Searching for Dark Matter produced in association with top quarks with the CMS experiment at the LHC

Dissertation

zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Physik

der Universität Hamburg

vorgelegt von

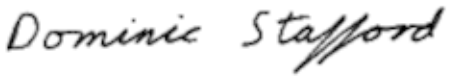
Dominic William Stafford

Hamburg
2023

Gutachter/innen der Dissertation:	Prof. Dr. Christian Schwanenberger Prof. Dr. Gudrid Moortgat-Pick Prof. Dr. Björn Penning
Zusammensetzung der Prüfungskommission:	Prof. Dr. Jochen Liske Prof. Dr. Christian Schwanenberger Dr. Alexander Grohsjean Prof. Dr. Gregor Kasieczka Prof. Dr. Gudrid Moortgat-Pick
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Jochen Liske
Datum der Disputation:	19.12.23
Vorsitzender des Fach-Promotionsausschusses PHYSIK:	Prof. Dr. Günter H. W. Sigl
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Wolfgang J. Parak
Dekan der Fakultät MIN:	Prof. Dr.-Ing. Norbert Ritter

Declaration on oath

I hereby declare in lieu of oath that I have written this dissertation myself and that I have not used any auxiliary materials or sources other than those indicated.



Hamburg, 1st September 2023

Acknowledgements

During my time as a PhD student I have been fortunate to work with many talented scientists and kind people, and in this section I will attempt to recognise and thank all those who have made a significant contribution. First and foremost I would like to thank my supervisors, Alexander Grohsjean and Christian Schwanenberger; their assistance encompassed not only numerous physics insights, but also supporting me in my return to the UK during the early months of the Covid pandemic and advice on how to navigate the academic world, from my first talks to future career advice. Next thanks must go to Danyer Perez Adan, who was a model postdoctoral researcher supporting a PhD student, always being available for advice or discussion, thoroughly checking my work and identifying potential problems. My collaborator in designing the "Pepper" analysis framework was Jonas Rübenach, whose software expertise and rigorous coding practices were instrumental in making the framework what it is today. Laurids Jeppe not only optimised the selection of the $t\bar{t}Z$ control region described in section 7.2, but was a helpful office mate, providing a sounding board for ideas and assistance on catching bugs. I would also like to thank Nicole Stefanov, who explored the use of different kinematic reconstruction algorithms for the $t\bar{t}+\text{DM}$ signal, and provided numerous insights when starting this work; Afiq Anuar, who provided helpful advice on a number of topics, must notably the smoothing of systematic uncertainties, and Andrej Saibel, who introduced me to the role of Herwig contact within CMS.

From outside of the DESY group I would like to thank everyone who worked on the search for dark matter in association with top quarks, especially Cedric Prieels, Deborah Pinna and Victor Shang.

I would also like to thank Simon Plätzer for offering me the opportunity to work on implementing dark showers into Herwig, supervising this project and providing many insights into both the code and the physics it simulates. This project also gave me the opportunity to work with Suchita Kulkarni, who brought a lot of knowledge of dark shower phenomenology. Additionally I would like to acknowledge the contributions of Aidin Masouminia, Andreas Papaefstathiou and Andrzej Siódmok, who implemented the parton shower element of this model.

Finally, I would like to thank everyone in the DESY CMS group, who together make a supportive and engaging working community, and my friends who have made the past four years enjoyable and enriching, in particular Mikel Mendizabal Morentin, for being there when needed. The very last word must go to my parents, Caroline Greaves and Brendan Stafford, whose support guidance and wisdom has been unfailing; thank you.

Abstract

A search is presented for dark matter (DM) produced in association with top quarks in final states with two leptons with data collected by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider between 2016 and 2018. This search targets the single top quark + DM (t + DM) signal in addition to the top quark pair + DM ($t\bar{t}$ + DM) process for the first time in the dileptonic channel. This analysis includes numerous improvements with respect to the previous iteration performed using the 2016 CMS dataset. A new kinematic reconstruction algorithm is used to reconstruct the $t\bar{t}$ pair in $t\bar{t}$ + DM events, which is challenging since both the DM and the neutrinos from the top decays are not observed. Neural networks trained on a number of sensitive variables are used to improve the discrimination of signal from background, and new control regions are introduced to estimate the rate of the $t\bar{t}Z$ and Drell-Yan backgrounds. These improvements increase the sensitivity of this search by 40-65% (depending on the DM mediator mass considered) for the 2016 dataset; also including the 2017 and 2018 data gives a further improvement of 40%.

Limits are set on models where the DM couples to top quarks via a scalar mediator, excluding mediator masses up to 150 GeV (230 GeV expected), and on models with a pseudoscalar mediator, excluding mediator masses up to 150 GeV (260 GeV expected), for the signal cross sections for the benchmark Yukawa-like couplings. This search will be combined with corresponding searches in 0 and 1 lepton final states from CMS to produce a result that will be the most sensitive search for this model published to date.

Lastly, the implementation and validation of a new "dark shower" model in the Herwig Monte Carlo generator is presented. This is an alternative type of dark matter model in which the DM is part of a "dark sector" with strong interactions between the particles. This leads to showers of dark particles which form bound states, some of which are stable and provide the DM candidate, and others of which are unstable and decay to Standard Model (SM) particles. This can lead to interesting experimental signatures of jets with a different substructure to the SM backgrounds, and the sensitivity of some new jet substructure variables to this model are explored.

Zusammenfassung

Eine Suche für dunkle Materie (DM), produziert zusammen mit top Quarks in Endzuständen mit zwei Leptonen, mit Daten gesammelt am Compact Muon Solenoid (CMS) Experiment am Large Hadron Collider von 2016 bis 2018 wird präsentiert. Diese Suche visiert zum ersten Mal im dileptonischen Kanal das einzelne top Quark + DM ($t + \mathrm{DM}$) Signal zusätzlich zum top Quark Paar + DM ($t\bar{t} + \mathrm{DM}$) Prozess an. Diese Analyse enthält mehrere Verbesserungen im Vergleich zum vorläufigen Durchlauf dieser Analyse, die den 2016 CMS Datensatz benutzt hat. Ein neuer Kinematische-Rekonstruktion-Algorithmus wird benutzt, um das $t\bar{t}$ Paar in $t\bar{t} + \mathrm{DM}$ Events zu rekonstruieren, was fordernd ist, weil sowohl die DM als auch die im Zerfall der top Quarks produzierten Neutrinos nicht beobachtet werden. Neuronale Netze, trainiert auf mehreren sensiblen Variablen, werden benutzt, um die Diskrimination des Signals gegen den Hintergrund zu verbessern, und neue Kontrollregionen werden definiert, um die Rate der $t\bar{t}Z$ und Drell-Yan Prozesse abzuschätzen. Diese Verbesserungen erhöhen die Sensitivität dieser Suche um 65-80% (abhängig von der Masse des DM Mediators) für den 2016 CMS Datensatz; hinzufügen der 2017 und 2018 Daten führt zu einer weiteren Verbesserung von 40%.

Wegen dieser Verbesserungen und dem größeren Datensatz, ist die Sensitivität diese Suche 65-80% (abhängig von der Masse des DM Mediator) höher als beim vorläufigen Durchlauf dieser Analyse, die den 2016 CMS Datensatz benutzt hat.

Obergrenzen werden für Modelle, in dem die DM an top Quarks über ein skalaren Mediator koppelt, und für Modelle mit ein pseudoskalaren Mediator gesetzt; skalare Mediator-Massen bis zu 150 GeV (230 GeV erwartet) und pseudoskalare Mediator-Massen bis zu 150 GeV (260 GeV erwartet) werden für Signal-Querschnitte mit Benchmark-Yukawa-Artigen Koplungen ausgeschlossen . Diese Suche wird mit den entspechenden Suchen in den 0- und 1- Leptonen-Endzuständen kombiniert werden, um ein Ergebnis zu produzieren, das wird die sensibelste bisher veröffentlichte Suche für dieses Modell sein.

Schließlich wird die Implementation und Überprufung eines neues “dunkler Schauer” Modell in dem Monte Carlo Generator “Herwig” präsentiert. Dies ist eine alternative Art von DM-Modell, in dem die DM ein Teil eines “dunklen Sektor” mit starken Wechselwirkungen zwischen den Teilchen ist, was zu Schauern von dunkle Teichen führt, die gebundene Zustände bilden; einige davon sind stabil und dienen als ein DM-Kandidat, anderen davon sind instabil und zerfallen in Standardmodell (SM) Teilchen. Das kann zu interessanten experimentellen Signaturen mit Jets mit verschiedener Substruktur im Vergleich zu den SM Hintergründen führen, und die Sensitivität einiger neuer Jet-Substruktur-Variablen in diesem Modell wird untersucht.

