are used as well. Leptons and photons are for example vetoed in order to reject backgrounds, and are used to define the many control regions used for the background prediction.

5.1.1 Jets

The method described in Section 4.3.7 is used to reconstruct the used PF jets, and the jet energy corrections are applied as well. Furthermore, the jets are required to have $p_T > 30 \text{ GeV}$, and to pass the loose jet identification and pileup jet identification described in Section 4.3.7. Additionally, a jet cleaning is applied by requiring a jet charged hadron energy fraction CHF > 0.1 and a jet neutral hadron energy fraction NHF < 0.8 . The b- jets are tagged as using the Combined Secondary Vertex algorithm described in Section 4.3.8, with the loose working point.

5.1.2 Missing transverse momentum and hadronic recoil

The missing transverse momentum E_T^{miss} is reconstructed as detailed in Section 4.3.10. For the control regions, the E_T^{miss} is redefined in order to imitate the E_T^{miss} shape in the signal region. This hadronic recoil U is obtained by removing the leptons or the photon present in the event from the E_T^{miss} computation.

5.1.3 Leptons

Loose identification criteria, described in Section 4.3.3, are used to veto electrons. Additionally, the electrons are required to have $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$. Similarly, muons are required to pass the loose muon identification described in Section 4.3.5 and to have $p_T > 10 \text{ GeV}$ and $|\eta| < 2.4$. In the case of tau leptons, they are required to have $p_T > 15 \text{ GeV}$ and $|\eta| < 2.3$. They should also pass the tau identification criteria, which require a jet with an identified subset of particles with a mass consistent with the decay products of a hadronic tau and which are isolated with a pileup corrected isolation cut requiring less than 5 GeV of energy deposits within a radial cone of $\Delta R < 0.3$.

The selection of electrons and muons for the control regions is however stricter: they are required to pass the tight identification.

5.1.4 Photons

Photons are required to have $p_T > 15 \text{ GeV}$, $|\eta| < 2.5$, and to pass the loose identification criteria described in Section 4.3.3, in order to be considered for the photon veto. For the photon + jets control region, photons are required to pass the tight photon identification in order to be considered.

5.2 Trigger selection

In order to select events that have the monojet signature displayed in Figure 5.1, the trigger requires either $E_T^{miss} > 90 \text{ GeV}$, where E_T^{miss} is the magnitude of the negative vectorial sum of the p_T of all particles at trigger level, or $H_T^{miss} > 90 \text{ GeV}$, where H_T^{miss} is calculated as the magnitude of the negative vectorial sum of the momenta of all jets with $p_T > 20 \text{ GeV}$. In order to avoid collecting events that contain noise signals coming from the detector, tight requirements are placed on the jets used in the H_T^{miss} computation. Jet energy corrections are already applied at the HLT-level and the jet NHF is required to be smaller than 0.9. Muons are not taken into account to compute E_T^{miss} and H_T^{miss} , so that the same trigger can be used to select the events for the muon control regions used for the background prediction.

This trigger is studied using events passing the single electron trigger with a threshold of $p_T > 23 \text{ GeV}$, and applying an extra offline selection requiring an electron with $p_T > 40 \text{ GeV}$ passing the tight identification and a jet with $p_T > 100 \text{ GeV}$ and $|\eta| < 2.5$. The trigger efficiency is computed by determining the fraction of events which additionally pass the signal triggers, as a function of the hadronic recoil. The efficiency is above 98% for events passing the analysis selection described in Section 5.3, as shown in Figure 5.2.

To select events in the electron control regions, a single electron trigger was used with a threshold at $p_T > 27 \text{ GeV}$. The efficiency of this trigger is determined as a function of p_T and η . This is done by “tagging” one electron with $p_T > 40 \text{ GeV}$ and $|\eta| < 2.1$ which passes the tight selection requirements. A second electron with $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$ is then selected, while the invariant mass of this pair is required to be between 60 and 120 GeV, to correspond to the Z boson mass. This tag-and-probe method ensures that backgrounds are removed and allows to measure the efficiency by simply taking the fraction of probes that passes the single electron trigger. The obtained trigger efficiency turn-on curves as a function of the electron

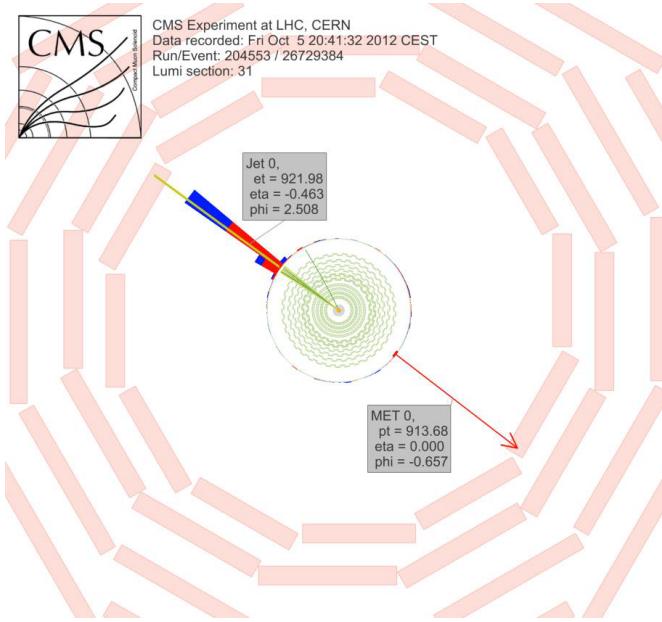


Figure 5.1: Event display showing the monojet final state.

transverse momentum are shown in Figure 5.3 for two η bins, covering the ECAL barrel and endcaps separately and leaving out the poorly instrumented region in between.

Finally, the events in the photon control region are selected with a single photon trigger requiring an isolated photon with $p_T > 175$ GeV. The performance of this trigger is measured in data selected using a single photon trigger with a lower p_T threshold, and the turn-on is shown in Figure 5.4.

5.3 Event selection

Once the events have been fully reconstructed and the jets in the event have been corrected as described in Section 4.3.7, an event selection is applied in order to efficiently select signal events and reduce the contribution from backgrounds, such as the production of a leptonically decaying W boson in association with jets, semileptonic diboson decays, QCD multijets, and top quark production. The background contribution coming from events with a Z boson produced together with a number of jets, where the Z boson decays to two neutrinos, is irreducible as it produces exactly the same signature as the expected signal. The missing transverse momentum E_T^{miss} is required to be larger than 200 GeV in order to be safely above the trigger turn-on. Additionally, the leading jet is required to have $p_T > 100$ GeV in order to balance the missing transverse momentum, and $|\eta| < 2.5$. A cut on the difference in azimuthal angle between the E_T^{miss} and the first four leading jets of $\Delta\phi(jet, E_T^{miss}) > 0.5$ is also applied. This is done to suppress the QCD background from mismeasurements of jet momentum or detector noise, which could introduce missing transverse momentum in the event, in the same direction as the mismeasured jet. The events are further cleaned by applying quality filters to remove events coming from beam or instrumental backgrounds. Finally, events containing a lepton, a photon, or a b-jet are vetoed as well. Figure 5.5 shows the E_T^{miss} distribution for data and MC, after applying the described selection.

Events containing a lepton are vetoed to suppress the electroweak backgrounds, such as $W(l\nu) + \text{jets}$ and semileptonic diboson decays, while the photon veto is added to suppress the $Z(\nu\nu) + \text{photon} + \text{jets}$ and $W(l\nu) + \text{photon} + \text{jets}$ background processes, and to ensure there is no overlap with a similar dark matter search which investigates the final state consisting of missing momentum and a photon. This rejects less than 1% of the signal. Finally, the b-jet veto

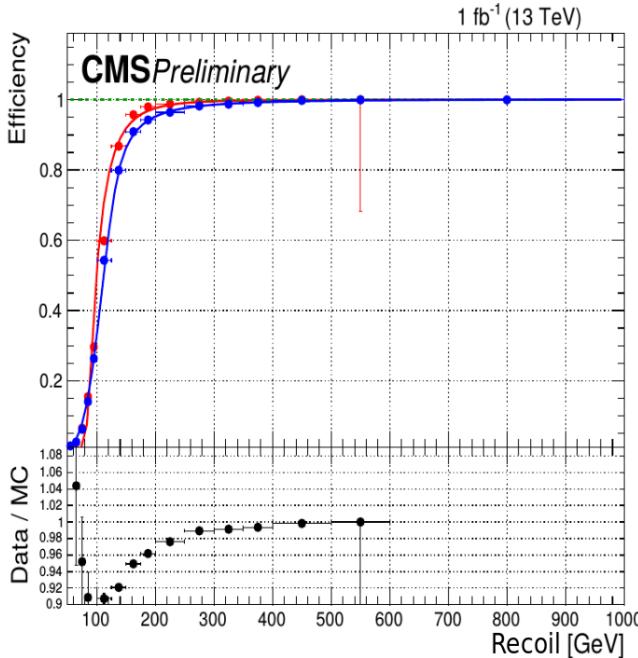


Figure 5.2: The efficiency of the used signal triggers as a function of the hadronic recoil, in MC (red) and data (blue).

reduces the background from events with a single top quark or top quark pairs by a factor 3 and only reduces the signal by 5 to 10%, depending on the type and mass of the mediator. No additional veto on the number of jets is applied.

5.4 Background estimation

The dominant background comes from events with a Z boson produced together with a number of jets, where the Z boson decays to two neutrinos. This produces the same signature of jets with missing momentum as the signal, and results in an irreducible background. The second largest background consists of $W + \text{jets}$ events with a leptonically decaying W boson. This background is already suppressed by the lepton veto, but a fraction of these events remain when the lepton is either not identified or outside of the detector acceptance. The remaining background events come from top quark decays, which are suppressed by the b-jet veto, semileptonic diboson (WW , WZ , and ZZ) decays, and QCD multijet events. The two main background contributions are estimated from five control regions in data consisting of dimuon, dielectron, single muon, single electron, and photon + jets events. The contributions from top quark decays and semileptonic diboson decays are estimated using simulated samples, while the QCD multijet background is estimated using a data-driven approach.

5.4.1 The Z and W background estimation

The traditional control region for the Z boson background is the dimuon control region. This region is dominated by $Z \rightarrow \mu\mu$ events, which are very similar to the $Z(\nu\nu) + \text{jets}$ background events, the only difference being the decay mode. The production mode and kinematics in the control region are very similar, as well as the acceptance. However, the branching ratio of the Z boson into two muons is 6 times smaller than the branching ratio to two neutrinos. As a result of this, and due to the muon selection efficiency, the dimuon control region contains about 10 times less Z boson events than the signal region. In order to improve this statistical limitation,

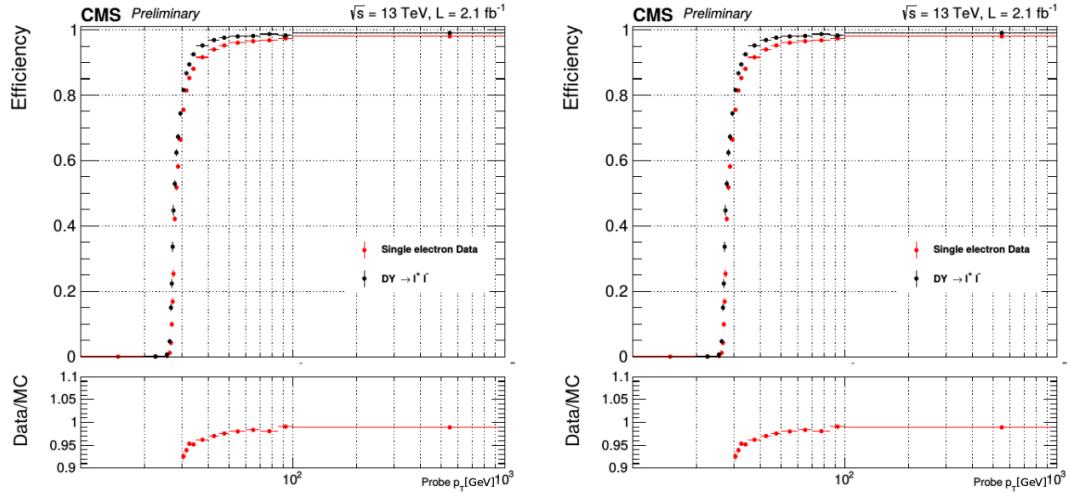


Figure 5.3: The efficiency of the single electron trigger in data (red) and MC (black) for $|\eta| < 1.4442$ (left) and $1.566 < |\eta| < 2.5$ (right).

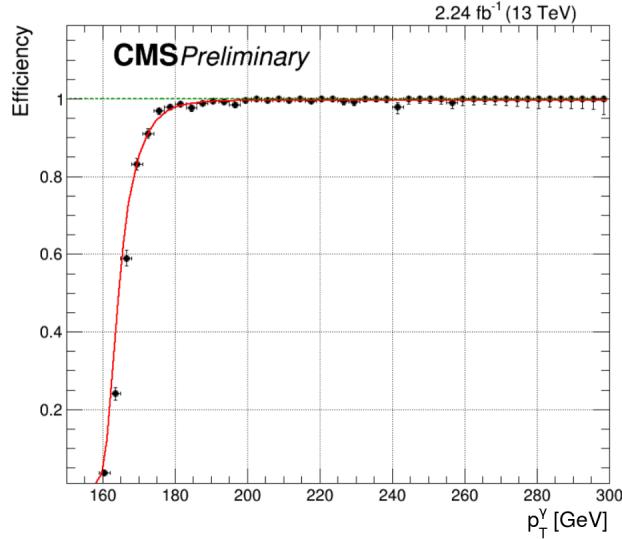


Figure 5.4: The efficiency of the single photon trigger measured in data.

other control regions have been added as well.

The yield of $Z(\nu\nu)$ and $W(l\nu) + \text{jets}$ events in the signal region is therefore estimated from five control regions by using the ratio between data and MC in the control region, per bin of the hadronic recoil distribution. As illustrated in Figure 5.6, the $W(l\nu) + \text{jets}$ background is estimated from the single lepton control regions, and the $Z(\nu\nu) + \text{jets}$ background is predicted from the dilepton and photon + jets control regions. Five transfer factors are therefore needed to connect the control regions to the signal region. An additional transfer factor is used between the $Z(\nu\nu) + \text{jets}$ and $W(l\nu) + \text{jets}$ backgrounds to factor in the experimental as well as theoretical correlations between these two backgrounds.

For the prediction using $Z \rightarrow \mu\mu$ events in the dimuon control region for example, the

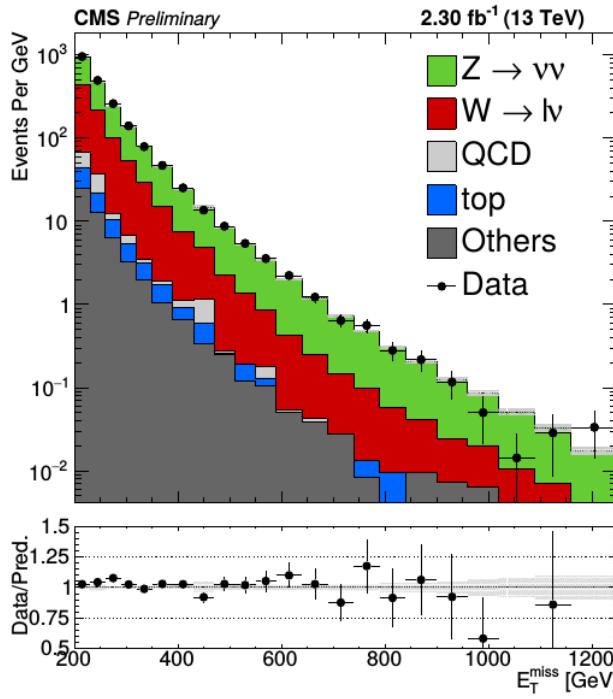


Figure 5.5: The missing transverse momentum distribution after applying the described event selection, for data and MC.

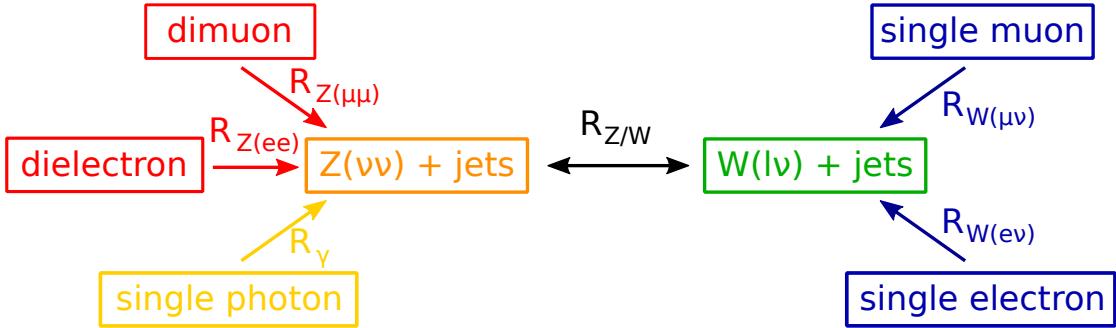


Figure 5.6: Schematic overview of the five control regions used to estimate the $Z(\nu\nu)$ and $W(l\nu) + \text{jets}$ backgrounds with the transfer factors.

predicted yield of $Z \rightarrow \nu\nu$ events is given by

$$N_{Z(\nu\nu)} = \frac{N_{Z(\mu\mu)}^{\text{data}}}{N_{Z(\mu\mu)}^{\text{MC}}} N_{Z(\nu\nu)}^{\text{MC}} \quad (5.1)$$

$$= \frac{N_{\mu\mu}^{\text{data}} - N_{\text{Bkgd}}}{N_{Z(\mu\mu)}^{\text{MC}}} N_{Z(\nu\nu)}^{\text{MC}} \quad (5.2)$$

$$= \left(N_{\mu\mu}^{\text{data}} - N_{\text{Bkgd}} \right) R_{Z(\mu\mu)}, \quad (5.3)$$

where the number of $Z(\mu\mu) + \text{jets}$ events in data $N_{Z(\mu\mu)}^{\text{data}}$ is given by the number events in the dimuon sample, removing the number of background events, and $N_{Z(\mu\mu/\nu\nu)}^{\text{MC}}$ represent the number of $Z(\mu\mu/\nu\nu) + \text{jets}$ events in MC. The five transfer factors, denoted by R , are derived

from simulation and take into account the impact of lepton acceptance and efficiency, as well as the additional E_T^{miss} requirement for the single electron control region. They also include the difference in branching ratio and the relation between the differential cross sections of the photon, W , and Z boson production as a function of the boson p_T . The transfer factors are computed as a function of the hadronic recoil, and are shown in Figure 5.7. Furthermore, the Z/W ratio shown in the bottom right plot of Figure 5.7 provides an additional constraint between the $Z(\nu\nu) + \text{jets}$ and $W(l\nu) + \text{jets}$ backgrounds.

The simulated samples used for the background estimation are generated at leading order (LO) using the MADGRAPH generator, and corrected to next-to-leading order (NLO). These corrections are crucial in order correctly represent the data since inclusively, the simulation is approximately 40% higher than the data when using only LO calculations. The NLO QCD k-factors are derived from samples generated at NLO with MADGRAPH5_aMC@NLO, while the NLO electroweak k-factors are obtained from theoretical calculations [150–153]. The differential cross section as a function of the boson p_T is shown in Figure 5.8 for photon, W , and Z production, and the obtained k-factors are displayed in the ratio plots. More details on the different control regions are given in the following.

Dimuon control region

In the dimuon control region the events are selected using the monojet triggers and applying the same requirements as described in Section 5.3 for the signal region, using the hadronic recoil instead of the missing transverse momentum, except for the muon veto. Additionally, exactly two muons with opposite charge should be identified using the loose identification, and at least one should also pass the tight selection requirements. The leading muon should have a transverse momentum larger than 20 GeV, and the second one should have $p_T > 10$ GeV. Finally, the dimuon mass should be between 60 and 120 GeV, corresponding to the Z boson mass.

Single muon control region

In order to model the second largest background, coming from $W(l\nu) + \text{jets}$ events, a single muon control region is used. This control region is in addition also used to constrain the $Z(\nu\nu) + \text{jets}$ background. The events in the single muon control region are required to pass the monojet triggers and event selection replacing the E_T^{miss} by the hadronic recoil obtained by removing the muon, except for the muon veto. One muon should then pass the tight selection requirements and have $p_T > 20$ GeV.

Dielectron control region

The dielectron control region is also used to constrain the $Z \rightarrow \nu\nu$ background. The events are selected by the single electron triggers. Similarly to the dimuon control region, the events are required to pass the monojet selection, except for the electron veto. Instead, exactly two electrons with $p_T > 10$ GeV are required to pass the loose identification described in Section 4.3.3. In addition, at least one electron should pass the tight selection requirements, and the leading electron is required to have $p_T > 40$ GeV in order to have a high probability to have passed the single electron trigger. Finally, the dielectron mass should be between 60 and 120 GeV, in order to be consistent with a Z boson decay. The jump that can be observed in the last bin of the resulting transfer factor in the middle left plot of Figure 5.7 is due to the isolation requirement of the single electron trigger. This was verified by removing the trigger selection in MC, yielding a flat transfer factor.