Contents

1 Introduction **10**

2 Standard Model of Particle Physics **12**
2.1 The Standard Model Lagrangian . . . 12
2.2 The Electroweak Force . . . 15
2.3 Leptons . . . 16
2.4 The Strong Force . . . 17
2.5 Quarks and Hadrons . . . 17
2.6 The Top Quark . . . 18
2.7 The Higgs Boson . . . 19

3 Motivation for Dark Matter Searches at Colliders **20**
3.1 Evidence for Dark Matter . . . 20
3.1.1 Galactic Rotation Curves . . . 20
3.1.2 Motion of Galaxy Clusters . . . 21
3.1.3 Gravitational Lensing . . . 22
3.1.4 Formation of Structure in the Early Universe . . . 23
3.2 Weakly Interacting Massive Particles . . . 24
3.3 Dark Matter at Particle Colliders . . . 25
3.4 Spin-0 Mediator Dark Matter Model . . . 27
3.5 Hidden Valley Models . . . 29

4 Experimental Setup **31**
4.1 The Large Hadron Collider . . . 31
4.1.1 Luminosity and Pile-up . . . 32
4.2 The CMS Experiment . . . 33
4.2.1 Solenoid Magnet . . . 34
4.2.2 Tracker . . . 35
4.2.3 Electromagnetic Calorimeter . . . 36
4.2.4 Hadronic Calorimeter . . . 37
4.2.5 Muon System . . . 37
4.3 Event Reconstruction . . . 39
4.3.1 Trigger . . . 39
4.3.2 Object Reconstruction . . . 39

5 Monte Carlo Simulation **41**
5.1 Parton Distribution Functions . . . 41
5.2 Matrix Element Generators . . . 43
5.3 General Purpose Event Generators . . . 44
5.3.1 Parton Showers . . . 44
5.3.2 Hadronisation and Decays . . . 46
5.3.3 Underlying Event and Pile-up . . . 48

5.4 Detector Simulation . . . 48

6 Analysis Design **49**
6.1 Introduction . . . 49
6.2 Analysis Tools . . . 50
6.3 Datasets . . . 51
6.4 Trigger . . . 52
6.5 Object Selection and Corrections . . . 56
6.5.1 Electrons . . . 56
6.5.2 Muons . . . 57
6.5.3 Jets . . . 57
6.5.4 B-tagged Jets . . . 58
6.5.5 Missing Transverse Momentum . . . 59
6.6 Event Selection . . . 59
6.7 Event Corrections . . . 61
6.7.1 Top Quark p_T mismodelling . . . 61
6.7.2 Level 1 Trigger Prefiring . . . 62
6.7.3 Failure of HEM Modules . . . 63
6.8 Kinematic Reconstruction . . . 63
6.9 Signal Regions . . . 65
6.10 Sensitive Variables . . . 66
6.11 Neural Network Set-up . . . 69

7 Control and Validation Regions **75**
7.1 $t\bar{t}$ Validation Region . . . 75
7.2 $t\bar{t}Z$ Control Region . . . 75
7.3 DY Control Region . . . 79

8 Systematic Uncertainty Estimation **85**
8.1 Systematic Uncertainty Sources . . . 85
8.1.1 Theoretical Uncertainties . . . 85
8.1.2 Collider uncertainties . . . 86
8.1.3 Detector uncertainties . . . 86
8.2 Smoothing of Systematic Uncertainties . . . 88

9 Fit and Results **90**
9.1 Statistical Fit . . . 90
9.2 Simplified Dark Matter Results . . . 92
9.2.1 Analysis Optimisation . . . 92
9.2.2 Unblinding Checks . . . 95
9.2.3 Results . . . 98
9.2.4 Comparison with Existing Results . . . 102
9.3 Outlook . . . 104

10 Dark Showers with the Herwig Generator **106**
10.1 Dark Shower Models . . . 106
10.2 Herwig Implementation . . . 107
10.2.1 Hard Process . . . 107
10.2.2 Parton Shower . . . 107
10.2.3 Hadronisation . . . 108
10.2.4 Decays . . . 110
10.3 Validation . . . 111
10.3.1 Benchmark Model . . . 111
10.3.2 Parton Shower Validation . . . 112
10.3.3 Hadronisation Parameters . . . 114
10.4 Sensitive Variable Studies . . . 116
10.4.1 Missing Transverse Momentum . . . 116
10.4.2 Angularities . . . 117
10.4.3 Angular Correlation Functions . . . 118
10.5 Outlook . . . 119

11 Conclusion **120**

1 Introduction

For many years, astronomical and cosmological observations [1] have provided a growing body of evidence that all of the visible matter in the universe (stars, planets, interstellar gas, etc.) can only account for about one fifth of the matter [2] in the universe, with the remainder constituting "dark matter", which does not interact with light, and has only been observed via its gravitational effects. No known particle can account for this dark matter, and understanding the nature of this large unknown fraction of our universe is one of the major challenges of modern physics, with the effort spanning a range of disciplines from astronomy to particle physics.

Many cosmological models for dark matter [3] suggest it was produced in the very early universe via a high energy interaction, but this interaction rarely occurs in the universe today since the average energy of particles is much lower. However, it may be possible to probe these energies under laboratory conditions on earth using the Large Hadron Collider (LHC), which accelerates protons to energies comparable with those in the early universe around the time dark matter is believed to have been produced, and brings them together in collisions which may be able to produce new dark matter particles. The main challenge for detecting this process is that the dark matter thus produced is very unlikely to interact with the ordinary matter forming the detector, and hence cannot be directly measured. Instead, searches rely on detecting events where dark matter is produced in association with visible particles, which recoil in one direction, indicating invisible particles must have been produced and travelled in the other direction by momentum conservation. Detecting such events would give a concrete candidate for the particles which make up the majority of the mass of the universe, and potentially lead to further insights of how nature works at high energy scales.

This thesis is organised as follows: chapter 2 introduces the Standard Model of particle physics, which provides a fundamental description of the visible matter in the universe. Chapter 3 then reviews the body of astronomical and cosmological evidence for the existence of dark matter, and introduces some models of dark matter candidates which could be searched for at the LHC. Chapter 4 describes the CMS (Compact Muon Solenoid) experiment at the LHC, which recorded the data used in this work, and chapter 5 describes the Monte Carlo simulations with which this data was compared. Chapters 6 to 9 then describe a search for dark matter produced in association with top quarks in two lepton final states using data collected by the CMS experiment between 2016 and 2018. This analysis contains a large number of improvements with respect to previous CMS searches in this channel, including the addition of a new single top + dark matter signal to the original top quark pair + dark matter signal, the use of new kinematic reconstruction algorithms and neural network (NN) techniques for better signal discrimination, and the addition of new control regions to estimate the rates of certain backgrounds, and will form part of a combined CMS analysis across all decay modes which aims to be the most sensitive search for this dark matter model published to date. Finally, chapter 10 discusses dark showers, a more recent model in which dark matter is part of a "dark

sector", in which some particles are stable and comprise the observed dark matter in the universe, and others are unstable and decay back to Standard Model particles which can be detected in the CMS experiment. The implementation and validation of this model in the Herwig Monte Carlo generator is described, as are some preliminary studies of variables which could be used in future searches for these signatures.

2 Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) represents the culmination of decades of theoretical and experimental developments in the field of particle physics. Although it does not describe all of nature, with the most notable absence in the context of this work being dark matter, it provides a strong, self-consistent model of almost all observations within the field of particle physics, and has often predicted new particles years before their eventual discovery.