Single electron control region

The single electron control region is used to constrain the $W(l\nu) + \text{jets}$ background. In the single electron control region, the events are required to have one electron with $p_T > 40$ GeV passing the tight selection requirements, analogous to the single muon control region. In this region, a large amount of QCD background is however present due to jets being wrongly reconstructed as electrons. In order to reject most of those events, an

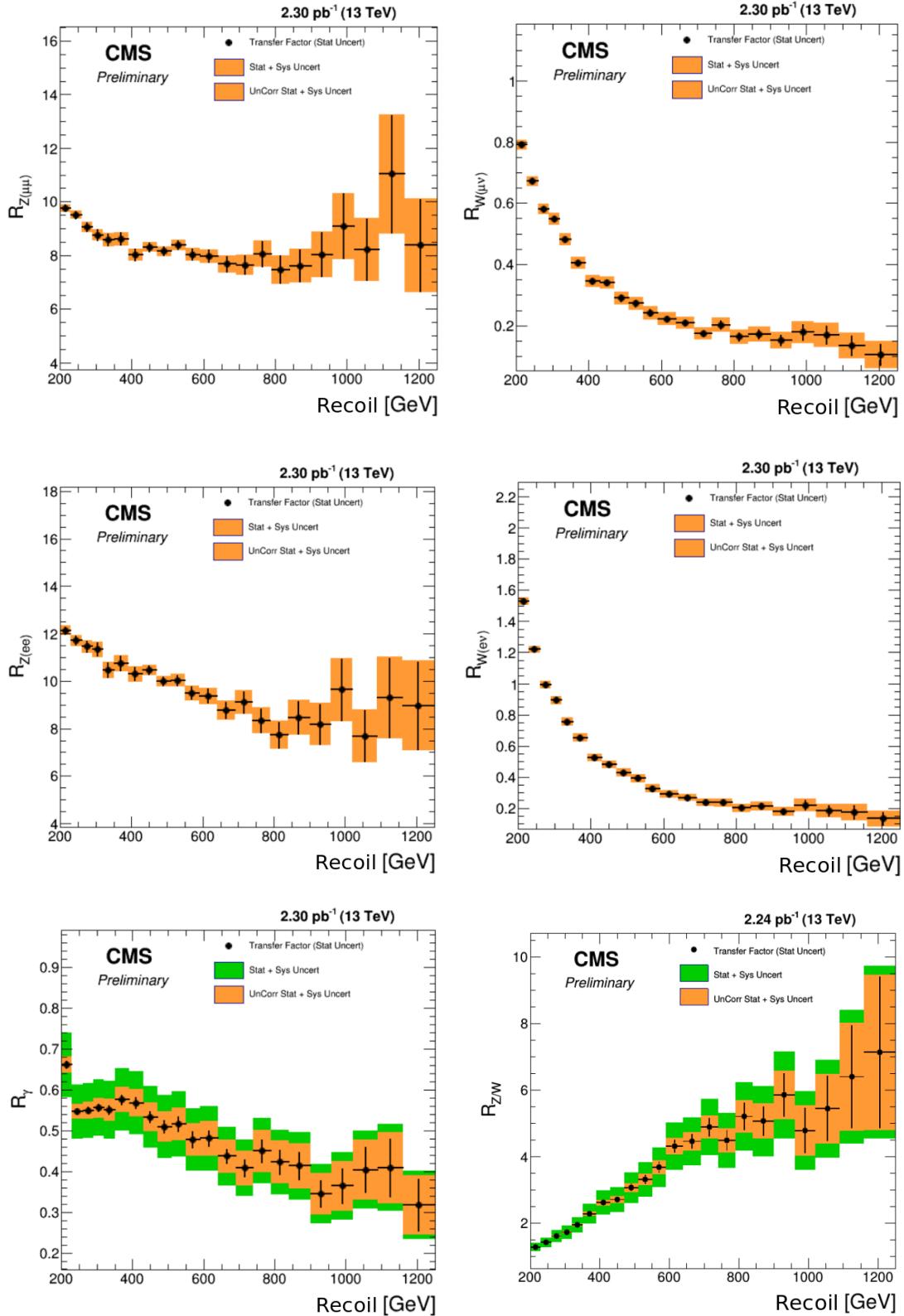


Figure 5.7: Transfer factors for the dimuon (top left), single muon (top right), dielectron (middle left), single electron (middle right), and photon + jets (bottom left) control regions. The ratio of the Z and W transfer factors is shown in the bottom right plot.

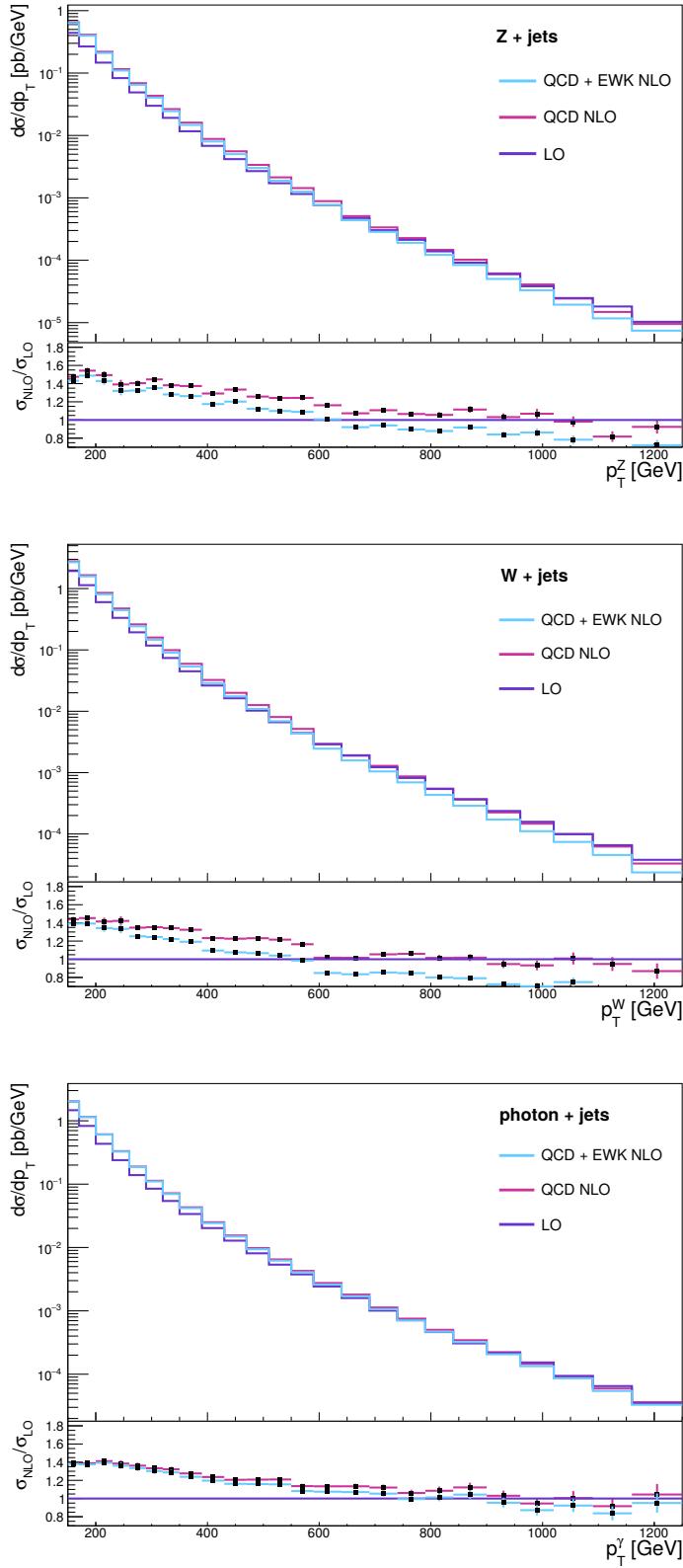


Figure 5.8: The differential cross section as a function of the boson p_T for photon, W , and Z production, using boson p_T -binned NLO samples. The resulting k-factors are shown in the ratio plots.

additional cut on the E_T^{miss} , which includes the single electron, is added at 50 GeV. This reduces the QCD background by an order of magnitude.

Photon + jets control region

Due to its large yield, the photon + jets control region provides the dominant constraint on the high- p_T part of the $Z(\nu\nu) + \text{jets}$ background. The selection of these events is done using the single photon triggers and applying the monojet selection, except for the photon veto. One photon is then required to pass the tight identification and to have $p_T > 175$ GeV. Additionally, it should be reconstructed inside the ECAL barrel ($|\eta| < 1.4442$) in order to achieve a high purity of at least 95%. Events with more than one photon passing the loose identification requirements described in Section 4.3.3 are rejected.

5.4.2 The QCD background estimation

While QCD multijet background events are generally well balanced in the transverse plane, missing transverse momentum can arise in the event due to jet energy mismeasurements, punch-through¹, uninstrumented or defective regions in the detector, hot spots, or neutrinos from decays of heavy-flavour mesons. Although these effects will very rarely yield a large missing transverse momentum, the QCD production cross section is very large compared to other processes, and some events can be selected at high missing transverse momentum. The event selection detailed in Section 5.3 was designed to suppress contributions from the QCD multijet background, reducing it to the percent level. However, this background is not well reproduced in the simulation, and thus requires a background estimation using a data control region.

The yield is predicted using the $\Delta\phi$ extrapolation method, the QCD background is estimated in data selected using the signal triggers and applying the events selection, but inverting the $\Delta\phi$ selection cut and instead requiring $\min \Delta\phi(\text{jet}, E_T^{miss}) < 0.5$. An extrapolation is then made to the signal region using transfer factors derived from simulated events generated at LO with MADGRAPH in several bins of H_T . The obtained estimation using this method yields a contribution that is a factor 2 larger than the prediction from simulation, as can be seen in Figure 5.9.

5.4.3 Simulation-based background estimation

Contributions are also expected from diboson production, from top quark decays, both from $t\bar{t}$ and single top production, and from $Z(l\bar{l}) + \text{jets}$ events where the leptons are not detected. These backgrounds are estimated from MC simulations.

Top quarks typically decay into a W boson and a b quark. When the W boson decays leptonically, a neutrino is produced, generating genuine missing transverse momentum. If the event is not removed by the b -jet veto and the lepton is not identified, this type of events contributes to the background in the signal region. However, due to the small production cross section and the applied event selection, only a small fraction of these events are selected. In order to estimate the contribution of this background, a $t\bar{t}$ sample has been produced at LO with MADGRAPH, and single top events were generated with POWHEG at NLO.

When one of the weak bosons produced in diboson events decays leptonically, generating one or more neutrinos, and the other one decays hadronically, jets and missing transverse momentum are produced. The samples used to estimate this background have been produced using PYTHIA.

Finally, when the leptons in $Z(l\bar{l}) + \text{jets}$ events are lost or out of the detector acceptance, these events can mimic the monojet signature as well. MC samples have been generated at LO using MADGRAPH in several bins of H_T in order to estimate the contribution from this sub-dominant background.

¹It can happen that a jet is not fully contained inside the calorimeters, and leaks into the muon system. In that case, part of the energy of the jet is lost, leading to missing transverse momentum.

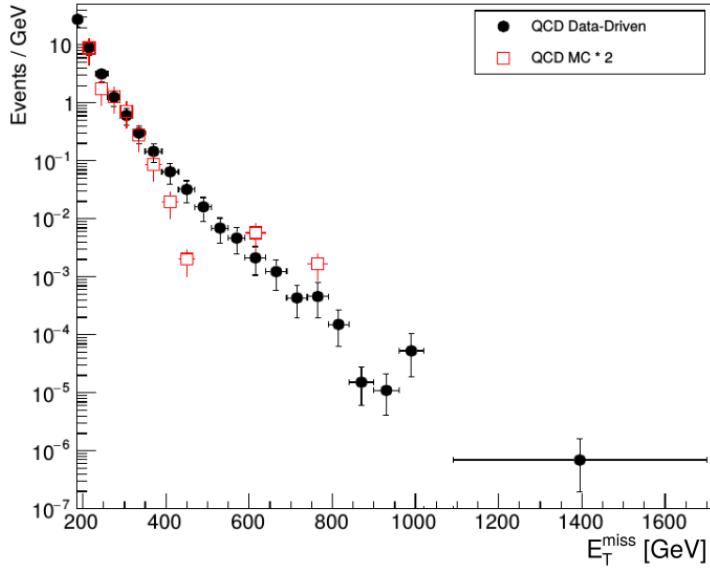


Figure 5.9: Predicted E_T^{miss} distribution of the QCD multijet background obtained using the $\Delta\phi$ extrapolation method, using data (black) and MC simulations (red). The simulation is scaled by a factor 2.

5.5 Systematic uncertainties

For the main backgrounds, multiple systematic uncertainties on the transfer factors are taken into account. Systematic uncertainties are added for the selection efficiency of muons, electrons, tau leptons, and photons, and for the photon purity in the photon + jets sample. Additionally, systematic uncertainties are added to take into account the electron and photon triggers. These uncertainties are fully correlated across the bins in recoil. The uncertainty on the modelling of the E_T^{miss} is dominated by the jet energy scale uncertainty, varying between 3 and 5%.

Uncertainties are also added from theory, to take into account variations of the factorisation and renormalisation scales, PDF uncertainties, and the NLO electroweak corrections [150–153]. The former 3 uncertainties are shown in Figure 5.10 for the $Z + \text{jets}$, $W + \text{jets}$, and photon + jets samples. The uncertainties are then propagated to the transfer factors, and are displayed in Figure 5.11. To evaluate the PDF uncertainty, the samples are reweighted with event-by-event scale factors representing the shift in the kinematic distributions from variations in the PDF. The transfer factors are then produced for each variation, and the RMS of the variation is taken as PDF uncertainty. Similarly, the renormalisation and factorisation scales are varied up and down by a factor 2, and the uncertainties are derived from the resulting transfer factors. For the electroweak corrections, the full correction is taken as an uncertainty.

The overall uncertainty on the QCD background varies between 50% and 150% depending on the E_T^{miss} region, and includes variations in the jet response and the statistical uncertainty on the transfer factors. The remaining subdominant top quark and diboson backgrounds are obtained directly from simulation. A systematic uncertainty of 20% is taken into account for the uncertainty on the production cross section of these processes. Additionally, a systematic uncertainty is added due to the uncertainty on the b-tagging applied for the b-jet veto, which amounts to 6% for the top quark background and 2% for the diboson background. The uncertainty on the modelling of the E_T^{miss} in simulation varies between 2% and 5%, depending on the E_T^{miss} . Finally, a systematic uncertainty of 2.7% [154] is added for all backgrounds derived from MC simulations, to take into account the uncertainty on the luminosity measurement.

Lastly, for the signal models described in Section 2.2.4 and simulated as summarised in Section 4.1.1, systematic uncertainties are included for the luminosity measurement (2.7%), the b-jet veto (2%), and the modelling of the E_T^{miss} distribution in simulation (2% to 5%). The

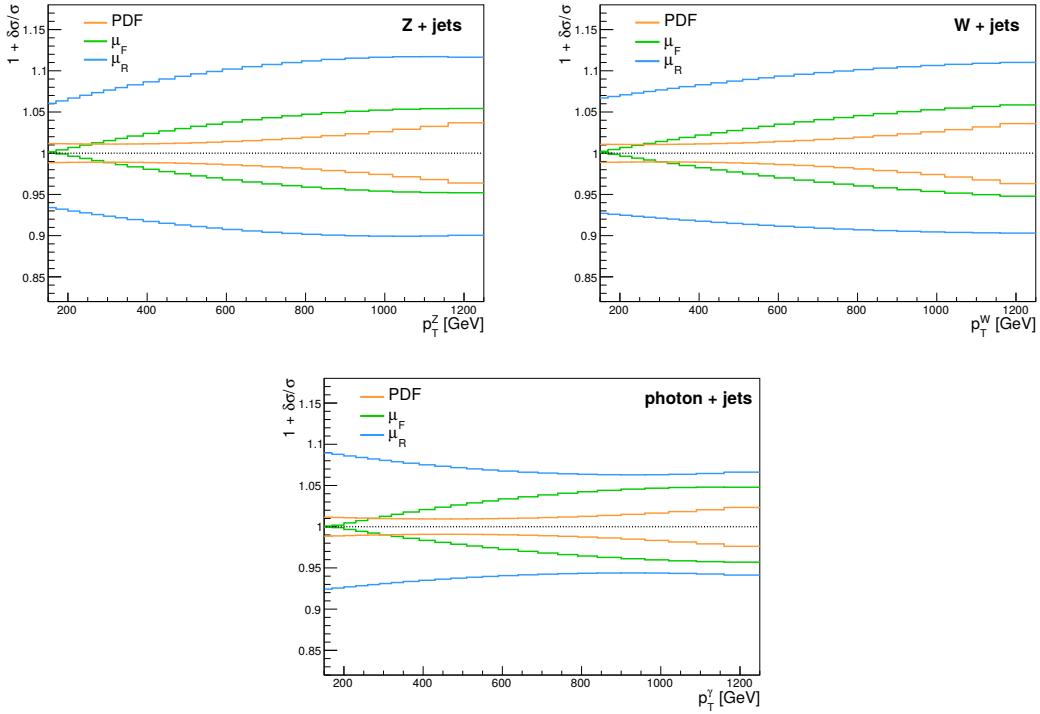


Figure 5.10: The PDF, renormalisation, and factorisation scale uncertainties for the $Z + \text{jets}$ (top left), $W + \text{jets}$ (top right), and photon + jets (bottom) samples. The uncertainties from the renormalisation and factorisation scales are obtained by separately varying them up and down by a factor 2.

systematic uncertainties coming from variations of the factorisation and renormalisation scales and PDF uncertainties amount to 20% for the vector and axial-vector signal samples and 30% for the scalar and pseudoscalar signals.

5.6 Results

The results are extracted by performing a binned fit to the missing momentum spectrum, fitting simultaneously over the five control regions and the signal region, under a given signal hypothesis. This is done using the CL_S criterion [155, 156] in the asymptotic approximation, described in Section 5.8, using the RooStat-based [157] Combine tool [158]. The systematic uncertainties described in Section 5.5 are modelled as nuisance parameters and their uncertainties are propagated as shape and normalization variations of the $Z(\nu\nu) + \text{jets}$ and $W(l\nu) + \text{jets}$ background. In Figure 5.12, the nuisance parameters and their normalised uncertainties are shown before and after the fit to the data in the control regions. Before the fit they are all centred at 0 and have an uncertainty of 1. After the fit, the two largest deviations originate from the statistical uncertainty on the Z/W transfer factor and the uncertainty on the muon scale factor, but no significant tension is present in the fit and most nuisance parameters are not biased or constrained by the fit. The photon + jets control region, which has the largest yields, drives the fit and the post-fit prediction therefore corresponds well to the data. A good agreement is obtained in all control regions, and the overall change in the transfer factors is less than 10%.

Table 5.1 gives the background prediction for the various background processes in bins of hadronic recoil, using a background-only fit. Correspondingly, the fit under the background-only hypothesis is shown in Figure 5.13 and a good agreement is observed between the data and the prediction. The uncertainty on the hadronic recoil is below 10% for all bins.

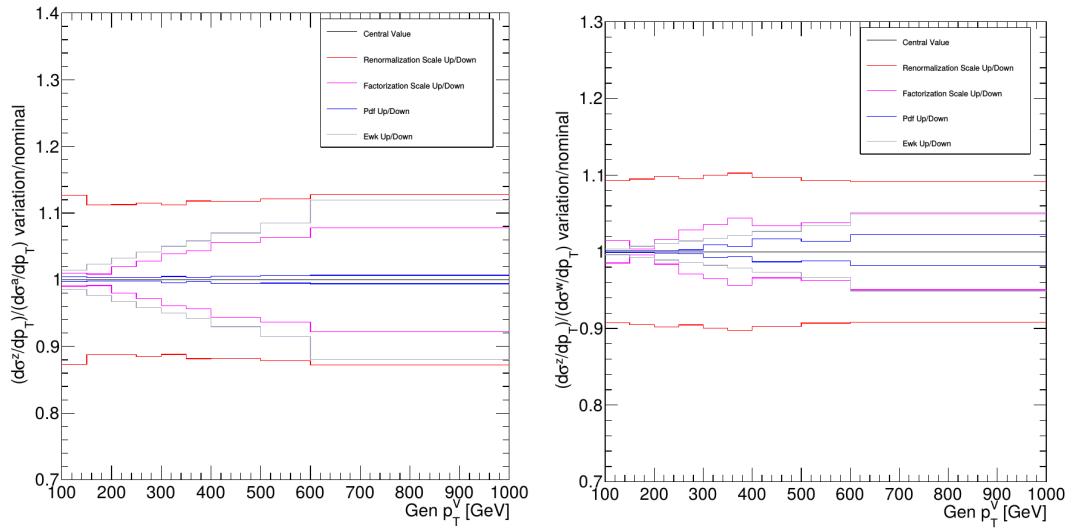


Figure 5.11: The theoretical uncertainties on the Z/γ (left) and Z/W (right) transfer factors.