2.1 The Standard Model Lagrangian

The Standard Model is a quantum field theory, that is, it describes the universe as a number of fields covering space-time. The possible excitations of these fields are quantised, and can be interpreted as particles localised in space. The SM particles, shown in figure 1, can be divided into three groups: first are the fermions, which have spin 1/2, and are the constituents of matter. These can be sorted into three generations of particles which are identical except for the mass. Within each generation there is a pair of quarks, which interact via the strong force, and a pair of leptons, which do not. The second type of particles are the vector bosons, which have spin 1, and are associated with gauge symmetries of the theory- symmetries under which the fields can be redefined arbitrarily as a function of space. These are responsible for the forces between the matter

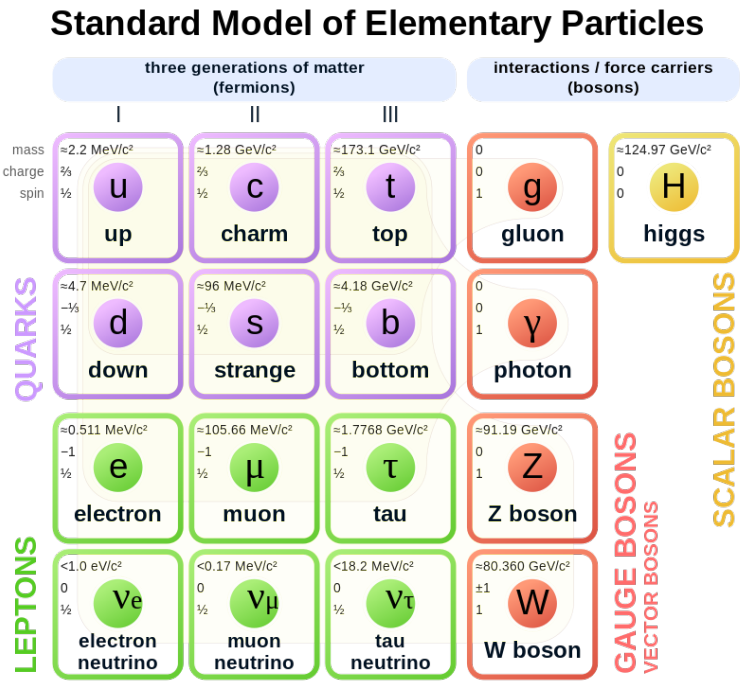


Figure 1: The particle content of the Standard Model. Source: [4]

particles. Finally there is one scalar boson, with spin 0, which arises from the Higgs mechanism, and will be discussed later.

These particles can interact with each other via local interactions, which are described in the SM Lagrangian, and can be written in a condensed form as [5]:

$$\mathcal{L}_{SM} = i\bar{\psi}\gamma^{\mu}D_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \psi_i y_{ij}\bar{\psi}_j\phi + \bar{\psi}_i y_{ij}\psi_j\phi^{\dagger} + |D_{\mu}\phi|^2 + V(\phi) \tag{1}$$

For the first term ψ are vectors of the fermion fields in spin space, γ^{μ} are the Dirac Matrices describing how these fields transform in this space, and D_{μ} is the covariant derivative:

$$D_{\mu} = \partial_{\mu} - i\frac{g}{2}Y_W B_{\mu} - i\frac{g\prime}{2}\tau_L^a W_{\mu}^a - i\frac{\alpha_s}{2}T_{ij}^a G_{\mu}^a \tag{2}$$

The derivative term on the right hand side describes how the fermion fields propagate in the absence of interactions, while the other terms give the interaction of these fields with the gauge fields - the field B_{μ} is associated with a $U(1)$ "hypercharge" symmetry, W_{μ}^a is associated with an $SU(2)$ "isospin" symmetry between the pairs of quarks and leptons in each generation, while G_{μ}^a is associated with an $SU(3)$ "colour" symmetry.

The second term of the lagrangian is formed of contractions of the field strength tensors, which are themselves formed by commuting the covariant derivative terms:

$$F_{\mu\nu} \propto [D_{\mu}, D_{\nu}] \tag{3}$$

Since the derivative terms commute with themselves, this gives rise to two types of terms, the ones formed by commuting the differentials with the fields, such as:

$$\partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu} \tag{4}$$

which describe the propagation of the fields in the absence of interactions, and the commutators of the fields, which give rise to interactions of these fields with themselves. Note however that all the members of the U(1) group commute, so the hypercharge field does not interact with itself.

The third and fourth terms in equation 1 give the interactions for the fermions and anti-fermions with a complex valued scalar field (the Higgs field) which is two-dimensional under the $SU_{isospin}(2)$ group, while the fifth term gives the interaction of this scalar field with the $SU(2) \times U(1)$ gauge fields (and a derivative term describing the evolution of this scalar field). In isolation these terms would give rise to four scalar bosons (since the field is two dimensional and complex valued), however there is a sixth term in the lagrangian, which gives a potential to this field. This potential has a non-zero minimum,

which breaks the $SU_{isospin}(2) \times U_{hypercharge}(1)$ symmetry to a simple $U_{charge}(1)$ symmetry, as will be described further in section 2.2. After this symmetry breaking, there is only a single scalar boson, the Higgs boson, (the other degrees of freedom being absorbed into the longitudinal modes of the vector bosons which acquire mass) and the third, fourth and fifth terms in equation 1 both describe the interaction of the scalar boson with the other particles, and give mass to the matter particles and three of the bosons associated with the broken $SU(2) \times U(1)$ symmetry.

The dynamics described by the Standard Model Lagrangian are very complex, and in general there is no known way to compute the full evolution of all these interacting fields. However, one can note that each interaction term comes with a power of the relevant coupling constant (g, $g\prime$ or α_s for the interactions of the gauge fields with themselves or the fermions, y_{ij} for the interaction of the higgs field with the other fields), and hence, provided these coupling constants are significantly less than 1, one can first calculate a "free state" solution for each field propagating without interactions, considering only the derivative and mass terms, then consider interactions between the particles as perturbations of these solutions. One can therefore represent interactions between particles, such as the scatterings at particle colliders, as diagrams of particles propagating and interacting at vertices, which are called Feynmann diagrams. One can associate an amplitude for the scattering to each diagram: each vertex corresponds to a term with an additional power of the coupling constant for the force which governs the interaction (or, for a few vertices such as the 4 - gluon vertex, two additional powers), and each line gives a propagator term for the corresponding particle. One can the calculate the cross section of each scattering at a particular order in perturbation theory by considering all combinations of diagrams in the squared amplitude which have the corresponding powers of the coupling constant. To calculate the amplitude to leading order (LO), one need only consider the diagrams which connect the incoming to outgoing particles without loops or additional outgoing particles, according to the interactions already described. One such interaction, for a pair of fermions scattering into another pair of fermions, is shown in figure 2. This proceeds via a vector boson mediator, since the fermion fields do not directly interact with themselves.

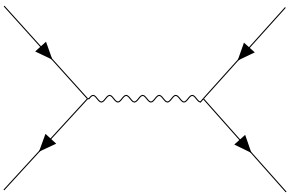


Figure 2: A leading order diagram of a $2 \to 2$ scattering of fermions (straight lines with arrows), mediated by a vector boson (wiggly line).

To get a more accurate prediction, one can consider next-to-leading order (NLO) diagrams, which have an additional particle. There are two types of these diagrams: real emission diagrams with one additional vertex where an additional particle is emitted, and virtual diagrams where the additional particle forms a loop, with two additional

vertices, which interferes with the LO diagram to enter into the amplitude squared at the same order as the real emission diagram. Diagrams of both these types are shown in figure 3.

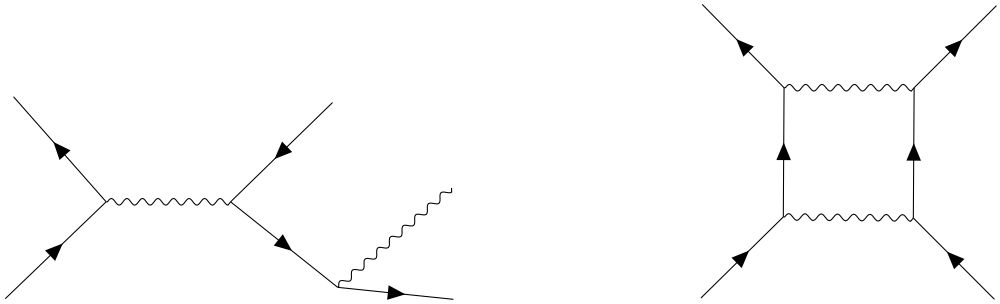


Figure 3: Next-to-leading order diagrams of a $2 \to 2$ fermion scattering. Left: Real emission of a vector boson. Right: Virtual (loop) diagram.

Unfortunately if one naively calculates the amplitudes from these diagrams over all possible momenta one obtains divergent results; the loop diagrams diverge when integrating up to infinite momenta in the loop, while the real emission diagrams diverge either in the case that the emitted particle tends towards zero energy (a soft divergence), or tends to being parallel to another emitted particle (a collinear divergence). These divergences come from the assumption that the free fields are completely non-interacting, when in fact they can have corrections from other fields. After reformulating the theory in terms of renormalised fields, i.e. including corrections from interactions with other fields as they propagate, and the couplings between these fields, the divergences in the loop and real-emission diagrams cancel, giving a finite result. This procedure has two notable consequences: firstly the masses and coupling strengths change as a function of the energy scale at which they are evaluated, since one probes the corrections to these fields to different degrees at different energies, and secondly when calculating amplitudes one must introduce an energy scale at which this evaluation is carried out: the renormalisation scale, which should be selected to be close to the energy of the highest energy particle in the LO process.