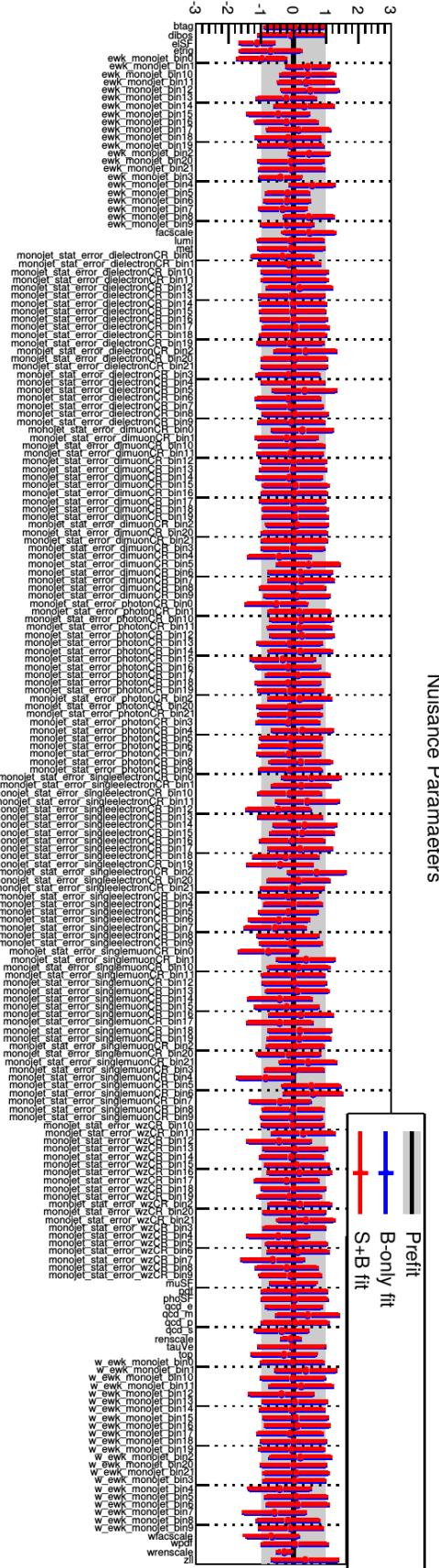


Figure 5.12: The post-fit nuisance parameters and their uncertainties, compared to the pre-fit values.

process	E_T^{miss} [200 - 230] GeV	E_T^{miss} [230 - 260] GeV	E_T^{miss} [260 - 290] GeV	E_T^{miss} [290 - 320] GeV	E_T^{miss} [320 - 350] GeV	E_T^{miss} [350 - 390] GeV	E_T^{miss} [390 - 430] GeV	E_T^{miss} [430 - 470] GeV	E_T^{miss} [470 - 510] GeV	E_T^{miss} [510 - 550] GeV	E_T^{miss} [550 - 590] GeV	E_T^{miss} [590 - 640] GeV
$Z(\nu\nu) + \text{jets}$	497	266	149	83.9	49.9	30.1	17.1	9.54	6.19	4.00	2.49	1.55
$W(l\nu) + \text{jets}$	399	193	95.6	50.7	27.3	13.9	6.87	3.86	2.18	1.30	0.73	0.38
$Z(ll) + \text{jets}$	6.88	3.10	1.26	0.61	0.33	0.10	0.05	0.02	0.01	0.01	0	0
top	18.8	8.92	3.85	1.89	1.12	0.61	0.24	0.23	0.01	0.07	0.02	0
QCD	16.9	10.3	1.28	0.99	0.17	0.10	0.13	0.41	0.01	0	0.03	0
dibosons	8.36	5.23	2.58	1.43	0.85	0.55	0.35	0.17	0.13	0.05	0.05	0.03
photon + jets	7.67	3.36	2.12	0.99	0.66	0.32	0.21	0.12	0.09	0.05	0.05	0.01
total post-fit	955 ± 6	489 ± 3	256 ± 2	141 ± 2	80.3 ± 1.2	45.7 ± 0.8	25.0 ± 0.6	14.4 ± 0.4	8.62 ± 0.32	5.49 ± 0.24	3.37 ± 0.17	1.97 ± 0.11
total pre-fit	925	470	240	136	79.5	45.4	24.5	14.7	8.45	5.30	3.37	2.02
observed	953	492	259	140	78.8	46.9	25.2	13.6	8.73	5.40	3.55	2.22

process	E_T^{miss} [640 - 690] GeV	E_T^{miss} [690 - 740] GeV	E_T^{miss} [740 - 790] GeV	E_T^{miss} [790 - 840] GeV	E_T^{miss} [840 - 900] GeV	E_T^{miss} [900 - 960] GeV	E_T^{miss} [960 - 1020] GeV	E_T^{miss} [1020 - 1090] GeV	E_T^{miss} [1090 - 1160] GeV	E_T^{miss} > 1160 GeV
$Z(\nu\nu) + \text{jets}$	0.90	0.56	0.44	0.27	0.16	0.09	0.06	0.04	0.02	0.02
$W(l\nu) + \text{jets}$	0.22	0.12	0.11	0.06	0.03	0.02	0.01	0.01	0	0
$Z(ll) + \text{jets}$	0	0	0	0	0	0	0	0	0	0
top	0	0	0	0.01	0	0	0	0	0	0
QCD	0.01	0	0	0	0	0	0	0	0	0
dibosons	0.03	0.01	0	0	0	0.01	0	0	0	0
photon + jets	0	0.01	0.01	0	0	0	0	0	0	0
total post-fit	1.16 ± 0.08	0.70 ± 0.05	0.55 ± 0.05	0.34 ± 0.04	0.20 ± 0.02	0.11 ± 0.02	0.07 ± 0.02	0.05 ± 0.01	0.03 ± 0.01	0.02 ± 0.01
total pre-fit	1.19	0.73	0.48	0.31	0.20	0.13	0.09	0.05	0.03	0.02
observed	1.22	0.64	0.56	0.28	0.22	0.12	0.05	0.01	0.03	0.03

Table 5.1: Pre-fit and post-fit background predictions in the signal region and observed yield given as number of events divided by the bin width, per E_T^{miss} bin. The predictions and uncertainties are obtained from the background-only simultaneous fit in the signal and control regions.

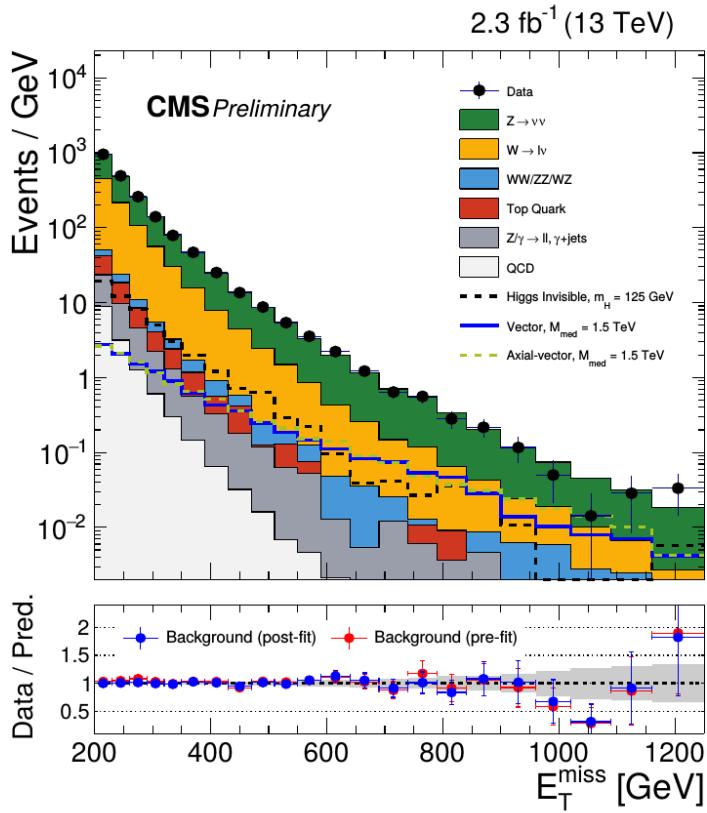


Figure 5.13: Post-fit missing transverse momentum distribution for the expected backgrounds and the observed data in the signal region. The grey band indicates the post-fit uncertainty obtained from the background-only simultaneous fit in the signal and control regions.

5.7 Improvements going from the 2015 to 2016 analysis

Several iterations of the monojet analysis were performed during Run 2 in 2015 and 2016, improving the achieved sensitivity with the increase in available data and additional developments to the analysis strategy, and adding new interpretations. In 2015, a first analysis with data at a centre-of-mass energy of 13 TeV was carried out [159], using data corresponding to an integrated luminosity of 2.1 fb^{-1} . In this first iteration of the search, the NLO k-factors applied to correct the LO samples used for the estimation of the background from Z and W bosons are derived from inclusive NLO samples, which are not binned in boson p_T . These k-factors are shown in Figure 5.14, and can be compared to the k-factors obtained from p_T -binned samples used in the following iterations of the analysis and shown in Figure 5.8.

In 2016, the analysis was extended to the mono-V channel [160], which can give rise to a monojet-like signature. At high p_T , the production of a W or Z boson which decays hadronically, can be effectively reconstructed as a single jet of large cone radius. However, in this chapter the focus is on the monojet channel and the added improvements during 2016. For this analysis, data collected in 2015 was used, corresponding to 2.3 fb^{-1} . This is the analysis reported in this chapter. The used NLO k-factors were derived using the p_T -binned samples, as described in Section 5.4.1. This contribution was a first step in the improvements I added to the analysis. I performed extensive validations, comparing the obtained k-factors to the previous results.

A substantial difference in the next iteration of the 2016 analysis [79] is the increase in collected data, which allowed to reduce the statistical uncertainties and set stronger limits on the

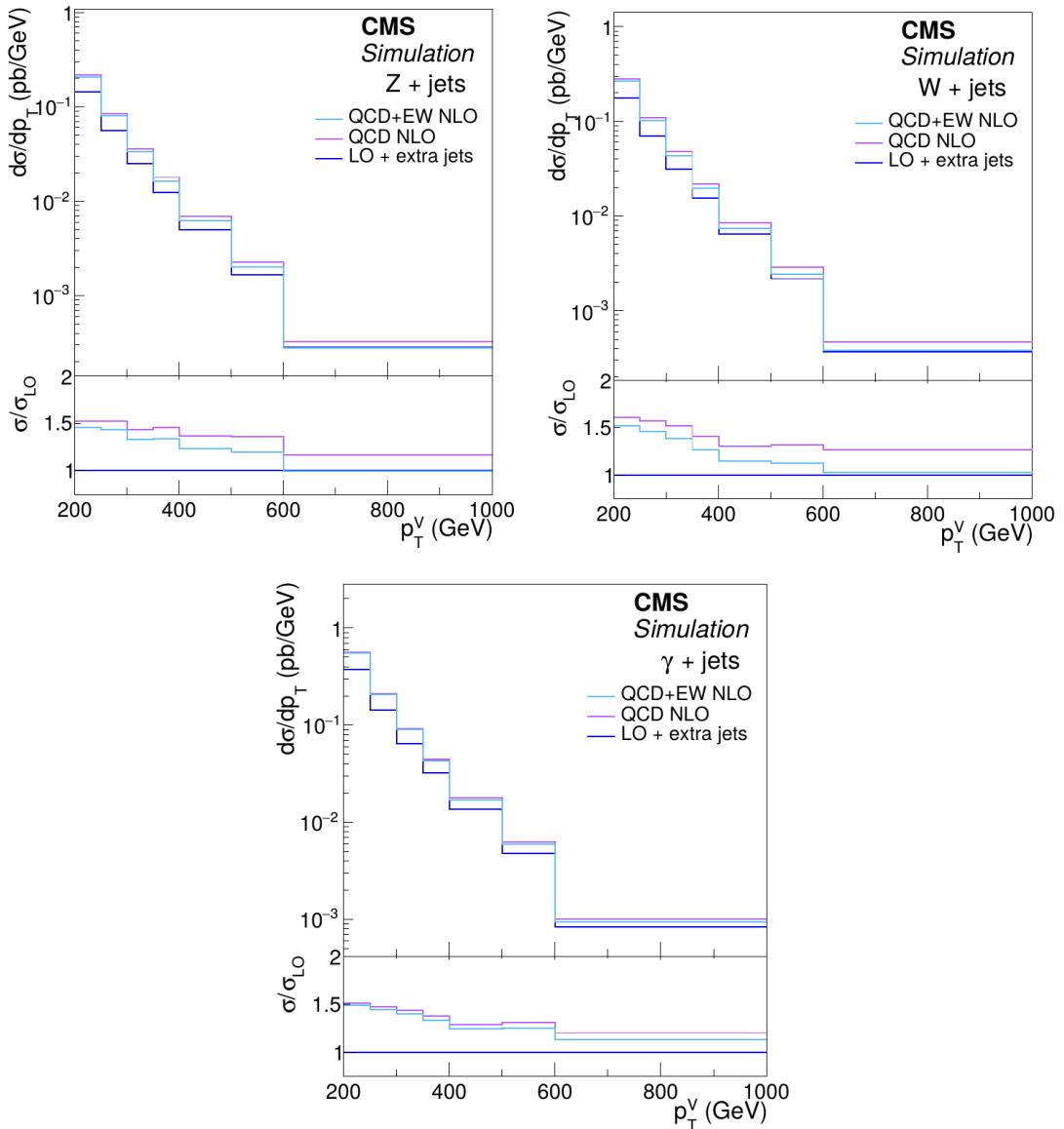


Figure 5.14: The differential cross section as a function of the boson p_T for photon, W , and Z production, using inclusive NLO samples. The resulting k-factors are shown in the ratio plots.

considered models. For these results, data collected in the first half of 2016 and corresponding to an integrated luminosity of 12.9 fb^{-1} were available. One of the improvements that were added is the direct use of MC samples generated at next-to-leading order for the estimation of the main backgrounds. This was possible by generating larger samples that are binned in W boson p_T , and are now used more widely in other analyses. As a result of my second contribution to the improvement of the background prediction, no k-factors need to be applied to the $W + \text{jets}$ sample and no additional systematic uncertainties are to be introduced for this. Figures 5.15 and 5.16 show the obtained reduction in the overall post-fit uncertainty on the predictions in the 5 control regions, which is clearly visible at high missing transverse momentum. The largest uncertainties of about 50% in the high E_T^{miss} region are reduced to 20% or less, which is driving the analysis sensitivity at the highest mediator masses.

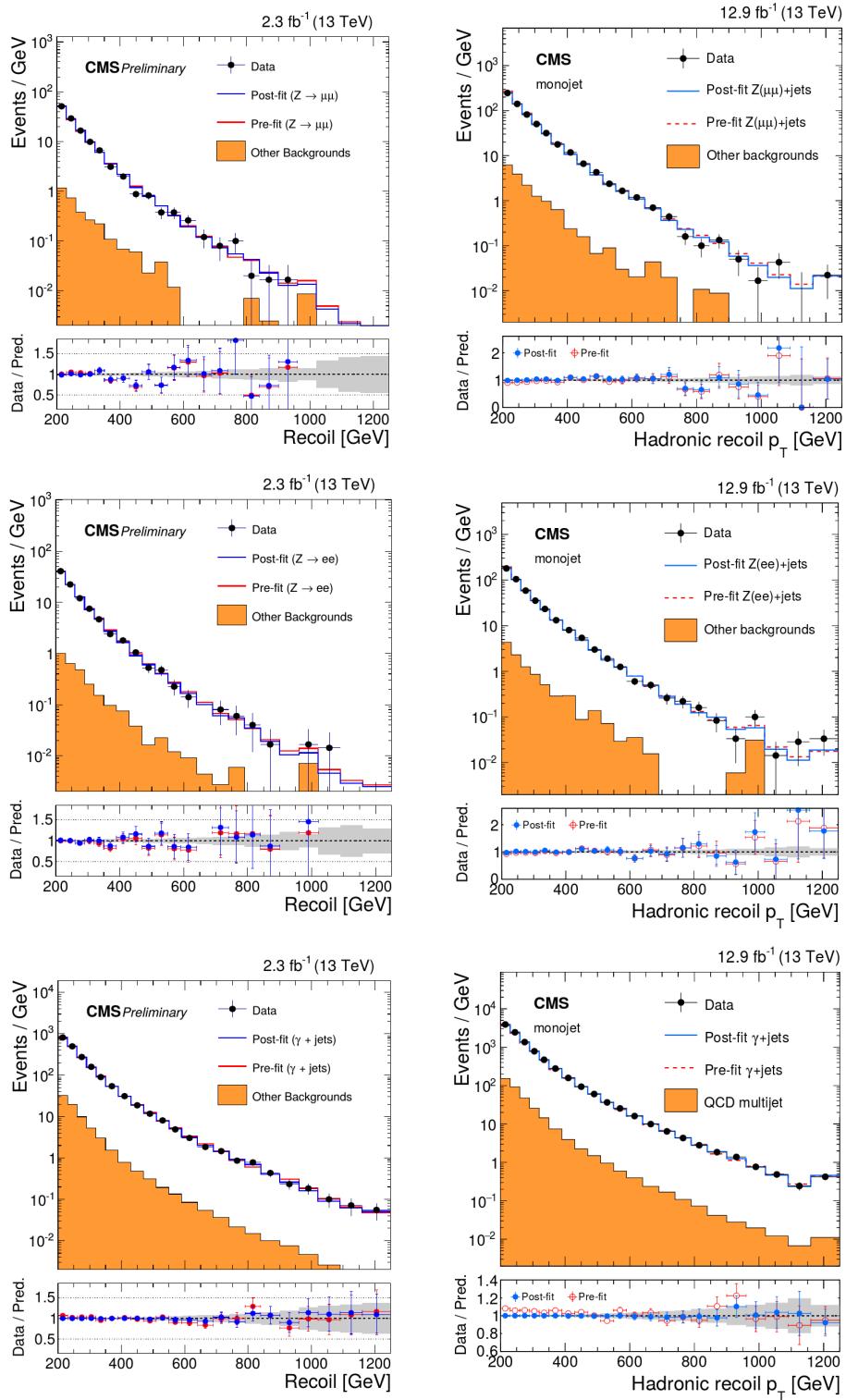


Figure 5.15: Comparison of the hadronic recoil p_T distribution between the data and MC simulation in the dimuon (top), dielectron (middle), and photon + jets (bottom) control regions, showing the reduction in overall post-fit uncertainty (grey band) going from the early 2016 analysis (left) to the analysis using data from the first half of 2016 (right). The fit was performed simultaneously over all the control and signal regions, assuming the absence of any signal. The last bin includes all events with hadronic recoil $p_T > 1160$ GeV. Figures taken from [160] and [79].

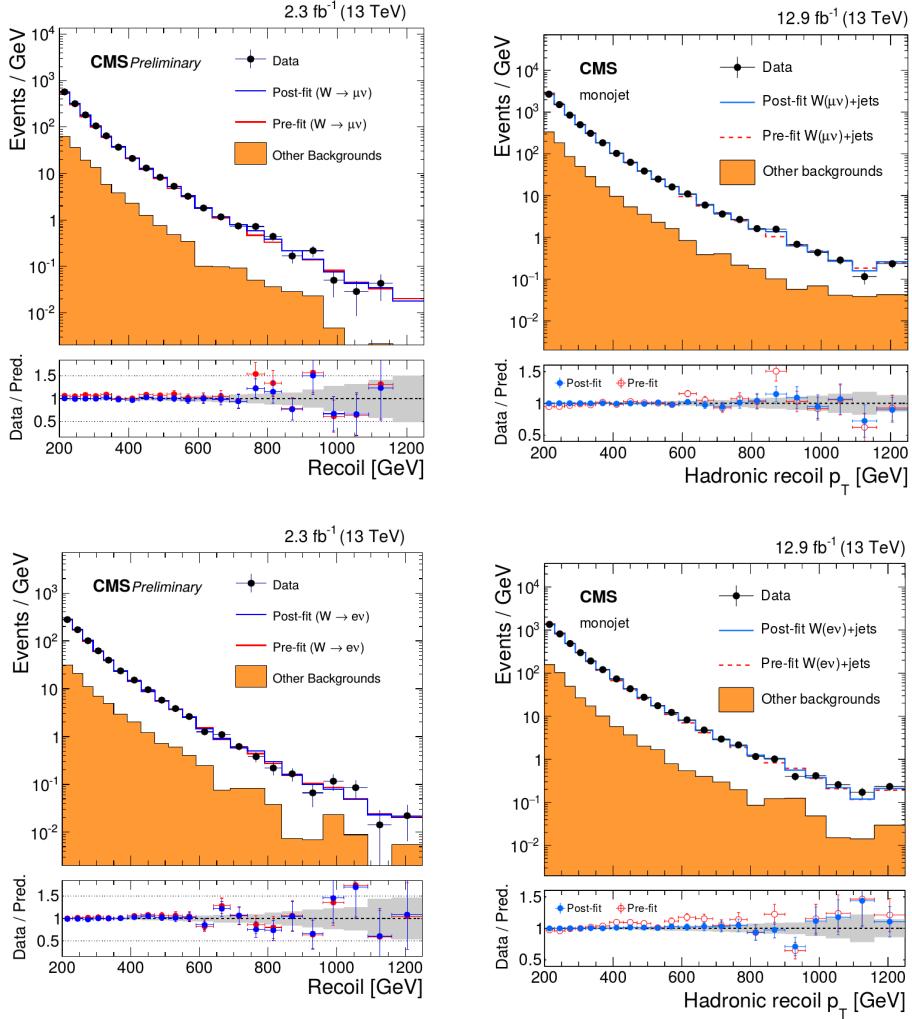


Figure 5.16: Comparison of the hadronic recoil p_T distribution between the data and MC simulation in the single muon (top) and single electron (bottom) control regions, showing the reduction in overall post-fit uncertainty (grey band) going from the early 2016 analysis (left) to the analysis using data from the first half of 2016 (right). The fit was performed simultaneously over all the control and signal regions, assuming the absence of any signal. The last bin includes all events with hadronic recoil $p_T > 1160$ GeV. Figures taken from [160] and [79].

5.8 Limit setting

As no significant excess was observed, upper limits are placed on the ratio of the signal cross section to the predicted cross section. This limit setting on new physics models is done with respect to the signal plus background hypothesis, comparing it to the background only hypothesis in which no new physics is present. Typically, a specific variable, for which different outcomes are predicted in the different hypotheses, is used to build a test statistic. The distribution of the test statistic is constructed by generating pseudo-experiments for each hypothesis. In case the expected number of events is large, the asymptotic approximation [161] is used to reduce the large computing time, replacing the pseudo-experiments by a single so-called Asimov data set.