The following sections describe the different components of the SM in more detail; all numerical values are from reference [6], unless otherwise stated.

2.2 The Electroweak Force

As has previously been mentioned, the addition of the Higgs field breaks the $SU(2)_{weak} \times U(1)_{hypercharge}$ symmetry by giving masses to the particles, resulting in the pairs of quarks and leptons no longer having the same masses. After this breaking there is a new $U(1)_{charge}$ symmetry which is preserved, which is associated with the well-known electromagnetic (EM) field A_μ, which is associated with a massless vector boson called the photon, γ, which couples to all electrically charged particles. The three vector bosons

associated to the broken symmetry, meanwhile, become massive- there is one Z boson, with a mass of about 91 GeV, and two $W^{\pm}$ bosons, which are antiparticles of each other and have a mass of about 80 GeV. These bosons are referred to as mediating the weak force, since, while this force has a similar strength to the electromagnetic force at high energies, it is much weaker at low energies due to the high masses of these mediators. The W bosons are particularly notable, as fermions can change flavour by emitting these particles. Before the breaking of the $SU(2) \times U(1)$ symmetry, such interactions could only change the flavour within a particular generation, however the Higgs sector links particles of different generations, so once the theory is formulated in terms of the mass eigenstates, transitions between the generations are also possible. The couplings between quark mass eigenstates from different generations are given by the CKM matrix, which is mostly flavour diagonal, i.e. couplings between quarks of the same generation are much stronger than those between generations. Dedicated neutrino experiments have found that there is a corresponding mixing in the lepton sector parameterised by the PNMS matrix, which was not originally predicted in the SM. This matrix has been found to be significantly less diagonal than the CKM matrix, however this is rarely considered in collider experiments since neutrinos are generally not observed.

2.3 Leptons

The leptons are the fermions which only interact via the electroweak forces. Within each generation, there is a charged lepton, and an uncharged neutrino. The lightest charged particle is the electron, e, which is stable with a mass of 551 eV; the next is the muon, μ, with a mass of 106 MeV, which is normally taken to be stable since its relatively long mean lifetime of $2 \times 10^{-6} s = 300\text{m}/c$ means it rarely decays inside the detector. The heaviest charged lepton is the tau, τ, with a mass of 1.78 GeV, which decays rapidly via the weak force, 35% of the time to one of the lighter charged leptons and two neutrinos, as shown in figure 4, the rest of the time to final states containing hadrons (the bound states of quarks, which will be described in more detail in section 2.4).

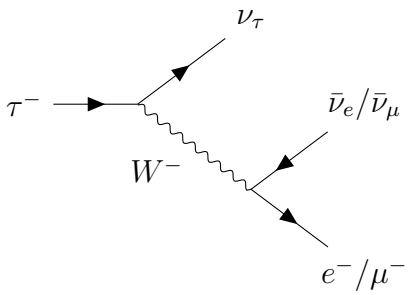


Figure 4: Decay of a tau lepton to a tau neutrino and a lighter lepton and the associated antineutrino.

Finally in the lepton sector there are the neutrinos. Since they are electrically neutral, these particles are usually not observed at particle colliders, and their presence can

only be inferred from a momentum imbalance of the other particles. Since they are not observed, these particles are normally considered in the states which are flavour diagonal to the charged leptons, ν_e, ν_μ and ν_τ, though dedicated neutrino experiments have observed mixing between these states, demonstrating that, like the other particles, they propagate as mass eigenstates.

2.4 The Strong Force

The strong force is associated with the $SU(3)_{colour}$ symmetry of the Standard Model, described by a theory called quantum chromodynamics (QCD). In this theory the quarks carry one of three colour charges, commonly called "red", "green" and "blue" in analogy to the three primary colours of visible light, while their anti-particles carry corresponding anti-colours. The vector bosons, called gluons, are massless and transform in the adjoint representation of the group and there are hence eight of these mediators, which to leading approximation can be considered to carry both a colour charge and an anti-charge.

The strong force is, as its name suggests, stronger than the other SM forces. To give an indicative value, at the mass of the Z boson, the coupling strength of the strong force is about $\alpha_s(M_Z) \approx 0.1$, whereas for the electromagnetic force $\alpha_{EM}(M_Z) \approx 0.008$. However, the most significant difference from the other forces is that whereas they slowly increase in strength as the energy scale increases, α_s increases as the energy scale decreases. This has two major consequences: the first is that for lower energies, one cannot ignore higher order terms in the perturbative expansion, and hence any particle with a colour charge will produce a "shower" of low energy radiation, giving a "jet" of particles rather than a single particle. The second consequence is that below about 1 GeV all colour charges are confined into colour-neutral bound states called hadrons. There are two possible ways this occurs: one can either have a quark of a given colour paired with an anti-quark of the same colour, which is referred to as a meson, or one can have three (anti-) quarks, one of each (anti-) colour, which form an (anti-) baryon. Since the approximation $\alpha_s << 1$ no longer holds in this region, one cannot compute the properties or formation of these hadrons using perturbation theory, so empirical models are used instead (see section 5.3.2).

2.5 Quarks and Hadrons

As with the leptons, there are six flavours of quarks, all of which are electrically charged and can change flavour via the weak force. However, the quarks differ from the leptons in that they are charged under the $SU(3)_{colour}$ group, and hence are confined into hadrons at low energy scales. The three lightest quarks, the up, down and strange, have masses much lighter than the confinement scale, and hence the exact values are hard to define. These form a large number of possible hadrons, which will not be detailled here, however one deserves a special mention: the proton, p, is the lightest baryon, and as such is stable

and forms a large part of the matter in the universe. As a result, it is the particle most commonly accelerated in hadron colliders, such as the LHC (see 4.1) used in this work.

The next two quarks, the charm and the bottom, are significantly more massive, with masses of approximately 1.3 GeV and 4.2 GeV, respectively. As a result, they form more massive hadrons, the lightest of which for each flavour decay via the weak force since only this can change the flavour of the constituent quarks. As this decay proceeds via an off-shell W boson, it is a reasonably slow process, and the lightest b and c hadrons therefore have relatively long lifetimes. B hadrons, in particular, have a mean lifetime of approximately $1.5 \times 10^{-12} s = 0.4\text{mm}/c$, which means B hadrons can potentially be identified by having a decay displaced from the other interactions.

2.6 The Top Quark

The top quark differs from the other quarks in that it is significantly heavier again than the charm and bottom quarks, with its mass measured to be around 172.5 GeV [7]. As a result the top quark decays very rapidly to a quark (which is almost always a bottom quark since the CKM matrix is highly diagonal) and a W boson, and has not been observed to form any QCD bound states. One interesting consequence of this is that the spin information of the top quarks is transferred to the W boson, and thence to its decay products [8], which is not the case for the other quarks since the hadronisation process scrambles the quark spin. This is particularly relevant for leptonic top decays, i.e. events where the W boson decays to a charged lepton and a neutrino, as not only is the charged lepton very easy to identify, it is (neglecting very minor NLO QCD corrections) a perfect spin analyser for the top quark, allowing one to find the spin of the parent top quark from the direction of this lepton.

The probability for a top quark to decay to each flavour of leptons is 11%, however since the tau is itself unstable and decays to electrons and muons 35% of the time, the total decay fraction to electrons and muons including via taus is about 26%, with the remainder being decays to hadronic final states, mainly due to $W \to q\bar{q}'$ decays.

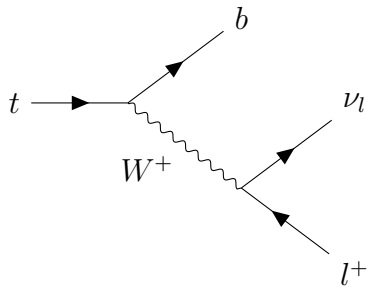


Figure 5: Leptonic decay of a top quark to a b quark, a charged anti-lepton l^+, and a neutrino of the corresponding flavour.

2.7 The Higgs Boson

The scalar Higgs boson has a mass of 125 GeV, and interacts with a Yukawa coupling, i.e., proportional in strength to the particle it is interacting with. As the most recently discovered particle in the Standard Model, the interactions of the Higgs boson have not been studied in as much detail, leaving open the possibility of beyond the Standard Model (BSM) interactions of the Higgs boson. One observation that particularly motivates these is that if any higher energy particles exist, they should add corrections to the Higgs boson mass via diagrams of the type shown in figure 6, which would make it significantly more massive than the other SM particles. This is referred to as the hierarchy problem, and is not an issue for the other particles as they are protected by a symmetry (the gauge symmetries for the vector bosons, the chiral symmetry between left- and right- handed helicities for the fermions), which means such diagrams can only make fractional changes to the mass of these particles. BSM theories which attempt to resolve the hierarchy problem often therefore introduce a new symmetry which also protects the Higgs mass from such corrections, which can lead to new scalars around the TeV scale.

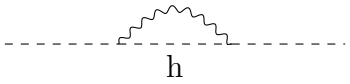


Figure 6: Possible diagram of a new vector boson interacting with the Higgs boson, increasing its mass.

3 Motivation for Dark Matter Searches at Colliders

Perhaps the greatest challenge to the Standard Model comes not from particle physics, but from astronomy, where a large body of evidence has accumulated that the majority of the matter in the universe is not made of the matter described in the SM, but of a new type of matter, referred to as dark matter (DM). This chapter will discuss some of the evidence for dark matter, how this motivates searching for a new DM particle at colliders, and outline some concrete models which will be considered in this thesis.

3.1 Evidence for Dark Matter

This section reviews some of the main astronomical evidence for dark matter (partially following [1]), and what properties a potential dark matter particle should have.

3.1.1 Galactic Rotation Curves

One of the earliest motivations for DM searches came from astronomical observations of the speed of objects in nearby galaxies, as measured by the Doppler shift of the light coming from stars and other objects. According to Newton's laws of motion, the speed of a gravitationally bound rotating object is:

$$v = \sqrt{\frac{GM(r)}{r}}, \tag{5}$$

where r is the radius of the orbit, G is the gravitational constant, and

$$M(r) = \int_0^r 4\pi r'^2 \rho(r') dr' \tag{6}$$

is the mass enclosed by the orbit, in terms of density the profile of the galaxy, $\rho(r)$. Observations of rotation speeds for stars within the main bodies of galaxies generally agree well with these predictions, however for satellite objects outside the body of the galaxy the velocities are expected to decrease as $v \propto r^{-1/2}$, since there is little mass visible at these large radii. Instead the velocities of these objects were observed [9, 10] to be generally constant as a function of radius at these large radii. This implies that there is a large halo of matter extending a long way beyond the visible galaxy, with a profile of $M(r) \propto r \implies \rho(r) \propto r^{-2}$. Such a profile could be explained by a weakly- or non-interacting "dust" held together by gravitational attraction. However, this dust is known not to interact with light, since then this halo would both absorb light passing through it, giving clear absorption lines in spectra of more distant objects, and warm up and emit thermal radiation, which would also be visible. An example of such a rotation curve, as well as the mass distributions of the visible components of the galaxy and that

attributed to DM, is shown in figure 7. This effect has now been observed for a very large number of galaxies- the rotation curves are not always exactly flat, as the DM distribution varies somewhat between galaxies, but almost all galaxies require a large component of DM to explain the observed rotation curves at high radii.

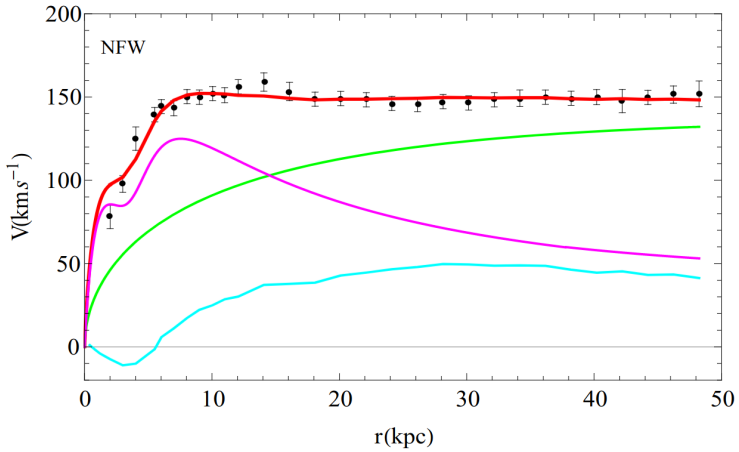


Figure 7: The rotation curve of the NGC 3198 galaxy (black data points), together with the predicted rotation curves from the stellar mass distribution (magenta), hydrogen gas (blue), the component assigned to dark matter (cyan) and the total (red). Source: [11]

3.1.2 Motion of Galaxy Clusters

While gravitational rotation curves provide a very clear argument for DM, the first evidence [12] came from studying larger objects, specifically galaxy clusters, which are groups of galaxies and intergalactic gasses held together by gravity, and represent some of the largest structures in the universe. While these more dispersed objects do not tend to orbit around a common center in the same way as galaxies, as bound systems their kinetic energy, T, should be related to their potential energy, V, by the virial theorem, $2T = V$. The kinetic energy is:

$$T = \frac{1}{2}\sum_i m_i v_i^2 = \frac{1}{2} M \langle v^2 \rangle \tag{7}$$

where m_i and v_i are the mass and velocity of the individual components of the cluster, M is the total cluster mass and $< v^2 >$ the average velocity squared, which can be estimated from the dispersion of red-shifts within the cluster, assuming the motion is roughly isotropic, so the velocities in the plane of the sky are comparable to those towards earth. Estimating the potential energy of the cluster is more challenging, since it depends on the strength of the attraction of all components on each other, however it can be written as:

$$V = \sum_i \sum_{j<i} \frac{G m_i m_j}{r_{ij}} = -\frac{\alpha G M^2}{r_h} \tag{8}$$

where G is the gravitational constant, r_{ij} is the distance between objects i and j, and r_h is the half-mass radius of the cluster, which is estimated from the radius containing half the luminous objects in the cluster, and α is a parameter based on the density profile of the cluster, estimated from kinematic fits. One can therefore estimate the mass of the cluster as:

$$M = \frac{r_h \langle v^2 \rangle}{\alpha G} \tag{9}$$

Using this approach on one nearby cluster, the coma cluster, gives that visible matter (mainly in the form of the large intergalactic gas clouds) only comprises around 10% [13] of the total cluster mass, with the rest coming from DM. While the difficulties of accurately modelling the cluster mean this number has a rather large uncertainty, it is useful since clusters are the largest observed structures in the universe, making them a useful proxy for the total DM density in the universe as a whole today.

3.1.3 Gravitational Lensing

Further evidence of dark matter can be obtained from gravitational lensing- the deflection of light from more distant background objects by the gravitational effects of foreground objects, as described by general relativity. In extreme cases, this can result in multiple images of the background objects, or even a "ring" of smeared images, as shown for example in figure 8. The strength of the lensing can be used as a measure of the mass of the galaxy, which again tends to be significantly larger than estimates of

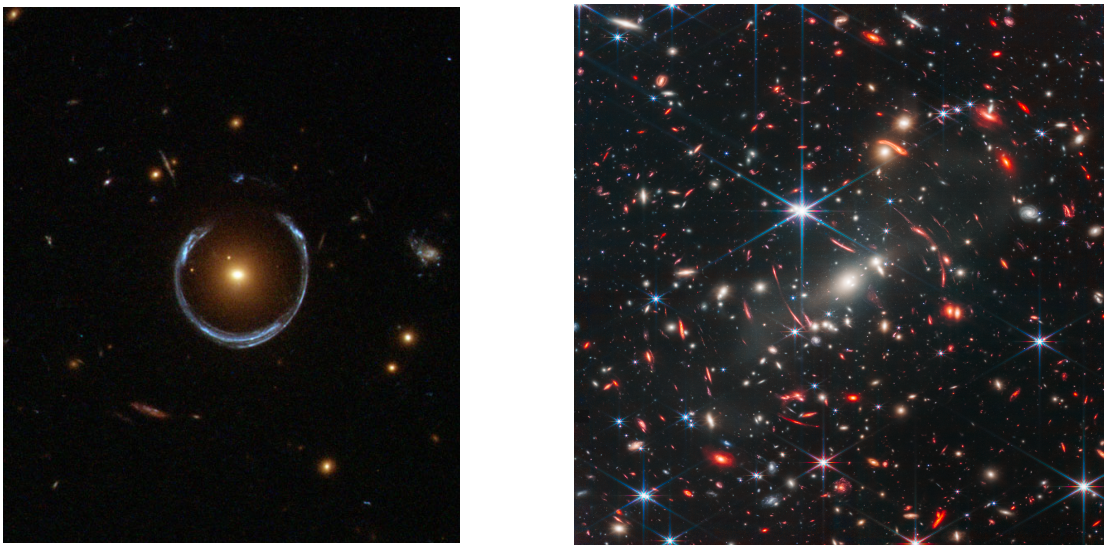


Figure 8: Examples of gravitational lensing: left - an image from the Hubble Telescope [14] showing a distant blue galaxy being lensed into broken ring by the redish galaxy in the center of the image which sits almost directly in front of it; right - an image from the James Webb Space Telescope showing many distant red galaxies being lensed by foreground objects (yellow and white galaxies).

the visible mass of the galaxy.