The CL_S method [155, 156] is then used to determine the compatibility of the signal hypothesis with the observed data and to derive the upper limits at a certain confidence level (CL). In this thesis, the LHC style test statistic [162] is used, given by

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})} \quad (0 \leq \hat{\mu} \leq \mu), \quad (5.4)$$

where μ is the signal strength, θ_μ are the nuisance parameters for a given μ , $\hat{\mu}$ is the maximum likelihood value of μ , and $\hat{\theta}_\mu$ are the corresponding nuisance parameters. The signal strength represents the fraction of the theoretical cross section corresponding to the considered hypothesis. When constructing the distribution of the used variable, the expected number of events in the i -th bin is then $\mu s_i + b_i$, with s_i and b_i the number of predicted signal and background events in that bin. $\mu = 0$ thus corresponds to the background only hypothesis, and $\mu = 1$ corresponds to the signal plus background hypothesis using the nominal theoretical cross section.

Adding a given observation of the test statistic, $q_{\mu, \text{obs}}$, from a measurement in data, the probability to find a data set with value of the test statistic equal or larger under the signal plus background hypothesis can be written as

$$p_{s+b} = \int_{q_{\mu, \text{obs}}}^{\infty} f_{s+b}(q_\mu | \mu) dq_\mu, \quad (5.5)$$

where $f_{s+b}(q_\mu | \mu)$ is the probability density function of the signal plus background hypothesis constructed from the pseudo-experiments. Similarly, the probability to find a data set equally or more incompatible with the background only hypothesis is given by

$$1 - p_b = \int_{-\infty}^{q_{\mu, \text{obs}}} f_b(q_\mu | 0) dq_\mu. \quad (5.6)$$

In these equations, p_{s+b} and p_b are the p-values associated to the two hypotheses. The CL_S value, which represents the compatibility of the signal hypothesis with the observed data, is defined as the ratio of these two probabilities:

$$CL_S = \frac{p_{s+b}}{1 - p_b} \quad (5.7)$$

Models can then be considered to be excluded at 95% CL when $CL_S \leq 0.05$.

5.9 Interpretation

The results of the second iteration of this search, using the mono-V final state as well, are interpreted in terms of simplified dark matter models assuming a vector, axial-vector, scalar, or pseudoscalar mediator decaying to a pair of fermionic dark matter particles, as described in Section 2.2.4. As no significant excess was observed, upper limits are placed at 95% confidence

level (CL) on the ratio of the signal cross section to the predicted cross section, $\mu = \sigma/\sigma_{th}$, following the method described in Section 5.8. These limits are shown as a function of the mediator mass (m_{med}) and the dark matter mass (m_{DM}) in Figures 5.17 and 5.18, for the four types of mediators. The regions where the 95% CL upper limits on μ are less than one are considered to be excluded.

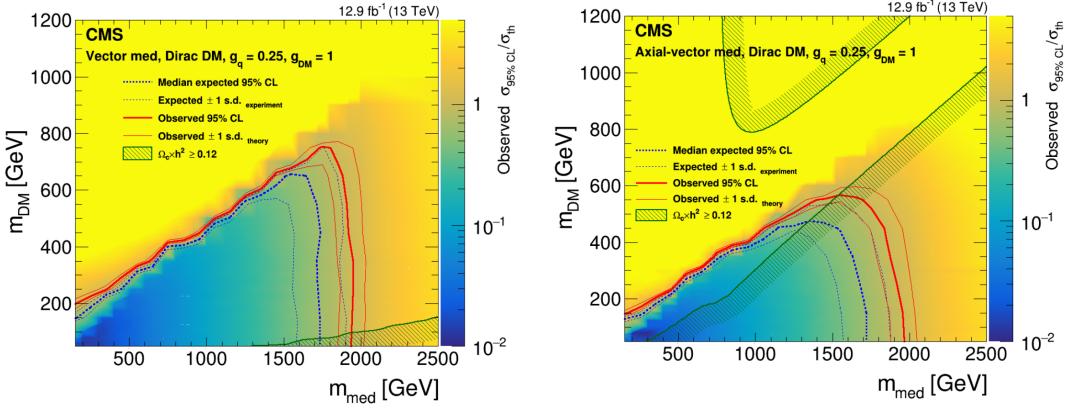


Figure 5.17: The 95% CL upper limits on the $\mu = \sigma/\sigma_{th}$ in the $m_{\text{med}} - m_{\text{DM}}$ plane, for a vector (left) or axial-vector (right) mediator. The cosmological constraints from the WMAP and Planck measurements of the CMB are shown as well. Figures taken from [79].

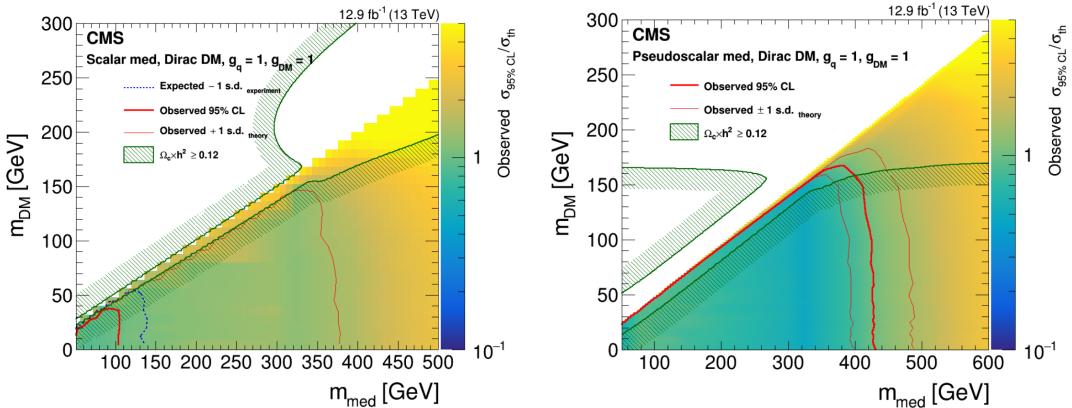


Figure 5.18: The 95% CL upper limits on the $\mu = \sigma/\sigma_{th}$ in the $m_{\text{med}} - m_{\text{DM}}$ plane, for a scalar (left) or pseudoscalar (right) mediator. The cosmological constraints from the WMAP and Planck measurements of the CMB are shown as well. Figures taken from [79].

For the vector and axial-vector models, mediator masses up to 1.95 TeV are excluded, while mediator masses up to 100 GeV and 430 GeV can be excluded for the scalar and pseudoscalar models, respectively. Figures 5.17 and 5.18 also show the constraint from the observed cosmological relic density of dark matter, which was determined using a thermal freeze-out mechanism and compared to the observed cold dark matter density, assuming that the considered dark matter candidate is the dominant component of the observed dark matter.

The obtained limits can also be translated into 90% CL upper limits on the dark matter-nucleon scattering cross section $\sigma_{\text{SI/SD}}$, in order to compare them to results from direct detection experiments. The exclusion contours in the $m_{\text{DM}} - \sigma_{\text{SI/SD}}$ plane are obtained following the approaches outlined in [163–165], where σ_{SI} stands for spin-independent and σ_{SD} for spin-

dependent dark matter-nucleon cross section. The resulting limits are shown in Figures 5.19 and 5.20.

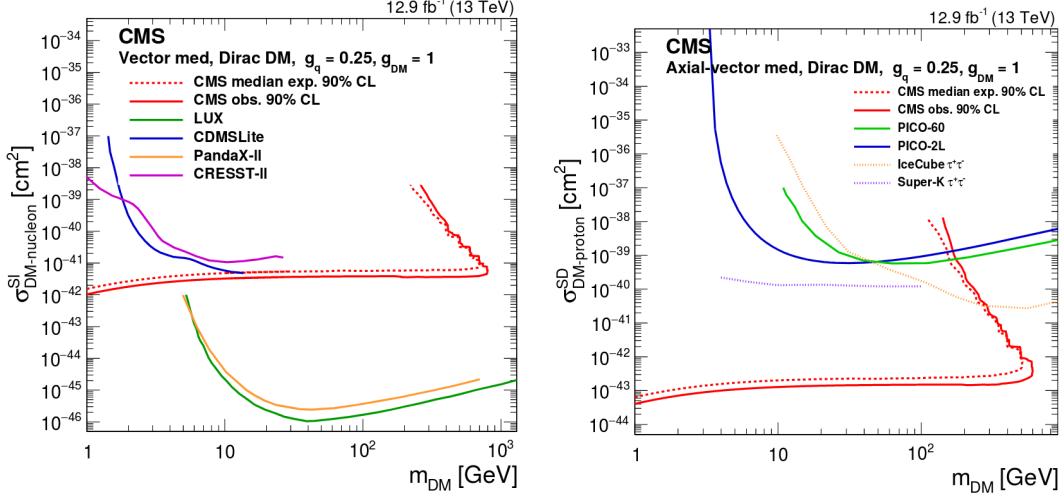


Figure 5.19: The 90% CL upper limits in the $m_{\text{DM}} - \sigma_{\text{SI/SD}}$ plane, for a vector (left) or axial-vector (right) mediator. Figures taken from [79].

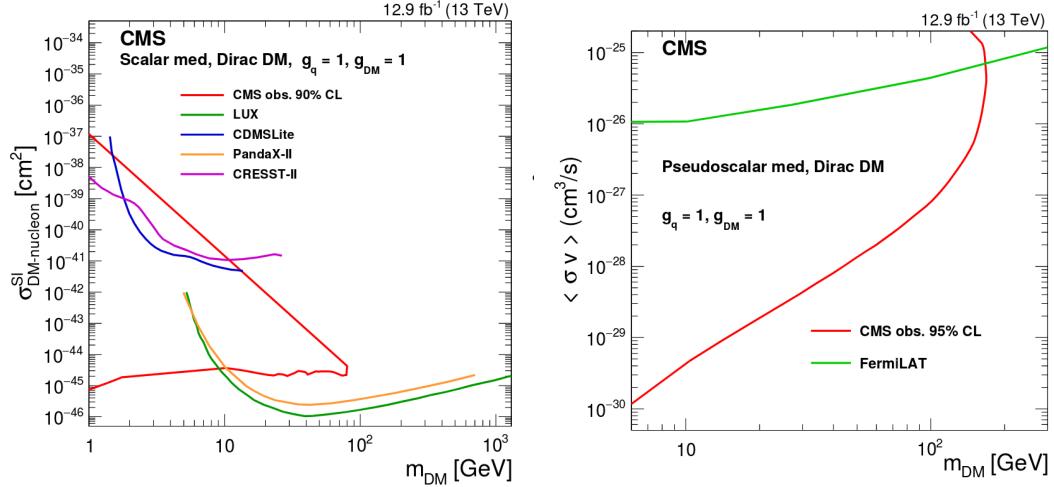


Figure 5.20: The 90% CL upper limits in the $m_{\text{DM}} - \sigma_{\text{SI/SD}}$ plane, for a scalar (left) or pseudoscalar (right) mediator. Figures taken from [79].

For the vector and scalar mediator cases, the derived 90% CL upper limits on the spin-independent cross section are compared to the results from the CDMSLite [166], LUX [167], CRESST-II [168], and PandaX-II [169] experiments. This shows that the monojet limits are complementary to the results from the direct detection experiments at low dark matter masses. The 90% CL upper limits on the spin-dependent cross section for the axial-vector mediator are compared to the PICO-2L [170], PICO-60 [171], IceCube [172], and Super Kamiokande experiments [173]. In the case of the pseudoscalar mediator, the obtained limits are compared to the indirect detection results from the Fermi-LAT experiment [174, 175]. In the considered scenario the dark matter annihilates into b quark pairs in the centre of the galaxy, producing gamma rays.

6

Search for SIMPs using Trackless Jets

The monojet dark matter search detailed in Chapter 5 can be complemented at high interaction cross sections by a different search which does not look for dark matter in the form of missing transverse momentum. Indeed, if the dark matter particles have an interaction cross section of the order of the strong interaction, they will interact in the detector, mainly in the calorimeters. The analysis described in this chapter is based on this scenario, and the considered simplified model is specified in Section 2.3.1. In this model, the dark matter particles are produced in pairs through a new strongly interacting mediator, and give rise to a pair of trackless jets as signature. Since the produced strongly interacting massive particles (SIMPs) are neutral, the resulting jets can be distinguished from QCD jets using the jet charged energy fraction (ChF). The signal region where the SIMPs are being looked for is therefore defined by requiring the two leading jets to have a low ChF. In the control region, which is used to predict and validate the expected background, one or both jets are required to have a large ChF.

Firstly, the applied jet and photon reconstruction and identification are described, as well as the specific treatment applied for the primary vertex selection. The triggers that were designed specifically for this search, exploiting the ChF, are outlined in Section 6.2. However, these triggers were found to be problematic and eventually a generic single jet trigger was used instead. Next, the event selection is detailed in Section 6.3 for both the signal and control regions. The strategy for the background estimation and the systematic uncertainties are discussed in Sections 6.4 and 6.5, respectively. Finally, in Sections 6.6 and 6.7, the results are shown and interpreted in terms of the SIMP simplified model described in Section 2.3.

6.1 Physics object reconstruction

In this analysis, jets with a very small charged energy fraction (ChF) are being searched for. Since these are rather peculiar jets containing no or very few tracks, a good primary vertex selection and photon identification play an important role in suppressing the main physics and reconstruction backgrounds.

6.1.1 Jets

For the jet reconstruction, the standard method described in Section 4.3.7 is used. Although the jets in the signal samples are expected to be neutral, it is beneficial to use PF jets because they directly provide an unambiguous association of tracks to jets. The standard jet energy

corrections are applied as well, while the standard jet identification criteria are not used, since several of the quality criteria would actually remove the neutral SIMP jets.

6.1.2 Photons

Since photons might be reconstructed as neutral jets, photon + jets events are an important background for the control as well as the signal region. The photons therefore need to be identified and rejected, which is done using the standard photon loose identification described in Section 4.3.3. Further photon rejection is achieved by analysis-specific selections on photon conversions and on the jet neutral electromagnetic energy fraction (NEMF), as described in Section 6.3.

6.1.3 Primary vertex

The standard primary vertex reconstruction described in Section 4.3.1 sometimes provides a wrong primary vertex, which arises from a pileup interaction. The choice of a wrong vertex is not a problem in the case of signal events, which will pass in particular the ChF cuts in the event selection detailed in Section 6.3 just as easily. However, a wrongly chosen vertex in a QCD background event leads to the jets having an artificially very low ChF, both in simulation and data, as the standard charged hadron subtraction (CHS) procedure will remove the tracks from the vertex of the true hard interaction. This makes such events appear signal-like. For the lowest jet charged energy fractions considered in this analysis, this background of events with a misidentified primary vertex becomes dominant with respect to the background from QCD events with a very rare jet fragmentation into predominantly neutral hadrons and photons. In Figure 6.1 an event display is shown that demonstrates such a wrong choice of vertex.

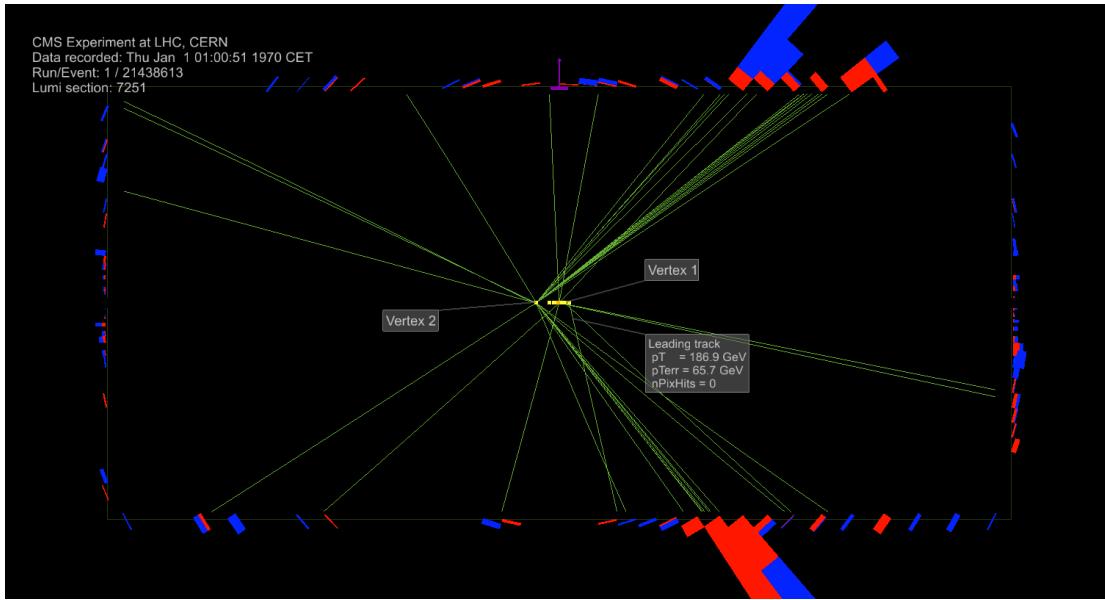


Figure 6.1: Event display showing an example of a wrong primary vertex selection in a MC simulated QCD event. Although "Vertex 2" is the real vertex of the hard collision, "Vertex 1" is selected because of the presence of a single high- p_T track with poor momentum resolution and no pixel hits attached. As a result, the two visible high- p_T jets (above 200 GeV), clearly with many tracks attached, are reconstructed with only 3% and 4% of charged energy fraction.

Investigating this problem, many events with a wrong vertex assignment were observed to have the highest- p_T track being of poor quality, with a high momentum with large uncertainty, and no pixel hits. This alone, though, does not provide a sufficient handle to suppress this

background, and a plain cut on number of pixel hits was verified to remove a lot of signal events as well. Studying the simulated QCD multijet events with a very low ChF coming from primary vertex misidentification by analysing event displays also showed that the true vertex is reconstructed as the second vertex in the list for the far majority of the cases. A second jet reconstruction was therefore produced, based on selecting the second entry in the list of primary vertices to be the event’s collision vertex, and rerunning the CHS as well. If this second vertex was the correct one, the jets will now have a large ChF in most of the cases, while the first event reconstruction yields low-ChF jets. In the event selection discussed in Section 6.3, it is then sufficient to ask both event reconstructions to pass the cut of low jet ChF, effectively suppressing this background induced from wrong primary vertex selection.

6.2 Trigger selection

Several triggers have been designed specifically for this analysis, as shown in Table 6.1. Four triggers were available to select signal events, denoted as “signal” in the table and providing a trade-off between a high ChF and a low jet p_T . The ChF cut is always applied on both jets. Two additional triggers, indicated with the “control” type, were used to determine the trigger efficiencies. These triggers were prescaled, meaning that only a fraction of the events are kept due to the otherwise very high rate of the trigger.

Trigger name	Type	Int. Lumi.
HLT_DiCentralPFJet170_CFMMax0p1_v*	Signal	33.1 fb^{-1}
HLT_DiCentralPFJet220_CFMMax0p3_v* ¹	Signal	5.91 fb^{-1}
HLT_DiCentralPFJet330_CFMMax0p5_v*	Signal	33.1 fb^{-1}
HLT_DiCentralPFJet430_v*	Signal	33.1 fb^{-1}
HLT_DiCentralPFJet170_v*	Control	0.101 fb^{-1}
HLT_SingleCentralPFJet170_CFMMax0p1_v*	Control	0.375 fb^{-1}

Table 6.1: Summary of the triggers designed for this analysis, with the corresponding integrated luminosity collected for each trigger.

These triggers were designed to select events containing jets with a low ChF, but keeping the jet p_T threshold as low as possible, while taking the limited trigger rate into account. Their performance was tested using simulated QCD events before using them in data taking. The trigger efficiency of `HLT_DiCentralPFJet330_CFMMax0p5` as a function of p_T and ChF is shown in Figure 6.2, using data and simulated QCD events. The plot on the left shows the efficiency as a function of p_T , which quickly reaches a plateau at 100% around 350 GeV, as expected. The efficiency as a function of ChF is shown on the right and also shows a plateau at 100% for low ChF. The slope of the turn-on is however less steep and only reaches the plateau around 0.35, for a threshold at ChF = 0.5 at trigger level. The signal efficiency measurement was first tried by taking a photon + jet sample, as photons could also mimic neutral jets. The photon was matched to the leading jet within a ΔR cone, and the efficiency was measured at low ChF. However, as the jet is reconstructed with a large electromagnetic energy fraction (EMF) and it does not mimic the SIMP signal.

Instead, one of the neutron signal samples described in Section 4.4 was used, generated with a SIMP mass of 700 GeV. The obtained trigger efficiency as a function of p_T and ChF is shown in Figure 6.3 for `HLT_DiCentralPFJet330_CFMMax0p5`, comparing the data and the signal-like events. This shows that the trigger reaches a plateau at only 60% signal efficiency.

¹Due to the unexpected high rate, this trigger was disabled after some time.