Perhaps the most compelling evidence provided by gravitational lensing, however, comes from the bullet cluster, which is actually two galactic clusters which have collided, resulting in the associated plasma gas clouds being slowed by the collisions, with one formed into a bullet-like shape, hence the name. A survey was performed of the mass distribution of these clusters by studying the lensing of background galaxies [15], which found that the bulk of the mass distribution is not centered on these plasma clouds, but is further apart (figure 9), suggesting that unlike the plasma, which was slowed by the collision, the dark matter halos passed through each other with minimal interactions. This result is particularly interesting for two reasons: first, it strongly supports the interpretation of the other results in terms of a new type of matter particle, rather than other theories such as modifications of gravity at large scales, which could explain the rotation curves of galaxies, but cannot account for a gravitational effect displaced from the visible matter as seen here, and second, it suggests that dark matter must interact with itself very weakly at large scales, in comparison to the electromagnetic force which slowed the colliding plasma clouds.

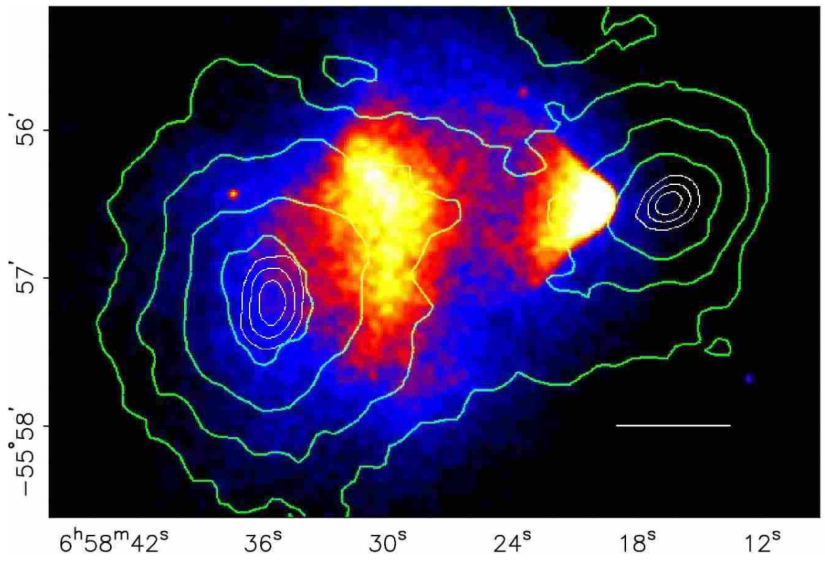


Figure 9: The bullet cluster. The colour map represents the plasma clouds, as observed by the Chandra X-ray observatory, while the contours represent the mass distributions as reconstructed from gravitational lensing. Source: [15]

3.1.4 Formation of Structure in the Early Universe

A further significant piece of evidence for dark matter comes from the cosmology. This describes the evolution of the universe and structures within it at the largest scales. The universe is believed to have begun in an event referred to as the "big bang" approximately 13.6 billion years ago, when the universe was compressed into a very small, dense

region, and has expanded ever since. At very early times, the universe was formed of a hot plasma of charged particles and photons, but as it expanded the energy of the particles decreased, eventually to the point that the charged particles bound into atoms, which allowed the photons to propagate freely as black-body radiation. This radiation has now been doppler-shifted to considerably longer wavelengths due to the expansion of the universe, and now appears as if it had been emitted by a black body with a temperature of $3K$, which is in the microwave region of the EM spectrum. This cosmic microwave background (CMB) is visible at the largest distance scales in all directions. Measurements of this background by the Planck [2] microwave observatory show that this spectrum is very homogeneous, varying on average only by 1 part in 10^5. These fluctuations are ten times too small to explain the galaxies, clusters and the other structures observed today, assuming the universe is only composed of this visible matter, which collapsed under gravity. However the formation of structure can be explained by the presence of dark matter which does not interact with the photon radiation (at least at the relatively late times and low energies at which the CMB was emitted), allowing larger variations to form, which gravitationally collapsed to form the structures observed today.

More detailed observations of the spectrum of variations of the CMB allow the determination of the relative densities of visible matter, dark matter, and dark energy, the energy associated with empty space, which causes the expansion of the universe to accelerate. Dark Matter is again found to represent approximately 80% of the matter in the universe, and 20% of the total energy-matter density [2]. The fact that the DM density at these early times is roughly comparable to that measured in nearby galaxy clusters indicates that the DM must be "cold", i.e. move at non-relativistic speeds, as if it were "hot" (relativistic), it would lose energy relative to the non-relativistic visible matter as the universe expanded, which slows the motion of all particles and hence reduces the energy of relativistic particles. Simulations of structure formation support this conclusion, as only cold dark matter can produce the observed structures, whereas hot dark matter would produce the large structures like galaxy clusters, but due to its high speed would not form smaller structures like galaxies.

3.2 Weakly Interacting Massive Particles

Although dark matter has not yet been directly detected, the detailed body of evidence outlined in the previous section gives strong indications of its properties, which guide search regions for future detection. Firstly, it must interact very weakly with Standard Model particles, since it has not been detected either by astronomical observations or experiments which search directly for dark matter interacting with large volumes of Standard Model matter. Furthermore, from the bullet cluster and simulation of structure formation, it should also only interact weakly with itself. Evidence from the CMB also shows that dark matter should be cold. A very simple way to realise this requirement is to assume that the dark matter is quite massive (with a mass of at least a few hundred

MeV), which means that provided it follows a similar thermal profile to the SM matter, it would be non-relativistic by the time the CMB was emitted. This information leads to the weakly interacting massive particle (WIMP) DM hypothesis, where the dark matter is assumed to have a mass > 100 MeV, and to couple via a massive mediator to the SM. DM would therefore have been able to interact with the SM particles in the very early universe, when the average energy of the particles was high enough that collisions between particles would often produce the heavy mediator, but would very rarely interact in the universe today, as particles have much lower energies and hence their collisions would not have sufficient energy to produce the mediator.

The most commonly hypothesis for how the modern DM density occurred is the "thermal freeze-out" model, in which DM was in thermal equilibrium with the SM in the very early universe, then as the universe expanded and cooled the DM density would decrease, as the rate at which the heavier DM particles produced SM particles exceeded the process in the opposite direction. However, as the universe expanded the rate of interactions in both directions would also decrease due to both the decreasing density of both SM and DM particles (and hence rate of collisions between them), and the decreasing cross section, since the mediator would be more off-shell due to the lower energy of the particles. At some point the rate of interactions would decrease to almost zero, leaving a population of DM particles which are no longer in equilibrium with the SM sector, as shown in figure 10. Larger cross sections for the interaction of SM and DM particles lead to a lower DM relic density, as the freeze-out occurs at later times. From the observed DM relic density one can therefore determine the cross section for DM production, and if one assumes the couplings of the mediator to both the DM and SM particles are of order 1, the mediator must have a mass in the range 100 GeV to 1 TeV to produce the observed density of DM [3]. This is dubbed the "WIMP miracle", as this mass range is exactly that which will be explored by the LHC (see section 4.1), motivating a broad search programme for production of DM.

It is worth noting, however, there are other models, for instance if the couplings are much smaller the mediator may be significantly lighter, while on the other hand some models do not assume the DM relic density is due to thermal freeze-out, but due to "freeze-in" of DM that was never in thermal equilibrium with the SM due to a very heavy mediator with very weak couplings. Such signatures may not be detectable at the LHC, but are searched for in other experiments, such as direct detection, which looks for interactions of the DM halo of our galaxy with SM matter on earth, or indirect detection, which looks for SM particles produced by DM annihilation or decays in astronomical signatures.

3.3 Dark Matter at Particle Colliders

As described in the previous section, the Large Hadron Collider is an excellent candidate for dark matter searches since it covers the most promising mass range for WIMP DM. The challenge, of course, is that while DM can be produced in the high energy collisions of the LHC, its very weak low energy interactions with the SM mean it is almost impos-

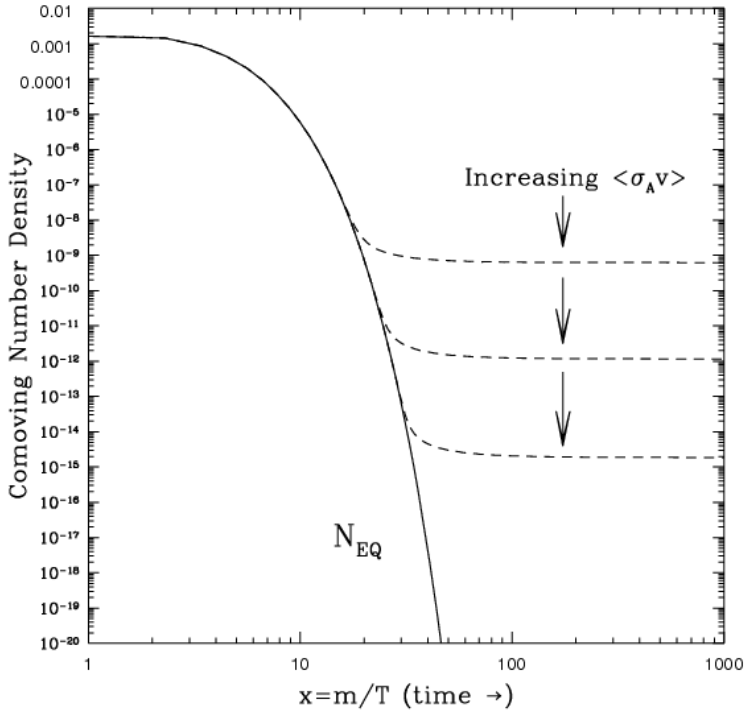


Figure 10: The density of DM as a function of time. The solid line indicates the density for no freeze-out, while the dashed lines indicate three different freeze-out scenarios for different cross sections. Source: [16]

sible to detect. Searches for DM therefore rely on processes where the DM is produced in association with SM particles, allowing the DM to be detected via momentum conservation, as the SM particles would recoil against the DM particles, so events with an imbalance in the total measured momentum would indicate that invisible particles with momentum opposite to the measured momentum imbalance had been produced. This requires measuring the momenta of all the visible particles, then the negative vector sum of these momenta will be equal to the vector sum of the momenta of the invisible particles, for two colliding particles of equal momenta. However for hadron colliders it is not possible to accurately measure the momentum imbalance in the direction of the beam, since a large amount of momentum is carried by proton remnants down the beam pipe, which unfortunately cannot be completely instrumented. DM searches therefore tend to search for events with a high negative vector sum of momentum transverse to the beam, which is refered to as missing transverse momentum, p_T^{miss}.

Searches for Dark Matter are categorised in terms of which Standard Model particles are produced in association with the DM, measured as p_T^{miss}. Originally, these searches were performed using effective field theories (EFTs), which assume that the DM mediator is significantly higher energy than the rest of the process, so can be "integrated out", giving an effective vertex in the Lagrangian which couples the DM particles to some

SM particles. However as discussed, the expected range of WIMP mediator masses is actually within the energy range probed by the LHC, so integrating out the mediator in this way potentially misses important kinematic information. On the other hand, there are a large number of full models which predict DM and in some cases other particles which can resolve other theoretical problems, however each of these models predicts slightly different kinematics and cross sections for different processes, making it difficult to organise a search programme covering the full potential phase space. A set of simplified benchmarks [17] were therefore proposed, which each introduce only a single new mediator particle and dark matter candidate, but cover the majority of the phase space described by more complex models.

3.4 Spin-0 Mediator Dark Matter Model

The majority of this work will concentrate on two such benchmark models, where the dark matter is a new fermion, which couples to the SM via either a new spin-0 mediator, which has a yukawa type coupling, similar to the Higgs boson 2.7. The mediator can either be a scalar (which has even parity, i.e. the wave-function is invariant under spatial inversion) or a pseudoscalar (which has odd parity, i.e. the wave-function becomes negative under spatial inversion). Such particles are often produced in models with extended Higgs sectors, for example the 2HDM+a model [18], which would simultaneously provide a DM candiate and resolve the hierachy problem discussed in 2.7. The fermion is taken to be a Dirac fermion (i.e. one with a complex valued wave-function, like the SM fermions); some UV complete models instead predict DM to be a Majorana fermion (with a real vauled wave-function), however one can still reinterpret these results in this context as this only changes the cross section rather than the kinematics.

For the scalar mediator the SM lagrangian is extended by the following terms:

$$\mathcal{L}_\phi \supset -g_\chi \phi \bar{\chi}\chi - \frac{\phi}{\sqrt{2}} \sum_{q=u,d,s,c,b,t} g_q y_q \bar{q} q - \frac{1}{\sqrt{2}} m_\phi^2 |\phi|^2 - m_\chi |\chi|^2, \tag{10}$$

where χ is the DM fermion, ϕ the scalar mediator, and y_q are the same yukawa couplings with which the SM Higgs boson couples to the quarks. For the pseudoscalar mediator, A, the lagrangian is similar, but with an extra $\gamma_5 = i \prod_{\mu=0}^{4} \gamma_\mu$ matrix, which flips the spin of the fermions:

$$\mathcal{L}_a \supset -g_\chi a \bar{\chi}\gamma_5\chi - \frac{ia}{\sqrt{2}} \sum_{q=u,d,s,c,b,t} g_q y_q \bar{q}\gamma_5 q - \frac{1}{\sqrt{2}} m_a^2 |a|^2 - m_\chi |\chi|^2, \tag{11}$$

In either case, the model has four free parameters: the mass of the dark matter particle, m_χ, the mass of the mediator, $m_{\phi/a}$, and the couplings of the mediator to the DM particle

g_χ and the SM fermions g_q. In practise, LHC searches have very little sensitivity to m_χ, as the kinematics do not change for $m_\chi < m_{\phi/a}/2$, while for $m_\chi > m_{\phi/a}/2$, the cross section decreases very rapidly, as the mediator is significantly off-shell. We therefore take $m_\chi = 1$ GeV, and focus on variations in mediator mass in the range 50-500 GeV. Since varying the couplings simply changes the expected cross section, we take $g_\chi = g_q = 1$, and set limits relative to these expected cross sections.

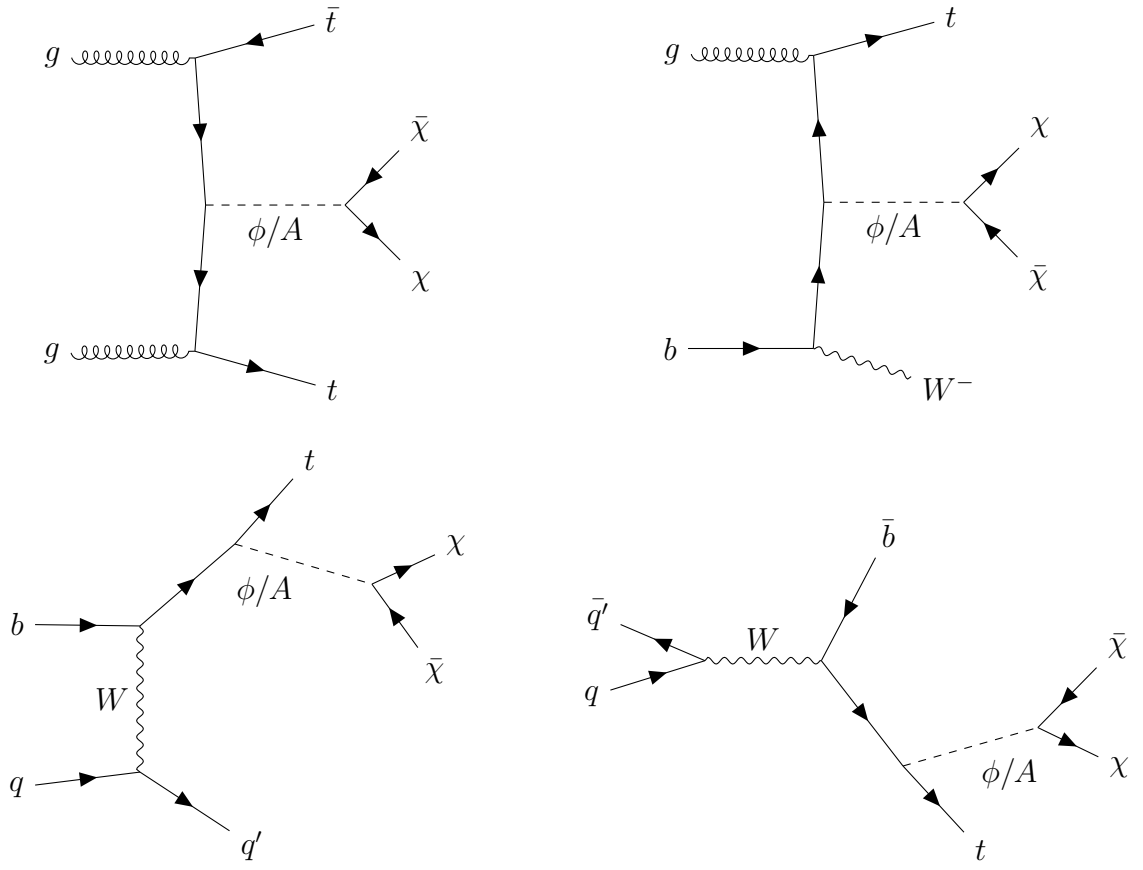


Figure 11: Diagrams for production of dark matter in association with top quarks. Top left: $t\bar{t}$+DM, top right: tW + DM, bottom left: t-channel top + DM, bottom right: s-channel top + DM