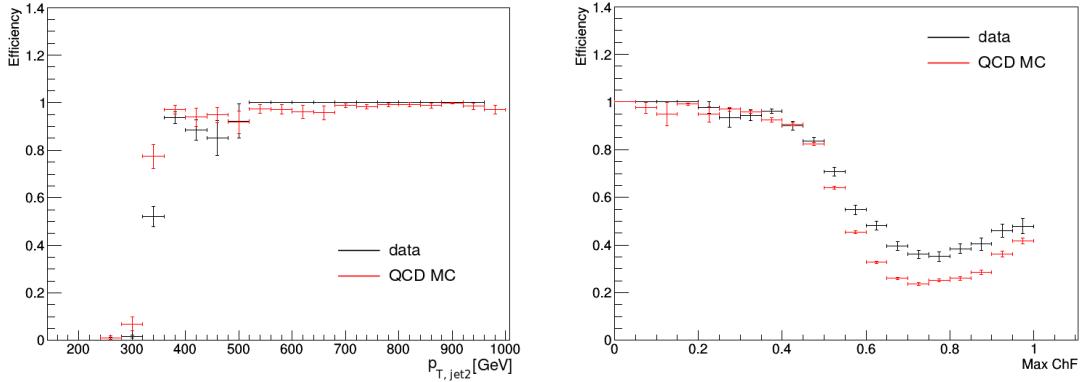


Figure 6.2: The HLT_DiCentralPFJet330_CFMMax0p5 trigger efficiency as a function of p_T (left) and ChF (right). Comparison between data (black) and QCD simulation (red).

The origin of the inefficiency was found to come from a hidden requirement asking at least one charged particle in the jets. This explains why the trigger efficiencies turn on at the expected values when using data or simulated QCD events, where jets nearly always contain some tracks, while it is largely inefficient in signal-like simulation which generates neutral jets.

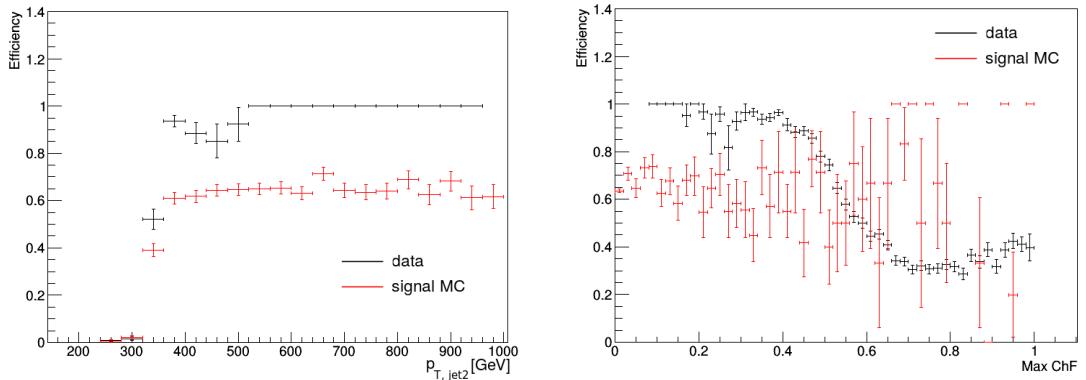


Figure 6.3: The HLT_DiCentralPFJet330_CFMMax0p5 trigger efficiency as a function of p_T (left) and ChF (right). Comparison between data (black) and signal-like MC (red).

After this problem had been uncovered, the single jet trigger HLT_PFJet 450 has been used for the analysis. It was selected because it has the lowest threshold among the unprescaled² single jet triggers. The trigger efficiency for the data was measured using single muon events, and is defined as

$$\epsilon = \frac{\text{Obs}(\text{HLT_PFJet 450 and HLT_IsoMu24})}{\text{Obs}(\text{HLT_IsoMu24})}, \quad (6.1)$$

where Obs in this case is the p_T spectrum of the leading jet for events that fired the single muon (and single jet) trigger. The jets are required to have a muon energy fraction smaller than 0.3 in order to avoid jets with large difference in the online and offline p_T . This difference can arise when the offline PF jet contains a high- p_T muon. Since the online jet is reconstructed using only information from the calorimeters, this can significantly change the total p_T of the jet between online and the offline PF jet. The trigger efficiency was also measured in the simulated QCD and

²Triggers with a lower p_T threshold exist, but these are prescaled. Prescaled triggers are typically not used to perform an analysis, as only a small fraction of the events that pass the trigger are kept, leading to large statistical uncertainties.

neutron signal samples. In this case the denominator is the p_T spectrum of leading jet without any trigger selection, while the numerator is the p_T spectrum of the events firing the single jet trigger. Figure 6.4 shows the turn-on curves for data, QCD events, and the signal-like events. The trigger efficiency was found to be 98% for jets with $p_T > 550$ GeV.

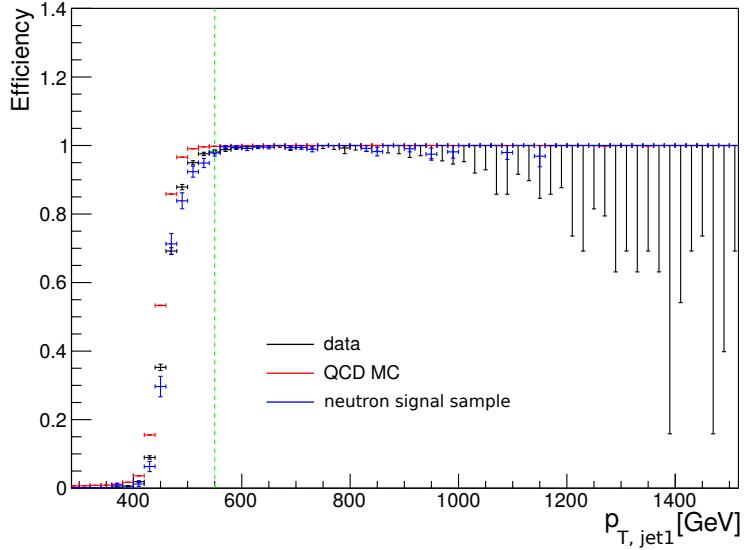


Figure 6.4: The HLT_PFJet450 trigger efficiency turn-on as a function of the offline leading jet p_T , for data, QCD events, and the signal-like events. The p_T cut used in the event selection is shown with a green dashed line.

6.3 Event selection

The event selection aims to select back-to-back dijet events with a low ChF. As a baseline selection, the two highest- p_T jets are required to have $p_T > 550$ GeV, in order to ensure the jets to be above the turn-on of the trigger. Furthermore, they are required to have $|\eta| < 2.0$, placing them fully within the tracking volume, thus suppressing backgrounds from jets that have a low ChF due to tracks falling out of tracker acceptance. As can be seen from the middle plots in Figure 6.5, this cut has only a small impact on the signal. Since the SIMPs do not undergo parton showering, while QCD partons undergo final state radiation, events with SIMPs have a lower number of jets than QCD multijet events, as can be seen from the top right plot in Figure 6.6. Events containing additional jets with $p_T > 30$ GeV in the full η acceptance of the CMS calorimeters on top of the two leading jets are therefore vetoed.

A photon veto is also applied to suppress photon + jets events. This is done by rejecting events with a photon within $\Delta R < 0.1$ of the leading or subleading jet, using the loose working point of the cut-based photon identification to identify photons, as described in Section 4.3.3. In some cases, however, jets have a large photon energy fraction, but the photon in the jet does not pass the loose identification requirements, for instance when there is a photon conversion in the tracker. In order to reject photon + jets events more efficiently, an additional cut is applied, as described in Section 4.3.3.

In order to avoid any problems related to the striking discrepancy between data and simulation observed for events with only one reconstructed vertex in the top left plot of Figure 6.6, at least two reconstructed vertices are required. Additionally, the azimuthal separation of the two selected jets is required to be $\Delta\phi > 2$ in order to obtain back-to-back jets, which rejects the peak visible in data at $\Delta\phi = 0$ in the bottom left plot of Figure 6.6. Finally, noise filters are applied

in order to reject beam halo or instrumental background, such as noise in the calorimeters.

Table 6.2 shows the number of events remaining in data, for simulated QCD events, and for 2 signal samples, after consecutively applying the described selection cuts. This shows that the background is already reduced by a factor 5, mainly by the cut on the number of jets, while a high efficiency is maintained for the signal. In Figure 6.5, data, QCD multijet simulation, and signal are compared after these selections, for the p_T , η and ChF of the two leading jets. Figure 6.6 shows the distribution of the number of vertices, the number of jets, $\Delta\phi(\text{jet1}, \text{jet2})$, and H_T , which is defined as the scalar sum of the transverse momenta of the two jets. In some cases, all the selection cuts except the cut on the variable being shown are applied. The bump and long tail that can be seen in the data for the $\Delta\phi(\text{jet1}, \text{jet2})$ distribution contain events coming from processes with a heavy vector boson, such as $Z + \text{jets}$ or $W + \text{jets}$, or $t\bar{t}$ events. The simulation instead shows a steeply falling spectrum since only QCD dijet events are shown. Contributions from $Z + \text{jets}$ or $W + \text{jets}$, and $t\bar{t}$ events were verified to be negligible in the signal region using simulated samples.

selection cut	yield			
	data	QCD MC	SIMP ($m_\chi = 1 \text{ GeV}$)	SIMP ($m_\chi = 1000 \text{ GeV}$)
$p_T^{j1,j2} > 550 \text{ GeV}$	2540420	3152550	604	5.6
$ \eta_{j1,j2} < 2.0$	2441240	2980320	581	5.5
# jets = 2	534053	587670	509	4.8
photon veto	531366	586674	507	4.8
# vertices ≥ 2	531244	586641	507	4.9
$\Delta\phi(j1, j2) > 2$	531207	586641	507	4.8
MET filters	528614	582184	493	4.7

Table 6.2: Number of events remaining after the listed selection cuts in data, QCD events, and 2 signal samples.

As will be detailed in Section 6.4, the background is predicted from a data control region where at least one of the leading jets has a high ChF, above 0.25. No further selection on the second reconstruction with respect to the second primary vertex is applied for the control region, since the presence of at least one jet with a large ChF avoids the problem of the wrong selection of primary vertex detailed in Section 6.1.

In the case of the signal region selection, both jets are required to have $\text{ChF} < x$, where x is the signal cut being considered. In this case, the cut is applied for both reconstructions starting from the first and second primary vertex.

6.4 Background estimation

The main background for this analysis is QCD multijets. This background is estimated from data, as the simulation does not describe the data well, especially at low ChF. The signal events can then be distinguished from this background using the jet ChF. A second background comes from photon + jets events. However, this background is efficiently removed by applying a photon veto.

6.4.1 Photon + jets

The photon + jets background was studied using a high- H_T MC sample generated at LO with **MADGRAPH5_aMC@NLO**, hadronized with **PYTHIA 8**, and simulated and reconstructed using the standard procedure described in Chapter 4. The used sample corresponds to about 27 fb^{-1} , which is larger than the data sample used for this analysis and was found to be sufficient to

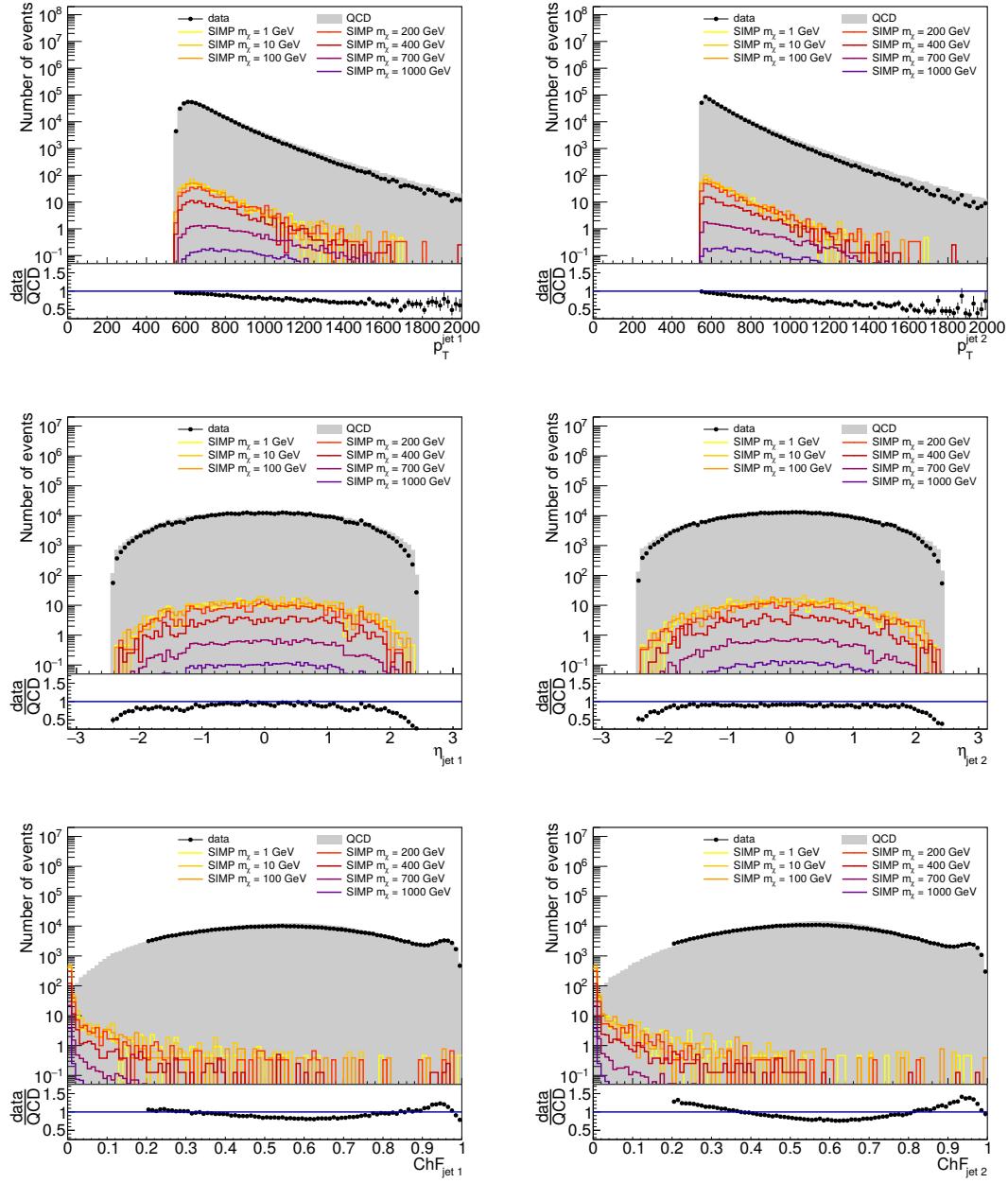


Figure 6.5: p_T , η and ChF of the leading (left) and subleading (right) jet. The selection cuts are applied, except for the cut on η in the corresponding plot.

evaluate its contribution in the signal region. This background is verified to be negligible after applying the signal region event selection, and well within any other systematic uncertainty on the background prediction. The photon veto works well, especially the cut on the jet neutral electromagnetic fraction, and no events from the used simulated photon + jets sample remain after applying a cut of $\text{ChF} < 0.1$. Additionally, in part of the events remaining just above that cut the photon is not identified because it is very close to a jet. These events are therefore already contained in the overlapping QCD multijets sample.

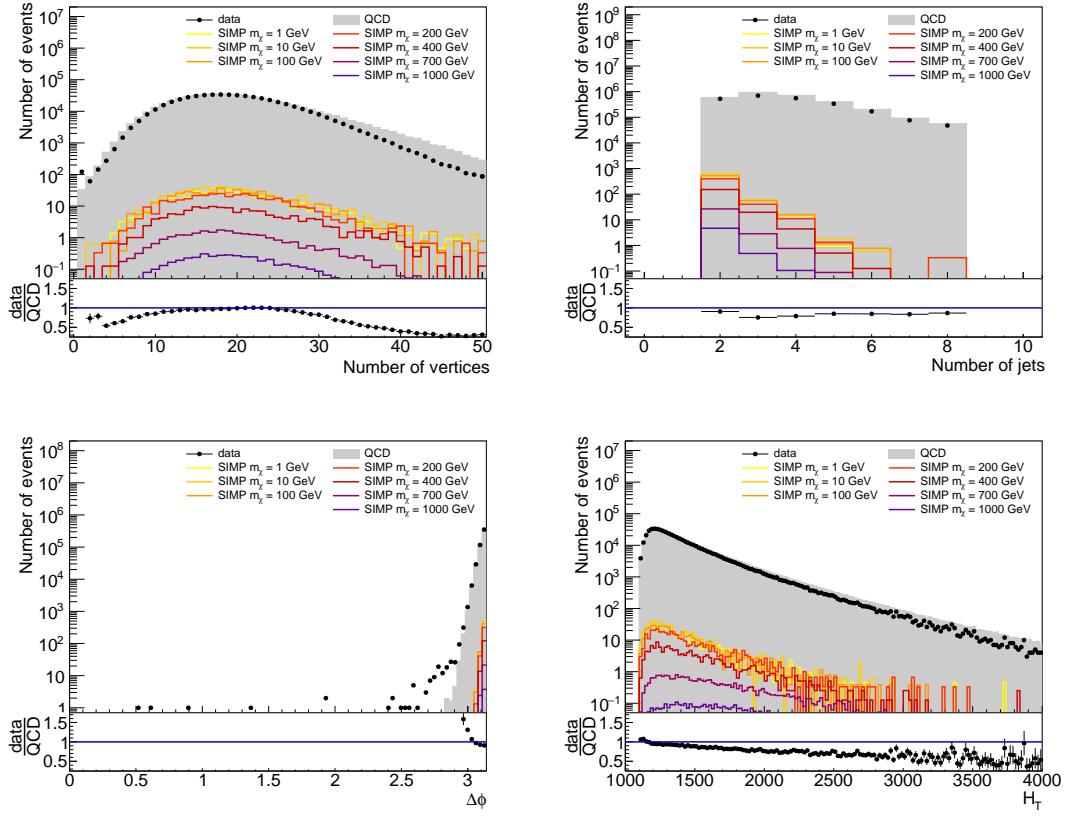


Figure 6.6: Number of vertices, number of jets in the full detector acceptance with $p_T > 30$ GeV, $\Delta\phi(\text{jet1}, \text{jet2})$, and H_T distributions, with selection cuts applied. The requirement on the number of vertices is not applied for the corresponding plot, the cut on the number of jets is not applied for the number of jets distribution, and the $\Delta\phi$ cut is similarly not applied for the corresponding distribution.

6.4.2 QCD multijets

The QCD multijet background is estimated from data, since the simulation does not describe the data well, especially at low ChF, as can be seen from Figure 6.7 which compares the ChF distribution in the control region to the QCD MC. In this plot the subleading jet is required to have a large ChF, in order to stay in the control region.

As a first step, the efficiency of the ChF cut is measured in the control region, by tagging one jet with high ChF and applying the ChF cut on the other jet. The efficiency is then given by the ratio of the number of events passing the cut divided by the total number of events selected in the control region. The measurement is performed in bins of jet p_T and η . The number of QCD events in the signal region is then predicted by using any QCD dijet event passing the selection cuts listed in Table 6.2 and applying the appropriate p_T and η dependent ChF cut efficiencies on the two leading jets. Figure 6.8 shows the measured efficiency as a function of the ChF cut for various bins in p_T^{jet} , integrating over the η_{jet} bins, and η_{jet} , integrating of the p_T^{jet} bins, as measured in the QCD MC. There is a strong dependence on the jet p_T , and a less pronounced dependence on the jet η at low ChF. The efficiencies are the highest for the $1.0 < |\eta| < 1.25$ bin, which can be attributed to this being the barrel-endcap transition region where most of the tracker material is located. As can be seen from the left plot in Figure 6.9, at generator level the ChF is independent of the jet p_T , as one would expect. However, after applying the jet reconstruction, a p_T dependence arises due to the known degradation of tracking efficiency in

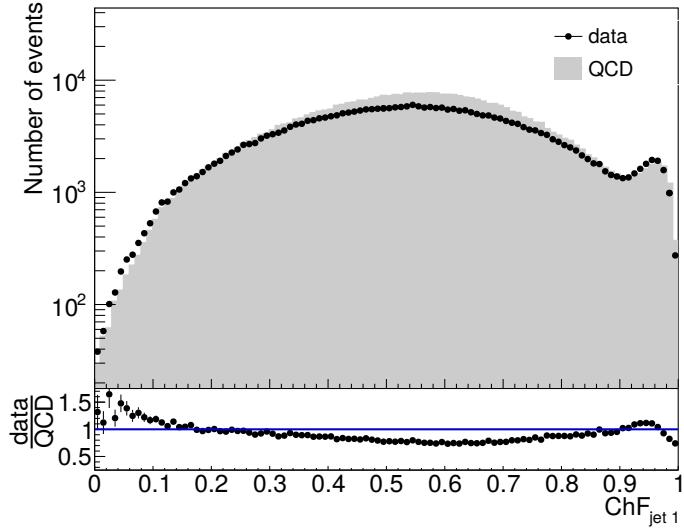


Figure 6.7: Data-MC comparison of the charged energy fraction of the leading jet, tagging events with subleading jet $\text{ChF} > 0.5$.

dense jet environments, which becomes more of an issue for very high p_T and thus collimated jets, as demonstrated in the right plot of Figure 6.9.

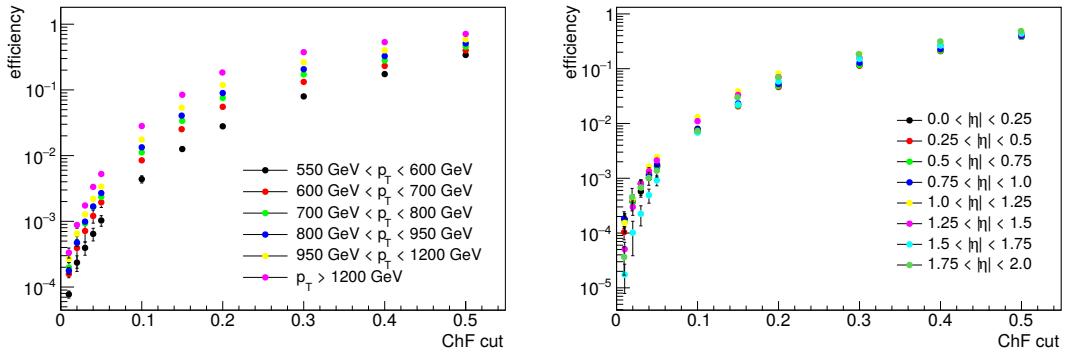


Figure 6.8: The efficiency of several ChF cuts in QCD MC, binned in p_T (left) and η (right).

A test of the background prediction method is performed using simulated events in order to validate it. This is done by comparing the MC truth and the so-called 1- and 2-leg predictions. The MC truth shows the yield after applying the ChF cuts on both jets. For the 1-leg prediction the ChF cut is applied on one jet and the event is then weighted by applying the measured p_T - and η -dependent efficiency for the other jet. For the 2-leg prediction, the efficiencies are applied for both jets, and no ChF cut is applied directly.

As a first check, the test was also performed at the generator level, using the GenJets, which are reconstructed from the generator level particles. This comparison is done in exclusive bins in $(\text{ChF}_{\text{jet}1}, \text{ChF}_{\text{jet}2})$, as illustrated in Figure 6.10. From Figure 6.11 one can see that there is a good agreement between MC truth, 1-, and 2-leg predictions. This shows that there are no relevant physics correlations between the 2 jets, an essential prerequisite for this background prediction to work well, and proves the p_T and η binning is adequate as well.

Next, the test is performed with reconstructed jets, as shown in the left plot of Figure 6.12, using the exclusive binning. For the MC truth, the ChF cut is applied on the 2 leading jets of

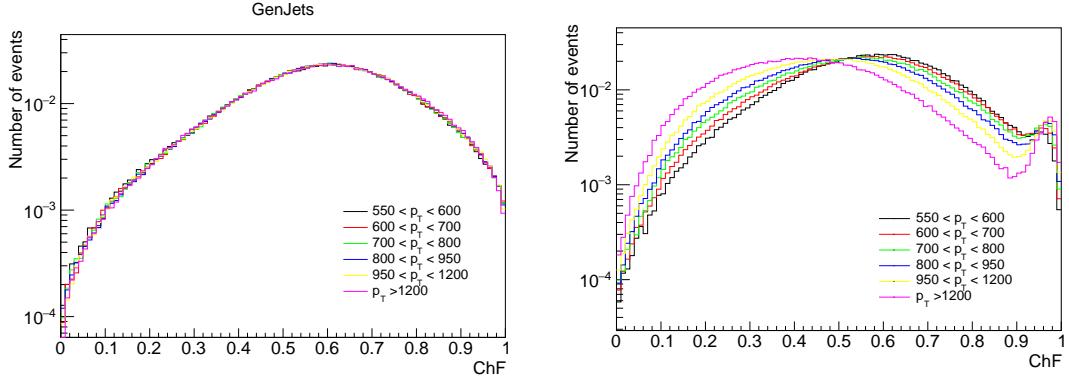


Figure 6.9: The ChF per bins of jet p_T , for generator level (left) and reconstructed (right) jets.

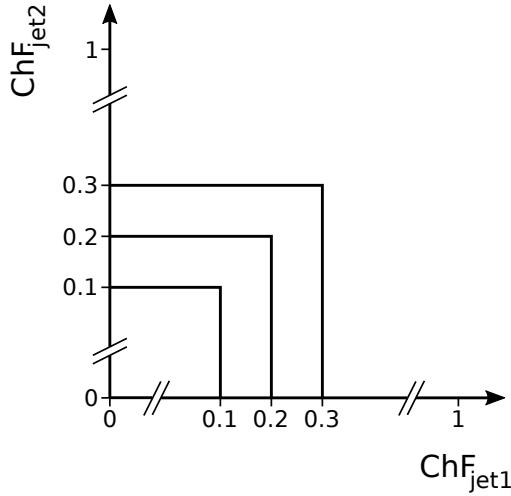


Figure 6.10: Illustration of the exclusive bins in the leading and subleading jet ChF, used for test of the background prediction method and the data vs. prediction comparisons.

the standard jet collection, as well as the 2 leading jets of the jet collection created when using the second vertex as primary vertex. This extra cut is a part of the signal region event selection described earlier, designed to remove events where the wrong primary vertex was chosen and the charged energy fraction of the jets is removed by CHS. However, there is still a small discrepancy between the MC truth and the prediction at the tightest ChF cuts. This is mainly due to a very small number of events where the wrong vertex was chosen, but where the correct one is not the second one. The level of agreement reached is however already sufficient for the considered data sample.

The test of the method is also performed in inclusive bins, as used in the signal region event selection, by applying the same cut on the ChF of both jets. This is shown in the right plot of Figure 6.12, with the applied ChF cut on the x-axis. For the MC truth, the statistical uncertainty is determined per HT-binned QCD sample, using asymmetric vertical bars with correct coverage for event counts with Poisson variates when less than 10 events remain, and the square root of the remaining number of events otherwise. In this way, the statistical uncertainty correctly reflects the contribution of HT bins with few or no events left. The total statistical uncertainty in the MC truth is then calculated by multiplying the uncertainty per HT bin by the corresponding weight for this HT bin, and adding them quadratically. The uncertainties on the 1- and 2-leg predictions are much smaller, as the statistical uncertainties from the efficiencies is negligible. The systematic uncertainty on the background prediction is then defined as the difference between the MC truth and the prediction, unless it is smaller than the statistical uncertainty on the MC

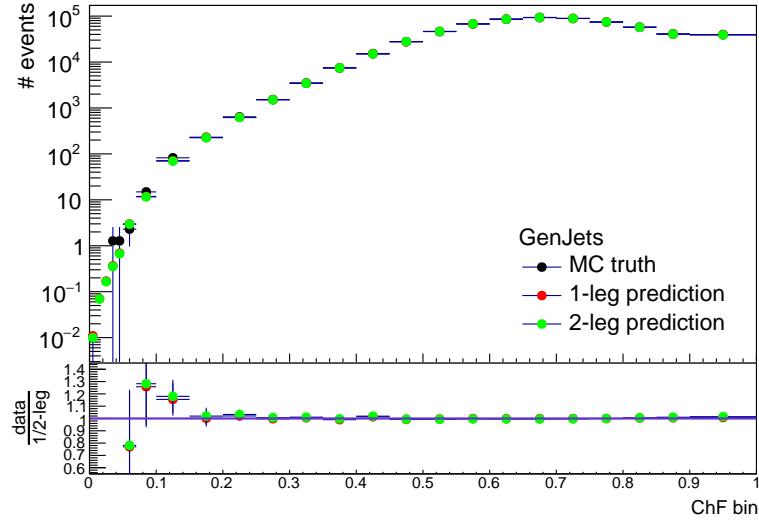


Figure 6.11: Test of the background prediction method using the GenJets in simulated events. The number of events is scaled to 16.1 fb^{-1} , corresponding to the amount of data used for this analysis.

truth. In that case, the uncertainty on the MC truth is taken as systematic uncertainty. For most ChF cut values, the statistical uncertainty dominates. As a result the background prediction is limited by the MC sample size.

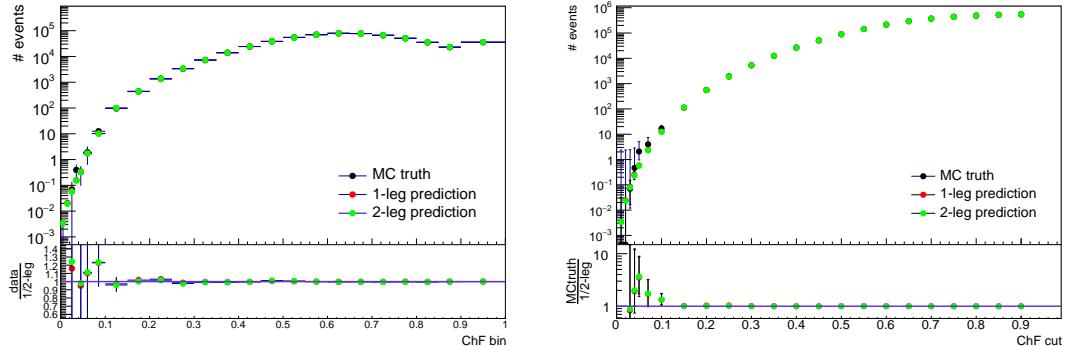


Figure 6.12: Test of the background prediction method with simulated events, using an exclusive (left) or inclusive (right) binning in ChF. The number of events is scaled to 16.1 fb^{-1} , corresponding to the amount of data used for this analysis.

The QCD background prediction in data is shown in Figure 6.13 as a function of the exclusive ChF bins, using the `HLT_PFJet450` single jet trigger. The data is blinded in the signal region below $\text{ChF} = 0.05$. A distinction is made between the run periods B to F (left) and G to H (right). There is a clear deviation below a ChF of 0.4 for run periods B to F, while the agreement is very good down to $\text{ChF} = 0.05$ for run periods G and H. The main difference between these 2 datasets is the issue with the Tracker APV pre-amplifier saturation, described in Section 3.2.6.1, which was solved from run period G onwards.

To first order, the measured ChF cut efficiencies should absorb the effects of tracking inefficiencies caused by the APV pre-amplifier saturation problem. However, this effect is also reflected in the distribution of the number of vertices. As is shown in Figure 6.14, a subtle effect causes the data and the prediction from data to disagree in the distribution of the number of ver-

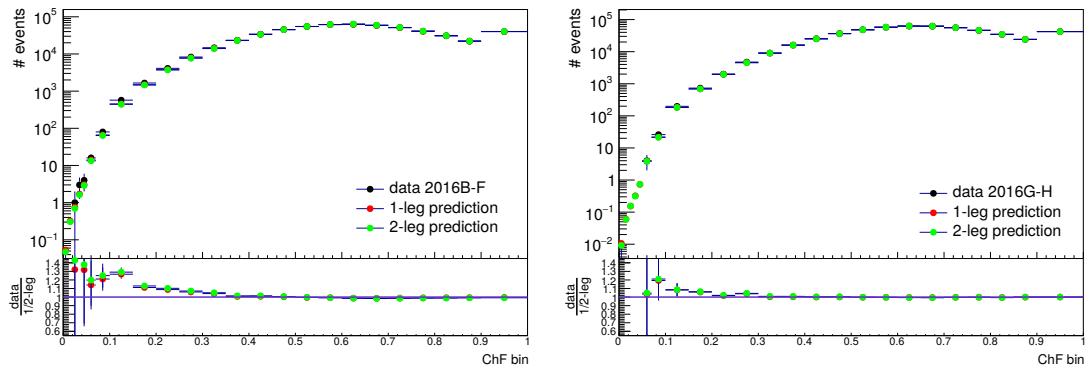


Figure 6.13: The 1- and 2-leg predictions from data, as well as the data as a function of the exclusive ChF bins, for run periods B-F (left) and G-H (right).

tices for run periods B to F, while a good agreement is obtained for run periods G to H. As there are not enough data to derive the ChF efficiencies reliably in bins of number of vertices as well as p_T and η , a recovery of the data in this way is not possible. A reweighting was also performed, using a fit to the ratio of data divided by the 2-leg prediction per ChF bin. Figure 6.15 shows the data versus data prediction comparison in run period E, applying the reweighting based on the number of vertices in the event, per ChF bin, and a clear improvement can be observed for the 2-leg prediction. In contrast, the 1-leg prediction has not been reweighted and still shows the original disagreement. This reweighting can however not be used in the signal region, where very few events remain. As a result, the analysis is performed with run periods G and H only, corresponding to an integrated luminosity of 16.1 fb^{-1} . The rejected data, from run periods B to F, corresponds to an integrated luminosity of 17 fb^{-1} .

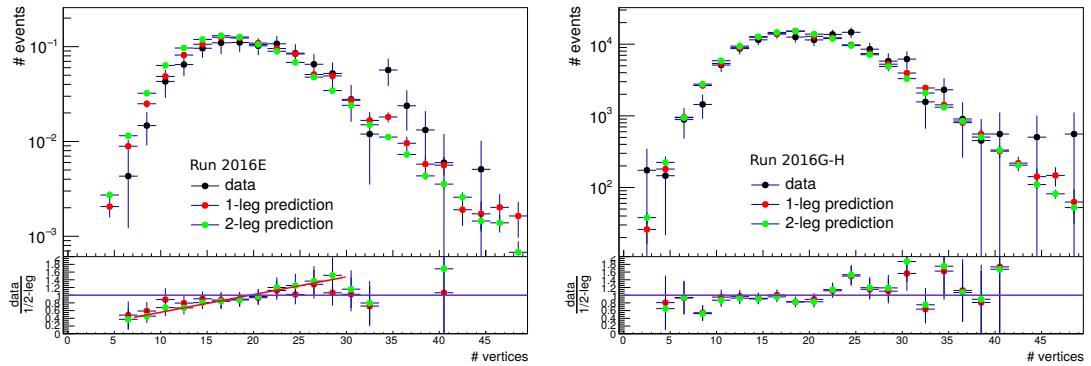


Figure 6.14: The distribution of the number of vertices for data, 1-leg, and 2-leg prediction using data from run period E (left) and run periods G-H (right).

6.5 Systematic uncertainties

For the signal prediction, systematic uncertainties are included for the luminosity, the jet energy corrections, and the trigger inefficiency. The systematic uncertainty for the luminosity amounts to 2.5% [176]. The systematic uncertainty coming from the jet energy corrections is computed by varying the jet energy by the correction and recalculating the yield after applying the selection cuts and the ChF cut. Depending on the SIMP mass and the ChF cut, this uncertainty varies between 0.4% and 3.9%. A systematic uncertainty is also included to take into account the

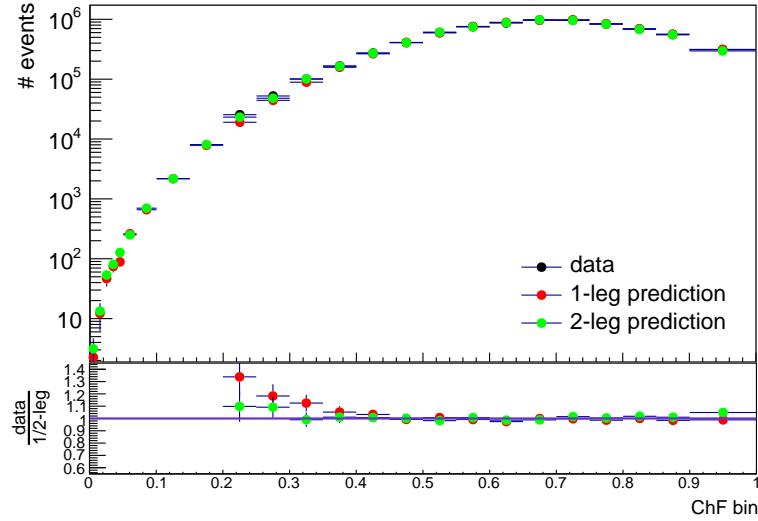


Figure 6.15: The 1- and 2-leg predictions from data, as well as the data (above ChF = 0.2) as a function of the exclusive ChF bins for run period E, reweighting the 2-leg prediction to data based on the number of vertices in the event, per ChF bin.

trigger inefficiency at 550 GeV due to the turn-on. This is done by taking 100% of the maximal inefficiency as the uncertainty, which gives a 2% systematic uncertainty for the signal. This does not take into account the fact that the turn-on was determined for one jet only and the inefficiency is strongly reduced when two jets with a similar p_T are present in the event. However, some signal events have one of the two jets with EMF = 0. In this case the jet which does not contain electromagnetic energy would not fire the single jet trigger, which requires the jet to have EMF > 0, and these events become single jet events from the trigger point of view. The 2% uncertainty is therefore conservative as it represents the worst case scenario. The photon and conversion veto was found to be 100% efficient on the signal, and as a result this systematic uncertainty is negligible. The effect of pileup was considered to be negligible as well, as the distribution of the number of vertices is very similar for the data and SIMP samples. As an example the data is compared to the SIMP sample with $m_\chi = 1000$ GeV in Figure 6.16, which shows that there is a good agreement in the bulk of the distribution with some deviations for a few events with a high number of primary vertices only.

As mentioned in Section 6.4, the main systematic uncertainty on the background prediction is obtained from the closure test in MC, by taking the difference between the MC truth and the prediction, unless it is smaller than the statistical uncertainty on the MC truth, in which case the uncertainty on the MC truth is taken as systematic uncertainty. This uncertainty varies between 4% and 3400%, depending on the ChF cut. As for the signal, the trigger inefficiency due to the turn-on at 550 GeV is also taken into account, yielding an additional 2% systematic uncertainty also for the background.

6.6 Results

Table 6.3 shows the number of predicted and observed events, per considered ChF cut. The prediction is done using the 1-leg data prediction, as this provides slightly smaller uncertainties, as every event can be used twice by applying the ChF first on one jet and then on the other. The statistical uncertainty, as well as the systematic uncertainty from the closure test described in Section 6.5, are given.

A cut of ChF < 0.05 is chosen as the final signal region selection, to reject most of the

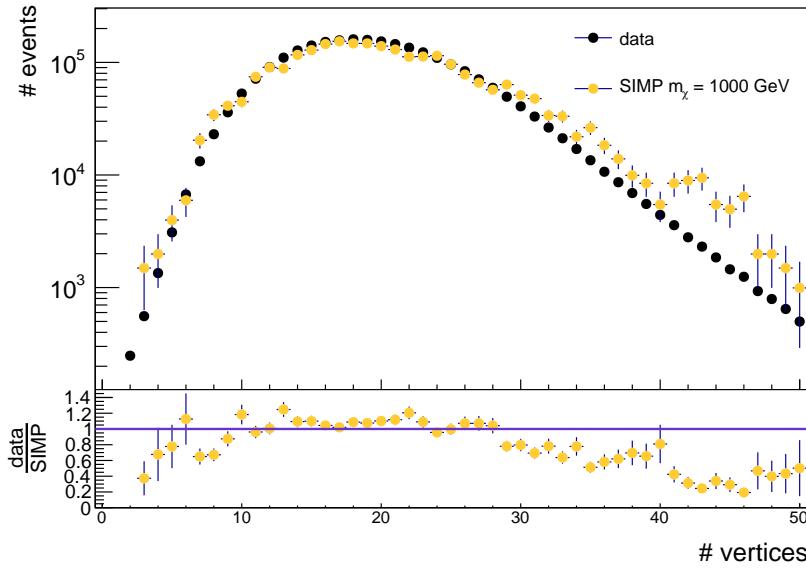


Figure 6.16: The distribution of the number of vertices in data compared to the SIMP signal with $m_\chi = 1000$ GeV.

QCD background. The background is then reduced to the level of about one event, and a large uncertainty does not have big consequences. Moreover, the uncertainty from the closure test is reduced when taking into account the smaller statistical uncertainty on the number of background events during the limit calculation. In addition, the expected sensitivity does not improve significantly at tighter ChF cuts, and the closure tests becomes statistically limited. The dominating uncertainty then comes from the closure test, and amounts to 250% in this case.

ChF cut	data prediction	QCD MC prediction	observed	SIMP signal [m_χ]	
				1 GeV	1000 GeV
0.2	902 ± 5 (stat.) ± 38 (syst.)	546.5 ± 0.6	969	456	4.4
0.15	210 ± 2 (stat.) ± 18 (syst.)	111.1 ± 0.4	229	442	4.3
0.1	26.9 ± 0.3 (stat.) ± 8.9 (syst.)	12.6 ± 0.2	30	426	4.2
0.07	5.1 ± 0.1 (stat.) ± 4.4 (syst.)	2.3 ± 0.2	4	408	4.0
0.05	1.28 ± 0.03 (stat.) $\pm^{3.24}_{-1.28}$ (syst.)	0.6 ± 0.1	0	395	3.9
0.04	0.55 ± 0.02 (stat.) $\pm^{2.81}_{-0.55}$ (syst.)	0.24 ± 0.09	0	387	3.8
0.03	0.22 ± 0.01 (stat.) $\pm^{7.68}_{-0.22}$ (syst.)	0.08 ± 0.07	0	377	3.7

Table 6.3: Number of predicted (using the 1-leg prediction from data) and observed events for several inclusive cuts on the ChF of both jets. The expected number of signal events is also given for the $m_\chi = 1$ GeV and $m_\chi = 1000$ GeV scenarios.

Model-independent limits are derived for a $\text{ChF} < 0.05$ cut, using the CL_S criterion [155, 156] with the LHC style test statistic [162] and using the full set of pseudo-experiments, as described in Section 5.8. This was done using the RooStat-based Combine tool, taking into account the systematic uncertainties detailed in Section 6.5, as well as the statistical uncertainties on signal and background predictions, and modelling them as nuisance parameters. All included systematic uncertainties are profiled with a lognormal prior, except for the uncertainty coming from the closure test, which is profiled with a gamma function since it arises from the limited number of remaining events. The resulting expected fiducial cross section is $\sigma_{\text{fid}}^{95\%} = \sigma \times A \times \epsilon = 0.17 \text{ fb}$. With zero observed events, the observed model-independent lower limit is found to be

$$\sigma_{\text{fid,obs}}^{95\%} = 0.18 \text{ fb.}$$

6.7 SIMP model interpretation

Limits are also derived on the production cross section for the SIMP simplified model, using the same method as described for the model-independent limits. The expected limits on the production cross section are shown for SIMP masses between 1 GeV and 1 TeV in Figure 6.17, using a cut of ChF < 0.05. In this case, the search is sensitive to all the generated SIMP mass points, up to 1 TeV.

Figure 6.18 shows the expected and observed limits when including the observation of zero events in the signal region. The expected and observed limits, and the theoretical cross section are given with respect to the generator level cuts applied in the signal sample generation, $p_T^\chi > 200 \text{ GeV}$ and $|\eta_\chi| < 2.5$. The shown theoretical cross section is also given per SIMP mass point in Table 4.1. In summary, no significant excess above the expected background is observed, and the considered SIMP simplified model is ruled out for SIMP masses up to 1000 GeV.

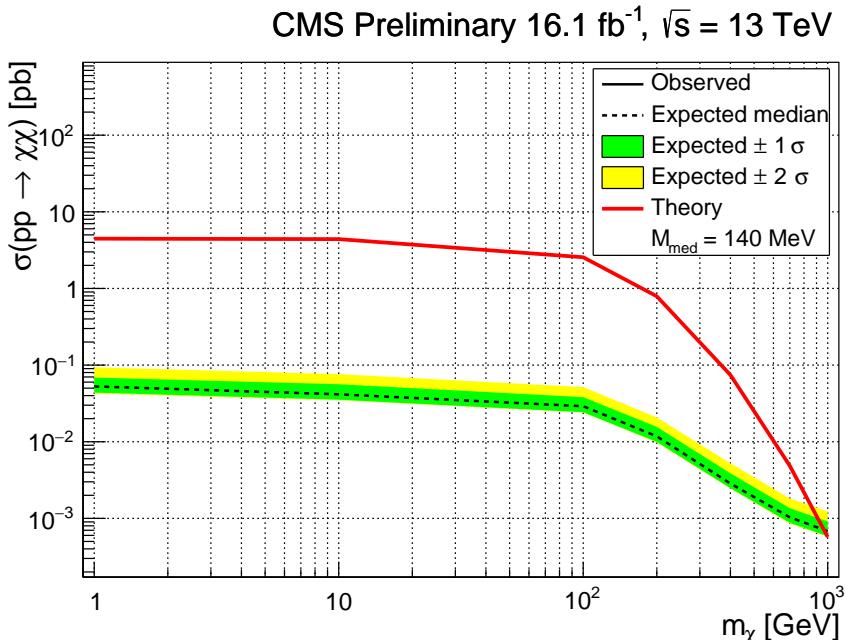


Figure 6.17: The expected limits on the production cross section, obtained without looking in the signal region, for SIMP masses between 1 and 1000 GeV, with 1σ and 2σ bands is shown, as well as the theoretical prediction (red), with respect to the generator level cuts ($p_T^\chi > 200 \text{ GeV}$ and $|\eta_\chi| < 2.5$).

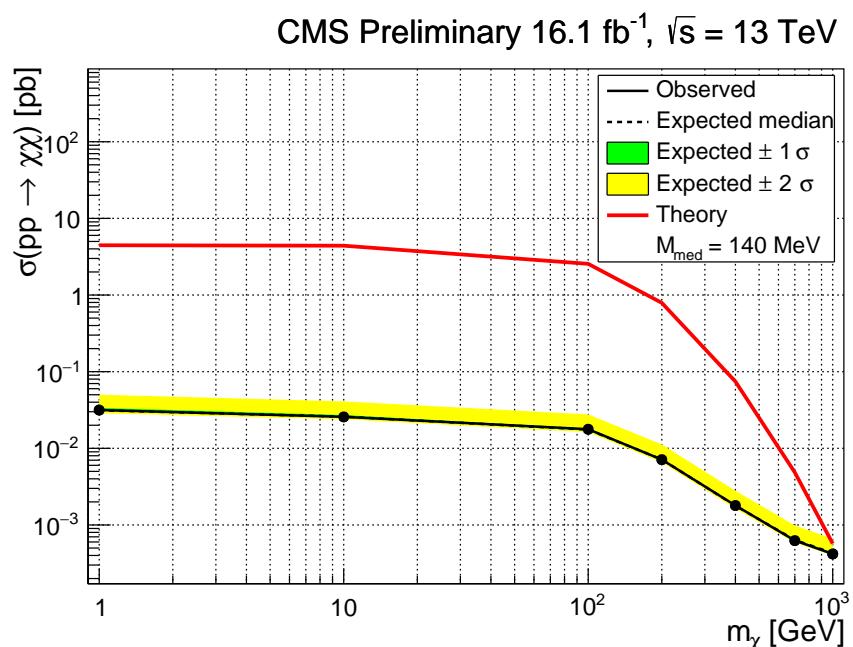


Figure 6.18: The expected and observed limits on the production cross section for SIMP masses between 1 and 1000 GeV, with 1σ and 2σ bands is shown, as well as the theoretical prediction (red), with respect to the generator level cuts ($p_T^\chi > 200$ GeV and $|\eta_\chi| < 2.5$).

7

Conclusions & Outlook

*This is the way the world ends
Not with a bang but a whimper.*

– T.S. Eliot, *The Hollow Men*

One of the remaining unsolved mysteries in the universe is the presumed existence of dark matter. Various astrophysical observations indicate the presence of an unknown type of matter next to the known visible matter. As it does not interact electromagnetically, only very little is known about this new type of matter, and many theories exist to explain its nature and origin. These theories can be tested through a variety of techniques and experiments. Depending on the exact nature of the dark matter particles, they could also be produced at colliders in high-energy collisions between Standard Model particles. Currently, the largest particle accelerator in the world is the LHC at CERN, which provides proton-proton collisions with a record centre-of-mass energy of 13 TeV at a very high rate. In this thesis, two complementary searches for dark matter have been performed with data collected by the CMS detector located at one of the interaction points of this collider.

For both analyses, simplified model interpretations were considered. While dark matter does not interact electromagnetically, in these models it is assumed to interact with the ordinary matter through a new force. In the first analysis, the dark matter candidates are expected to interact very weakly with Standard Model particles and leave the detector undetected. They can however be detected when they recoil against another object in the event. In this case, the studied signature is one or more jets together with missing transverse energy. This search was already performed during Run 1, and has now been improved for Run 2. One of the main developments was the refinement of the prediction of events with a Z boson decaying to two neutrinos in association with one or more jets, which led to an improved sensitivity. While the first iteration of this analysis with Run 2 data sets less stringent limits on the scenario with a vector mediator compared to the Run 1 analysis, the new results including the discussed improvements and using data corresponding to 12.9 fb^{-1} set stronger limits up to mediator masses of 1.95 TeV. Similar results are obtained for an axial-vector mediator. Scalar and pseudoscalar mediator masses are excluded up to 100 GeV and 430 GeV, respectively.

In the second analysis, a different simplified model is considered, where the dark matter particles are assumed to interact strongly with Standard Model particles, through a new, light mediator. The produced pair of stable, neutral dark matter candidates will then give a signature consisting of a pair of trackless jets. First, a feasibility study for this model and corresponding signature was performed and published. For this study, the analysis strategy was developed,

and the achievable sensitivity was investigated. This model had never been tested at colliders before, and as the feasibility study showed promising results, this signature was now studied for the first time at CMS in this analysis, using a type of jets that is typically rejected by the jet identification criteria. While this search is very sensitive to new physics and the expected background was very efficiently reduced, some complications emerged as well. As an example, care was taken to account for the possibility of a wrongly chosen primary vertex, as this problem in the reconstruction can mimic signal-like events. A different issue was related to data taking and originated from the tracker APV pre-amplifier saturation problem, which occurred during the first half of the 2016 data taking period at CMS. Due to the nature of this issue, the data affected by this problem were rejected for the analysis, and data corresponding to a total of 16.1 fb^{-1} were used. Moreover, the photon veto, which is crucial to reject background events coming from the production of a photon and a jet, was extended to take into account photon conversions as well. The outcome of this analysis yielded no observation of new physics, and all the considered dark matter masses, from 1 GeV up to 1 TeV, were excluded. Additionally, model-independent limits were derived as well, excluding production cross sections down to 0.18 fb .

Both results can be translated into limits on the dark matter-nucleon interaction cross section, in order to compare the results with other experiments. Depending on the dark matter mass, the monojet analysis can exclude cross sections between approximately 10^{-6} fb and 0.1 pb , for the scalar or vector mediator case. The limit on the high interaction cross sections arises at small mediator masses, where the used simplified model may be too simple since hadron resonances can change the picture. This region could consequently be further explored as a next step. In comparison, the trackless jets analysis excludes interaction cross sections of about $1 - 10 \text{ mb}$, thus complementing the monojet results at higher cross sections. Additionally, recent results which recast measurements from a low-threshold surface experiment looking for dark matter also exclude SIMPs up to a SIMP-nucleon interaction cross section of approximately 1 mb at SIMP masses above 100 MeV [177].

In the future, the monojet analysis will be able to cover a larger part of phase space, going to higher mediator and dark matter masses, which translates in improvements at small interaction cross sections. Projections can for example be made for the expected sensitivity at the high luminosity LHC, which is expected to deliver 3000 fb^{-1} of integrated luminosity at a centre-of-mass energy of 14 TeV. The expected reach is shown in Figure 7.1 for an axial-vector and pseudoscalar mediator. The nominal scenario assumes that the level of control of the E_T^{miss} distribution will remain the same as in the current analysis. Two more scenarios are shown, where the systematic uncertainties are reduced by a factor 2 and 4.

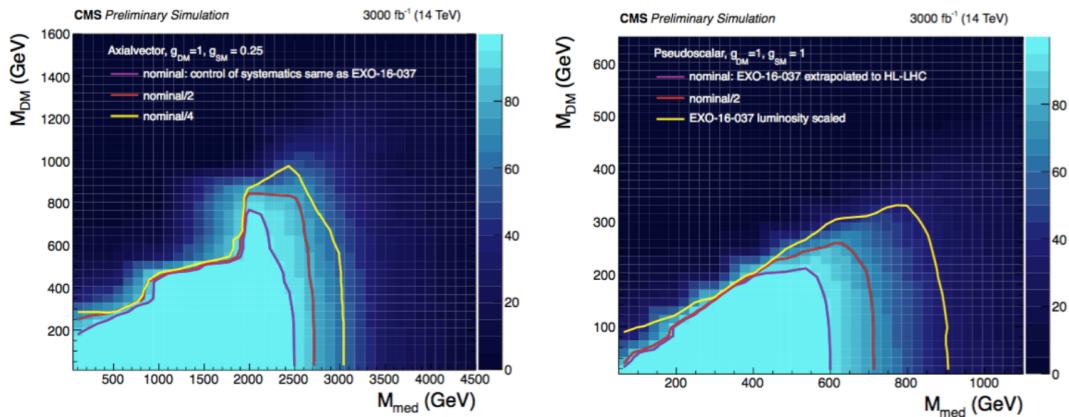


Figure 7.1: Prediction of the sensitivity of the monojet analysis in the $m_{\text{med}} - m_{\text{DM}}$ plane at the high luminosity LHC, using 3000 fb^{-1} of integrated luminosity, for an axial-vector mediator (left) and a pseudoscalar mediator (right). Figures taken from [178].

For the pseudoscalar mediator, the systematic uncertainties are dominated by the uncertainty on lepton identification and isolation efficiencies used in the selection of events in the control regions at low missing transverse momentum. At high missing transverse momentum on the other hand, the systematic uncertainties are dominated by the statistical uncertainty. For the nominal scenario shown in the right plot of Figure 7.1, the systematic uncertainties are scaled by the luminosity at high E_T^{miss} and scaled according to the predictions for the uncertainty on the lepton identification and isolation efficiencies at low E_T^{miss} . Additionally, a scenario is shown where the systematic uncertainties are reduced by a factor 2, as well as a scenario where the systematic uncertainties are scaled by the luminosity over the full E_T^{miss} range.

The obtained estimations show that mediator masses up to at least 2.5 TeV and 600 GeV could be excluded in the axial-vector and pseudoscalar case, respectively. This corresponds to an increase in range of about 30%.

In the case of the trackless jets search, the entire considered dark matter mass range is already excluded by the described results. As the interaction cross section in the model is restricted a narrow range in order to have a trackless jets signature, it was possible to exclude this model instead of excluding it up to some cross section only. In a next step, the mass range could be extended, or the search could be broadened by allowing more extra jets, or including missing transverse energy in the signature. Alternative signatures can also be obtained by assuming a different interaction cross section, such as emerging jets or a cluster of tracks in the muon systems. Aside from SIMPs, Hidden Valley models can also give rise to this kind of signature [179]. In these models, the interaction with the hidden sector can for example happen through rare Higgs boson decays. Additionally, trackless jets could also be produced by dark photons [180]. This model is now being studied as well, and the analysis can benefit from this first trackless jets analysis, among other things concerning the issue of the tracker APV pre-amplifier saturation. In the considered scenario, the neutral dark photon is produced in association with an ordinary photon. The resulting signature is then composed of a photon and a trackless jet with energy deposits in the ECAL or HCAL, or missing transverse energy, depending on the interaction strength.

References

- [1] K. G. Begeman, A. H. Broeils, and R. H. Sanders, “Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics,” *Mon. Not. Roy. Astron. Soc.*, vol. 249, p. 523, 1991.
- [2] V. Belokurov, N. W. Evans, P. C. Hewett, A. Moiseev, R. G. McMahon, S. F. Sanchez, and L. J. King, “Two New Large Separation Gravitational Lenses from SDSS,” *Mon. Not. Roy. Astron. Soc.*, vol. 392, p. 104, 2009.
- [3] G. F. Smoot *et al.*, “Structure in the COBE differential microwave radiometer first year maps,” *Astrophys. J.*, vol. 396, pp. L1–L5, 1992.
- [4] E. Komatsu *et al.*, “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation,” *Astrophys. J. Suppl.*, vol. 192, p. 18, 2011.
- [5] P. A. R. Ade *et al.*, “Planck 2013 results. XVI. Cosmological parameters,” *Astron. Astrophys.*, vol. 571, p. A16, 2014.
- [6] N. Daci, I. De Bruyn, S. Lowette, M. H. G. Tytgat, and B. Zaldivar, “Simplified SIMPs and the LHC,” *JHEP*, vol. 11, p. 108, 2015.
- [7] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory; 1995 ed.* Boulder, CO: Westview, 1995. [Online]. Available: <https://cds.cern.ch/record/257493>
- [8] Y. Nagashima, *Beyond the standard model of elementary particle physics.* Weinheim, Germany: Wiley-VCH Verlag, 2014. [Online]. Available: <http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.N34::2014>
- [9] 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.*, vol. B716, pp. 1–29, 2012.
- [10] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett.*, vol. B716, pp. 30–61, 2012.
- [11] S. L. Glashow, “The renormalizability of vector meson interactions,” *Nucl. Phys.*, vol. 10, pp. 107–117, 1959.
- [12] S. Weinberg, “A Model of Leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.
- [13] A. Salam and J. C. Ward, “Weak and electromagnetic interactions,” *Nuovo Cim.*, vol. 11, pp. 568–577, 1959.
- [14] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.
- [15] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.
- [16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

- [17] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Nonabelian Gauge Theories,” *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973.
- [18] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?” *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, 1973.
- [19] <http://www.quantumdiaries.org/2014/03/14/the-standard-model-a-beautiful-but-flawed-theory>.
- [20] Q. R. Ahmad *et al.*, “Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory,” *Phys. Rev. Lett.*, vol. 89, p. 011301, 2002.
- [21] Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998.
- [22] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints,” *Phys. Rept.*, vol. 405, pp. 279–390, 2005.
- [23] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helv. Phys. Acta*, vol. 6, pp. 110–127, 1933, [Gen. Rel. Grav.41,207(2009)].
- [24] P. Salucci and A. Boriello, “The intriguing distribution of dark matter in galaxies,” *Lect. Notes Phys.*, vol. 616, pp. 66–77, 2003.
- [25] A. A. Penzias and R. W. Wilson, “A Measurement of excess antenna temperature at 4080-Mc/s,” *Astrophys. J.*, vol. 142, pp. 419–421, 1965.
- [26] <https://briankoberlein.com/2015/06/15/science-in-the-raw>.
- [27] P. A. R. Ade *et al.*, “Planck 2013 results. I. Overview of products and scientific results,” *Astron. Astrophys.*, vol. 571, p. A1, 2014.
- [28] M. Azzaro, F. Prada, and C. M. Gutierrez, “Motion properties of satellites around external spiral galaxies,” *ASP Conf. Ser.*, vol. 327, p. 268, 2004.
- [29] M. Mateo, “Dwarf galaxies of the Local Group,” *Ann. Rev. Astron. Astrophys.*, vol. 36, pp. 435–506, 1998.
- [30] J. N. Bahcall, C. Flynn, and A. Gould, “Local dark matter from a carefully selected sample,” *Astrophys. J.*, vol. 389, pp. 234–250, 1992.
- [31] M. Milgrom, “A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis,” *Astrophys. J.*, vol. 270, pp. 365–370, 1983.
- [32] M. Milgrom, “A Modification of the Newtonian dynamics: Implications for galaxies,” *Astrophys. J.*, vol. 270, pp. 371–383, 1983.
- [33] M. Milgrom, “A modification of the Newtonian dynamics: implications for galaxy systems,” *Astrophys. J.*, vol. 270, pp. 384–389, 1983.
- [34] J. D. Bekenstein, “Relativistic gravitation theory for the MOND paradigm,” *Phys. Rev.*, vol. D70, p. 083509, 2004, [Erratum: Phys. Rev.D71,069901(2005)].
- [35] E. P. Verlinde, “Emergent Gravity and the Dark Universe,” *SciPost Phys.*, vol. 2, no. 3, p. 016, 2017.
- [36] L. Bergström, “Nonbaryonic dark matter: Observational evidence and detection methods,” *Rept. Prog. Phys.*, vol. 63, p. 793, 2000.

- [37] V. M. Lobashev, “The search for the neutrino mass by direct method in the tritium beta-decay and perspectives of study it in the project KATRIN,” *Nucl. Phys.*, vol. A719, pp. 153–160, 2003.
- [38] V. N. Aseev *et al.*, “An upper limit on electron antineutrino mass from Troitsk experiment,” *Phys. Rev.*, vol. D84, p. 112003, 2011.
- [39] C. Kraus *et al.*, “Final results from phase II of the Mainz neutrino mass search in tritium beta decay,” *Eur. Phys. J.*, vol. C40, pp. 447–468, 2005.
- [40] M. C. Gonzalez-Garcia and Y. Nir, “Neutrino masses and mixing: Evidence and implications,” *Rev. Mod. Phys.*, vol. 75, pp. 345–402, 2003.
- [41] S. Dodelson and L. M. Widrow, “Sterile-neutrinos as dark matter,” *Phys. Rev. Lett.*, vol. 72, pp. 17–20, 1994.
- [42] N. Yoshida, A. Sokasian, L. Hernquist, and V. Springel, “Early structure formation and reionization in a warm dark matter cosmology,” *Astrophys. J.*, vol. 591, pp. L1–L4, 2003.
- [43] L. J. Rosenberg and K. A. van Bibber, “Searches for invisible axions,” *Phys. Rept.*, vol. 325, pp. 1–39, 2000.
- [44] A. A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, “Where are the missing Galactic satellites?” *Astrophys. J.*, vol. 522, pp. 82–92, 1999.
- [45] B. Moore, S. Ghigna, F. Governato, G. Lake, T. R. Quinn, J. Stadel, and P. Tozzi, “Dark matter substructure within galactic halos,” *Astrophys. J.*, vol. 524, pp. L19–L22, 1999.
- [46] L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, “Ultralight scalars as cosmological dark matter,” *Phys. Rev.*, vol. D95, no. 4, p. 043541, 2017.
- [47] S. Hawking, “Gravitationally collapsed objects of very low mass,” *Mon. Not. Roy. Astron. Soc.*, vol. 152, p. 75, 1971.
- [48] B. P. Abbott *et al.*, “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Phys. Rev. Lett.*, vol. 116, no. 6, p. 061102, 2016.
- [49] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, “Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914,” *Phys. Rev. Lett.*, vol. 117, no. 6, p. 061101, 2016.
- [50] S. Clesse and J. Garcia-Bellido, “The clustering of massive Primordial Black Holes as Dark Matter: measuring their mass distribution with Advanced LIGO,” *Phys. Dark Univ.*, vol. 15, pp. 142–147, 2017.
- [51] S. Bird, I. Cholis, J. B. Munoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Racanelli, and A. G. Riess, “Did LIGO detect dark matter?” *Phys. Rev. Lett.*, vol. 116, no. 20, p. 201301, 2016.
- [52] D. Hooper, “Particle Dark Matter,” in *Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008*, 2010, pp. 709–764. [Online]. Available: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-09-025
- [53] K. Kainulainen and K. A. Olive, “Astrophysical and cosmological constraints on neutrino masses,” *Springer Tracts Mod. Phys.*, vol. 190, pp. 53–74, 2003.
- [54] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “The Hierarchy problem and new dimensions at a millimeter,” *Phys. Lett.*, vol. B429, pp. 263–272, 1998.

- [55] L. Randall and R. Sundrum, “A Large mass hierarchy from a small extra dimension,” *Phys. Rev. Lett.*, vol. 83, pp. 3370–3373, 1999.
- [56] E. W. Kolb, D. J. H. Chung, and A. Riotto, “WIMPzillas!” in *Trends in theoretical physics II. Proceedings, 2nd La Plata Meeting, Buenos Aires, Argentina, November 29-December 4, 1998*, 1998, pp. 91–105, [,91(1998)]. [Online]. Available: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf=98-325
- [57] D. N. Spergel and P. J. Steinhardt, “Observational evidence for selfinteracting cold dark matter,” *Phys. Rev. Lett.*, vol. 84, pp. 3760–3763, 2000.
- [58] A. De Rujula, S. L. Glashow, and U. Sarid, “CHARGED DARK MATTER,” *Nucl. Phys.*, vol. B333, pp. 173–194, 1990.
- [59] A. Kusenko and M. E. Shaposhnikov, “Supersymmetric Q balls as dark matter,” *Phys. Lett.*, vol. B418, pp. 46–54, 1998.
- [60] J. R. Ellis, “Particle candidates for dark matter,” *Phys. Scripta*, vol. T85, pp. 221–230, 2000.
- [61] L. Bergstrom, “Dark Matter Candidates,” *New J. Phys.*, vol. 11, p. 105006, 2009.
- [62] R. Bernabei *et al.*, “Final model independent result of DAMA/LIBRA-phase1,” *Eur. Phys. J.*, vol. C73, p. 2648, 2013.
- [63] R. Agnese *et al.*, “Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS,” *Phys. Rev. Lett.*, vol. 112, no. 24, p. 241302, 2014.
- [64] E. Armengaud *et al.*, “Constraints on low-mass WIMPs from the EDELWEISS-III dark matter search,” *JCAP*, vol. 1605, no. 05, p. 019, 2016.
- [65] G. Angloher *et al.*, “Results on low mass WIMPs using an upgraded CRESST-II detector,” *Eur. Phys. J.*, vol. C74, no. 12, p. 3184, 2014.
- [66] E. Aprile *et al.*, “XENON100 Dark Matter Results from a Combination of 477 Live Days,” *Phys. Rev.*, vol. D94, no. 12, p. 122001, 2016.
- [67] D. S. Akerib *et al.*, “Results from a search for dark matter in the complete LUX exposure,” *Phys. Rev. Lett.*, vol. 118, no. 2, p. 021303, 2017.
- [68] C. Amole *et al.*, “Dark Matter Search Results from the PICO-60 C₃F₈ Bubble Chamber,” *Phys. Rev. Lett.*, vol. 118, no. 25, p. 251301, 2017.
- [69] C. Fu *et al.*, “Spin-Dependent Weakly-Interacting-Massive-ParticleNucleon Cross Section Limits from First Data of PandaX-II Experiment,” *Phys. Rev. Lett.*, vol. 118, no. 7, p. 071301, 2017.
- [70] T. Marrodn Undagoitia and L. Rauch, “Dark matter direct-detection experiments,” *J. Phys.*, vol. G43, no. 1, p. 013001, 2016.
- [71] W. B. Atwood *et al.*, “The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission,” *Astrophys. J.*, vol. 697, pp. 1071–1102, 2009.
- [72] F. Aharonian *et al.*, “Observations of the Crab Nebula with H.E.S.S.,” *Astron. Astrophys.*, vol. 457, pp. 899–915, 2006.
- [73] J. Aleksic *et al.*, “Performance of the MAGIC stereo system obtained with Crab Nebula data,” *Astropart. Phys.*, vol. 35, pp. 435–448, 2012.

- [74] J. Holder *et al.*, “Status of the VERITAS Observatory,” *AIP Conf. Proc.*, vol. 1085, pp. 657–660, 2009.
- [75] M. Ageron *et al.*, “ANTARES: the first undersea neutrino telescope,” *Nucl. Instrum. Meth.*, vol. A656, pp. 11–38, 2011.
- [76] A. Achterberg *et al.*, “First Year Performance of The IceCube Neutrino Telescope,” *Astropart. Phys.*, vol. 26, pp. 155–173, 2006.
- [77] A. Kounine, “The Alpha Magnetic Spectrometer on the International Space Station,” *Int. J. Mod. Phys.*, vol. E21, no. 08, p. 1230005, 2012.
- [78] P. Picozza *et al.*, “PAMELA: A Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics,” *Astropart. Phys.*, vol. 27, pp. 296–315, 2007.
- [79] A. M. Sirunyan *et al.*, “Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 07, p. 014, 2017.
- [80] M. Aaboud *et al.*, “Search for dark matter produced in association with a hadronically decaying vector boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Lett.*, vol. B763, pp. 251–268, 2016.
- [81] F. Kahlhoefer, “Review of LHC Dark Matter Searches,” *Int. J. Mod. Phys.*, vol. A32, no. 13, p. 1730006, 2017.
- [82] O. Buchmueller, C. Doglioni, and L. T. Wang, “Search for dark matter at colliders,” *Nature Phys.*, vol. 13, no. 3, pp. 217–223, 2017.
- [83] ATLAS collaboration, “Search for long-lived, massive particles in events with displaced vertices and missing transverse momentum in 13 TeV pp collisions with the ATLAS detector,” 2017.
- [84] ATLAS collaboration, “Search for long-lived charginos based on a disappearing-track signature in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” 2017.
- [85] ATLAS collaboration, “Search for Higgs boson decays to Beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at $\sqrt{s} = 13$ TeV,” 2017.
- [86] A. M. Sirunyan *et al.*, “Search for new physics with dijet angular distributions in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 07, p. 013, 2017.
- [87] A. M. Sirunyan *et al.*, “Search for dijet resonances in protonproton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter and other models,” *Phys. Lett.*, vol. B769, pp. 520–542, 2017, [Erratum: *Phys. Lett.*B772,882(2017)].
- [88] M. Aaboud *et al.*, “Search for new phenomena in dijet events using 37 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Rev.*, vol. D96, p. 052004, 2017.
- [89] Y. Bai and A. Rajaraman, “Dark Matter Jets at the LHC,” 2011.
- [90] CMS Collaboration, “Search for heavy stable charged particles with 12.9 fb^{-1} of 2016 data,” 2016.
- [91] M. Aaboud *et al.*, “Search for metastable heavy charged particles with large ionization energy loss in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS experiment,” *Phys. Rev.*, vol. D93, no. 11, p. 112015, 2016.
- [92] <http://cms-results.web.cern.ch/cms-results/public-results/publications/EXO/index.html>.

- [93] G. Busoni, A. De Simone, E. Morgante, and A. Riotto, “On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC,” *Phys. Lett.*, vol. B728, pp. 412–421, 2014.
- [94] G. Busoni, A. De Simone, J. Gramling, E. Morgante, and A. Riotto, “On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC, Part II: Complete Analysis for the s -channel,” *JCAP*, vol. 1406, p. 060, 2014.
- [95] G. Busoni, A. De Simone, T. Jacques, E. Morgante, and A. Riotto, “On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC Part III: Analysis for the t -channel,” *JCAP*, vol. 1409, p. 022, 2014.
- [96] D. Abercrombie *et al.*, “Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum,” 2015.
- [97] G. Busoni *et al.*, “Recommendations on presenting LHC searches for missing transverse energy signals using simplified s -channel models of dark matter,” 2016.
- [98] J. S. Bullock, “Notes on the Missing Satellites Problem,” 2010.
- [99] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, “Too big to fail? The puzzling darkness of massive Milky Way subhaloes,” *Mon. Not. Roy. Astron. Soc.*, vol. 415, p. L40, 2011.
- [100] D. H. Weinberg, J. S. Bullock, F. Governato, R. Kuzio de Naray, and A. H. G. Peter, “Cold dark matter: controversies on small scales,” *Proc. Nat. Acad. Sci.*, vol. 112, pp. 12 249–12 255, 2014, [Proc. Nat. Acad. Sci.112,2249(2015)].
- [101] B. Famaey and S. McGaugh, “Challenges for Lambda-CDM and MOND,” *J. Phys. Conf. Ser.*, vol. 437, p. 012001, 2013.
- [102] J. F. Donoghue, E. Golowich, and B. R. Holstein, “Dynamics of the standard model,” *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 2, pp. 1–540, 1992, [Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.35(2014)].
- [103] C. Downum, T. Barnes, J. R. Stone, and E. S. Swanson, “Nucleon-meson coupling constants and form-factors in the quark model,” *Phys. Lett.*, vol. B638, pp. 455–460, 2006.
- [104] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac, “Constraints on the Self-Interaction Cross-Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56,” *Astrophys. J.*, vol. 679, pp. 1173–1180, 2008.
- [105] J. L. Feng, “Dark Matter Candidates from Particle Physics and Methods of Detection,” *Ann. Rev. Astron. Astrophys.*, vol. 48, pp. 495–545, 2010.
- [106] G. D. Mack, J. F. Beacom, and G. Bertone, “Towards Closing the Window on Strongly Interacting Dark Matter: Far-Reaching Constraints from Earth’s Heat Flow,” *Phys. Rev.*, vol. D76, p. 043523, 2007.
- [107] J. Rich, R. Rocchia, and M. Spiro, “A Search for Strongly Interacting Dark Matter,” *Phys. Lett.*, vol. B194, p. 173, 1987, [,221(1987)].
- [108] A. L. Erickcek, P. J. Steinhardt, D. McCammon, and P. C. McGuire, “Constraints on the Interactions between Dark Matter and Baryons from the X-ray Quantum Calorimetry Experiment,” *Phys. Rev.*, vol. D76, p. 042007, 2007.
- [109] G. D. Mack and A. Manohar, “Closing the window on high-mass strongly interacting dark matter,” *J. Phys.*, vol. G40, p. 115202, 2013.

- [110] R. H. Cyburt, B. D. Fields, V. Pavlidou, and B. D. Wandelt, “Constraining strong baryon-dark matter interactions with primordial nucleosynthesis and cosmic rays,” *Phys. Rev.*, vol. D65, p. 123503, 2002.
- [111] X.-l. Chen, S. Hannestad, and R. J. Scherrer, “Cosmic microwave background and large scale structure limits on the interaction between dark matter and baryons,” *Phys. Rev.*, vol. D65, p. 123515, 2002.
- [112] C. Dvorkin, K. Blum, and M. Kamionkowski, “Constraining Dark Matter-Baryon Scattering with Linear Cosmology,” *Phys. Rev.*, vol. D89, no. 2, p. 023519, 2014.
- [113] H. R. Gustafson, C. A. Ayre, L. W. Jones, M. J. Longo, and P. V. Ramana Murthy, “Search for New Massive Long-lived Neutral Particles,” *Phys. Rev. Lett.*, vol. 37, p. 474, 1976.
- [114] K. A. Olive *et al.*, “Review of Particle Physics,” *Chin. Phys.*, vol. C38, p. 090001, 2014.
- [115] E. A. Mobs, “The CERN accelerator complex. Complexe des accélérateurs du CERN,” Oct 2016, general Photo. [Online]. Available: <http://cds.cern.ch/record/2225847>
- [116] S. Chatrchyan *et al.*, “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.
- [117] D. A. Matzner Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung *et al.*, “CMS Technical Design Report for the Pixel Detector Upgrade,” 2012.
- [118] G. L. Bayatian *et al.*, “CMS Physics,” 2006.
- [119] S. Chatrchyan *et al.*, “Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 8, p. P09009, 2013, [JINST8,9009(2013)].
- [120] G. Abbiendi, “The CMS muon system in Run2: preparation, status and first results,” *PoS*, vol. EPS-HEP2015, p. 237, 2015.
- [121] A. Tapper and D. Acosta, “CMS Technical Design Report for the Level-1 Trigger Upgrade,” 2013.
- [122] J.-M. O. Andre *et al.*, “Performance of the CMS Event Builder,” no. CMS-CR-2017-034, Feb 2017. [Online]. Available: <http://cds.cern.ch/record/2252634>
- [123] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/DataQuality>.
- [124] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/ActiveChannelsSummary>.
- [125] E. Butz, private communication. Shown at 127th LHCC Open Session, https://indico.cern.ch/event/563488/contributions/2277293/attachments/1340356/2018262/CMS_LHCC_2016_09_21.pdf.
- [126] S. Höche, “Introduction to parton-shower event generators,” in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, 2015, pp. 235–295. [Online]. Available: <http://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf>
- [127] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, “New generation of parton distributions with uncertainties from global QCD analysis,” *JHEP*, vol. 07, p. 012, 2002.

- [128] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC,” *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009.
- [129] R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP*, vol. 04, p. 040, 2015.
- [130] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP*, vol. 07, p. 079, 2014.
- [131] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,” *JHEP*, vol. 11, p. 070, 2007.
- [132] V. N. Gribov and L. N. Lipatov, “Deep inelastic e p scattering in perturbation theory,” *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972, [Yad. Fiz. 15, 781(1972)].
- [133] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.” *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977, [Zh. Eksp. Teor. Fiz. 73, 1216(1977)].
- [134] G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language,” *Nucl. Phys.*, vol. B126, pp. 298–318, 1977.
- [135] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual,” *JHEP*, vol. 05, p. 026, 2006.
- [136] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics,” *Phys. Rept.*, vol. 97, pp. 31–145, 1983.
- [137] J. M. Campbell and R. K. Ellis, “MCFM for the Tevatron and the LHC,” *Nucl. Phys. Proc. Suppl.*, vol. 205-206, pp. 10–15, 2010.
- [138] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, “FeynRules 2.0 - A complete toolbox for tree-level phenomenology,” *Comput.Phys.Commun.*, vol. 185, pp. 2250–2300, 2014.
- [139] J. Allison *et al.*, “Geant4 developments and applications,” *IEEE Trans. Nucl. Sci.*, vol. 53, p. 270, 2006.
- [140] S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST*, vol. 9, no. 10, p. P10009, 2014.
- [141] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, “Technical Proposal for the Phase-II Upgrade of the CMS Detector,” 2015.
- [142] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting,” *J. Phys.*, vol. G34, p. N343, 2007.
- [143] A. Strandlie and R. Fruhwirth, “Discrimination between different types of material in track reconstruction with a Gaussian-sum filter,” *IEEE Trans. Nucl. Sci.*, vol. 53, pp. 3842–3849, 2006.
- [144] A. M. Sirunyan *et al.*, “Particle-flow reconstruction and global event description with the cms detector,” *JINST*, vol. 12, p. P10003, 2017.
- [145] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k t jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008. [Online]. Available: <http://stacks.iop.org/1126-6708/2008/i=04/a=063>

- [146] S. Chatrchyan *et al.*, “Identification of b-quark jets with the CMS experiment,” *JINST*, vol. 8, p. P04013, 2013.
- [147] V. Khachatryan *et al.*, “Reconstruction and identification of lepton decays to hadrons and at CMS,” *JINST*, vol. 11, no. 01, p. P01019, 2016.
- [148] B. Zaldivar, private communication.
- [149] V. Khachatryan *et al.*, “Search for dark matter, extra dimensions, and unparticles in mono-jet events in protonproton collisions at $\sqrt{s} = 8$ TeV,” *Eur. Phys. J.*, vol. C75, no. 5, p. 235, 2015.
- [150] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, “Electroweak corrections to hadronic photon production at large transverse momenta,” *JHEP*, vol. 03, p. 059, 2006.
- [151] S. Kallweit, J. M. Lindert, S. Pozzorini, M. Schönherr, and P. Maierhöfer, “NLO QCD+EW automation and precise predictions for V+multijet production,” in *Proceedings, 50th Rencontres de Moriond, QCD and high energy interactions: La Thuile, Italy, March 21-28, 2015*, 2015, pp. 121–124. [Online]. Available: <https://inspirehep.net/record/1372103/files/arXiv:1505.05704.pdf>
- [152] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, “NLO electroweak automation and precise predictions for W+multijet production at the LHC,” *JHEP*, vol. 04, p. 012, 2015.
- [153] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini, and M. Schönherr, “NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging,” *JHEP*, vol. 04, p. 021, 2016.
- [154] CMS Collaboration, “CMS Luminosity Measurement for the 2015 Data Taking Period,” 2016.
- [155] A. L. Read, “Presentation of search results: the cls technique,” *J. Phys.*, vol. G28, p. 2693, 2002.
- [156] T. Junk, “Confidence level computation for combining searches with small statistics,” *Nucl. Instrum. Meth.*, vol. A434, p. 435, 1999.
- [157] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, “The RooStats Project,” *PoS*, vol. ACAT2010, p. 057, 2010.
- [158] Higgs working group, “Documentation of the roostats-based statistics tools for the higgs pag,” <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- [159] CMS Collaboration, “Search for dark matter with jets and missing transverse energy at 13 TeV,” 2015.
- [160] CMS Collaboration, “Search for dark matter production in association with jets, or hadronically decaying W or Z boson at $\sqrt{s} = 13$ TeV,” 2016.
- [161] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, 2011, [Erratum: *Eur. Phys. J.*C73,2501(2013)].
- [162] ATLAS, CMS, LHC Higgs Combination Group Collaborations, “Procedure for the LHC Higgs boson search combination in summer 2011,” 2011.

- [163] O. Buchmueller, M. J. Dolan, S. A. Malik, and C. McCabe, “Characterising dark matter searches at colliders and direct detection experiments: Vector mediators,” *JHEP*, vol. 01, p. 037, 2015.
- [164] A. Kurylov and M. Kamionkowski, “Generalized analysis of weakly interacting massive particle searches,” *Phys. Rev.*, vol. D69, p. 063503, 2004.
- [165] J. Hisano, K. Ishiwata, and N. Nagata, “Gluon contribution to the dark matter direct detection,” *Phys. Rev.*, vol. D82, p. 115007, 2010.
- [166] R. Agnese *et al.*, “New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment,” *Phys. Rev. Lett.*, vol. 116, no. 7, p. 071301, 2016.
- [167] D. S. Akerib *et al.*, “Improved Limits on Scattering of Weakly Interacting Massive Particles from Reanalysis of 2013 LUX Data,” *Phys. Rev. Lett.*, vol. 116, no. 16, p. 161301, 2016.
- [168] G. Angloher *et al.*, “Results on light dark matter particles with a low-threshold CRESST-II detector,” *Eur. Phys. J.*, vol. C76, no. 1, p. 25, 2016.
- [169] A. Tan *et al.*, “Dark Matter Results from First 98.7 Days of Data from the PandaX-II Experiment,” *Phys. Rev. Lett.*, vol. 117, no. 12, p. 121303, 2016.
- [170] C. Amole *et al.*, “Improved dark matter search results from PICO-2L Run 2,” *Phys. Rev.*, vol. D93, no. 6, p. 061101, 2016.
- [171] C. Amole *et al.*, “Dark matter search results from the PICO-60 CF₃I bubble chamber,” *Phys. Rev.*, vol. D93, no. 5, p. 052014, 2016.
- [172] M. G. Aartsen *et al.*, “Improved limits on dark matter annihilation in the Sun with the 79-string IceCube detector and implications for supersymmetry,” *JCAP*, vol. 1604, no. 04, p. 022, 2016.
- [173] K. Choi *et al.*, “Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande,” *Phys. Rev. Lett.*, vol. 114, no. 14, p. 141301, 2015.
- [174] M. Ackermann *et al.*, “Constraining Dark Matter Models from a Combined Analysis of Milky Way Satellites with the Fermi Large Area Telescope,” *Phys. Rev. Lett.*, vol. 107, p. 241302, 2011.
- [175] A. A. Abdo *et al.*, “Observations of Milky Way Dwarf Spheroidal galaxies with the Fermi-LAT detector and constraints on Dark Matter models,” *Astrophys. J.*, vol. 712, pp. 147–158, 2010.
- [176] A. Kornmayer, “The CMS Pixel Luminosity Telescope,” no. CMS-CR-2015-121, Jun 2015. [Online]. Available: <https://cds.cern.ch/record/2039978>
- [177] J. H. Davis, “Probing sub-GeV mass SIMP dark matter with a low-threshold surface experiment,” 2017.
- [178] “Updates on Projections of Physics Reach with the Upgraded CMS Detector for High Luminosity LHC,” Oct 2016. [Online]. Available: <http://cds.cern.ch/record/2221747>
- [179] M. J. Strassler and K. M. Zurek, “Echoes of a hidden valley at hadron colliders,” *Phys. Lett.*, vol. B651, pp. 374–379, 2007.
- [180] E. Izaguirre and I. Yavin, “New window to millicharged particles at the LHC,” *Phys. Rev.*, vol. D92, no. 3, p. 035014, 2015.

List of Acronyms

A

ALICE A Large Ion Collider Experiment
ATLAS A Toroidal LHC ApparatuS

B

BU Builder Unit

C

CERN European Organization for Nuclear Research
CHS charged hadron subtraction
CMB Cosmic Microwave Background
CMS Compact Muon Solenoid
CSC Cathode Strip Chamber

D

DAQ data acquisition
DQM Data Quality Monitoring
DT Drift Tube

E

ECAL electromagnetic calorimeter

F

FED Front End Driver
FSR final state radiation
FU Filter Unit

G

GSF Gaussian-sum filter

H

HCAL hadronic calorimeter
HLT High-Level Trigger
HSCP heavy stable charged particle

I

IP	interaction point
ISR	initial state radiation

J

JER	jet energy resolution
-----	-----------------------

L

L1	Level-1
LEIR	Low Energy Ion Ring
LEP	Large Electron Positron
LHC	Large Hadron Collider
LHCb	LHC Beauty
LO	leading order

N

NLO	next-to-leading order
-----	-----------------------

P

PDF	parton distribution function
PF	particle flow
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster

Q

QCD	Quantum Chromodynamics
-----	------------------------

R

RF	radio-frequency
ROC	read-out chip
RPC	Resistive Plate Chamber
RU	Readout Unit

S

SIDM	self-interacting dark matter
SIMP	strongly interacting massive particle
SPS	Super Proton Synchrotron
SUSY	supersymmetry

T

TEC	Tracker EndCaps
-----	-----------------

TIB Tracker Inner Barrel
TID Tracker Inner Disks
TOB Tracker Outer Barrel

W

WIMP weakly interacting massive particle