Since the mediators have a Yukawa type coupling, it couples most strongly to the heaviest particle in the SM, the top quark, so a natural signature for this model is top quarks recoiling against some invisible particles. Possible LO diagrams for this process are shown in figure 11. The dominant process is production of DM with a top quark pair ($t\bar{t}$ + DM), however there are other processes in which a single top quark is produced, for which the cross section is generally lower than for the $t\bar{t}$+DM process, but decreases more slowly with increasing mediator masses (figure 12) due to the lower energy required to produce one top quark than two, making these processes particularly relevant for the higher mediator masses targeted in this search. The processes are collectively referred to as single top + DM (t + DM), but can be split into three different diagrams: in the

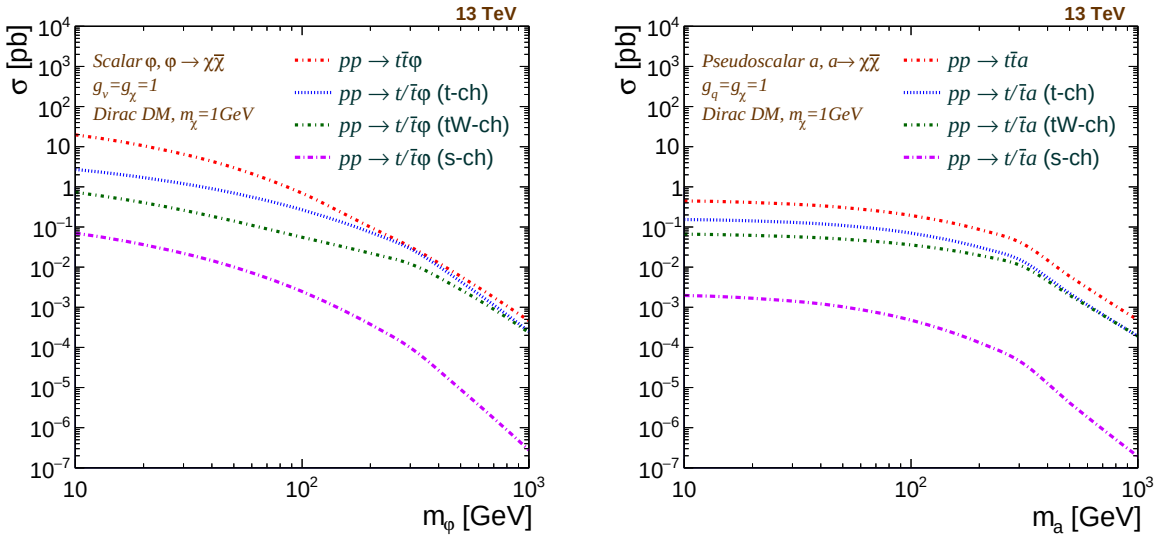


Figure 12: The cross section of the different diagrams for production of dark matter in association with top quarks, as a function of the mediator mass.

first process a single top quark is produced by a W boson decaying to an top quark and a b anti-quark (s-channel single top + DM), which has a cross-section two orders of magnitude below the $t\bar{t}$ + DM process across the entire mediator mass range as the W boson is required to be very off-shell. In the second process a b quark is converted to a top quark by exchange of a W boson with another "spectator" quark (t-channel single top + DM), while in the third the b quark converts via emission of a W boson (tW + DM). Each of these processes have a corresponding process where every particle is exchanged for it anti-particle, which are also considered here.

It should be noted these are not the only searches sensitive to this model, in particular searches for a single jet + p_T^{miss} can have higher sensitivity for higher mediator masses, especially for the pseudoscalar mediator, however this search is the most sensitive for lower mediator masses (particularly for scalar mediators).

3.5 Hidden Valley Models

Since no evidence has yet been found for dark matter at the LHC, there is now an ongoing effort to consider signatures which may have been missed by the simplified benchmarks described in the previous section. One such possibility is for the DM to not simply be composed of non-interacting fermions, but to strongly interact with itself, similarly to QCD in the SM. It is worth noting this does not conflict with the observation from astronomy that DM must interact only weakly with itself, as this new strong interaction would be confined inside hadrons, and hence not be relevant to objects on astronomical scales. In this model, DM would not be produced simply as a pair of stable particles,

but produce a shower of dark particles which then form dark hadrons. Some of these hadrons would need to be stable to act as dark matter candidates, however some might decay back to SM particles. This is referred to as a "hidden valley" model, since these signatures can escape constraints that would apply both from traditional p_T^{miss} based DM searches, since not all of the dark shower contributes to p_T^{miss}, and constraints from resonance searches in the spectra of SM particles, since these semi-visible jets will not form a clear resonance. The development of these models with the Herwig Monte Carlo Generator to facilitate future searches is described in section 10.

4 Experimental Setup

In order to probe the high energy scales of the Standard Model, one needs a particle collider to accelerate elementary particles to these high energies and collide them, and a detector to measure the resulting particles. This chapter outlines the collider and detector used in this work, as well as the process by which the raw data obtained from the detector is converted to a format which can be used for a DM search.

4.1 The Large Hadron Collider

The Large Hadron Collider at CERN is the largest particle accelerator in the world, which enables it to reach the highest energies, colliding protons at a center of mass energies of up to $\sqrt{s} = 13.6$ TeV. The two driving motivations for the construction of the LHC were firstly to search for and study the properties of the Higgs boson, the only SM particle undiscovered at the time, and secondly to search for physics beyond the Standard Model, in particular possible candidates for dark matter. However, to fully realise its potential, the LHC and its experiments carry out a broad physics programme, including studies of the top quark, B hadrons, and even collisions of heavy ions to allow studies of quark-gluon plasma.

The large hadron collider is a circular collider with a circumference of 27km [19], with two beams of protons rotating in opposite directions, separated into a series of "bunches" of protons. The majority of this circumference is occupied by superconducting electromagnets, which bend the high energy beam of protons round the circular collider. There is also a set of radio frequency (RF) electromagnetic cavities, which are used to accelerate the beams. The beams are accelerated by a series of pre-accelerators (see figure 13) and injected into the LHC at an energy of 450 GeV, and then accelerated to the collision energy by passing many times through the RF cavities, gaining 485 keV per turn. The rest of the circumference is filled with magnet systems for focusing and controlling the beam. There are four major experiments on the LHC ring; the first two are general purpose experiments, designed to facilitate analyses covering the entire LHC physics programme: ATLAS (A Toroidal Lhc ApparatuS) and CMS (Compact Muon Solenoid), with which this analysis was performed. In addition there is the LHCb (LHC beauty) experiment, designed to study the properties and decays of B hadrons, and ALICE (A Large Ion Collider Experiment), which is focused on studying heavy ion collisions.

The LHC started colliding proton beams for physics analysis at $\sqrt{s} = 7$ TeV in 2010, increasing to $\sqrt{s} = 8$ TeV in 2012. This data allowed the LHC to achieve the first of its physics objectives, with the ATLAS and CMS experiments announcing the discovery of the Higgs boson on the 4^{th} July 2012 [21, 22]. The LHC then shut down until 2015 to allow upgrading the collision energy to $\sqrt{s} = 13$ TeV, and ran at this energy until 2018, which is referred to as Run 2 data-taking. This run was used to perform extensive measurements of the Higgs boson and other SM particles, as well as a large number of searches for BSM physics, including the dark matter search described in this work. In

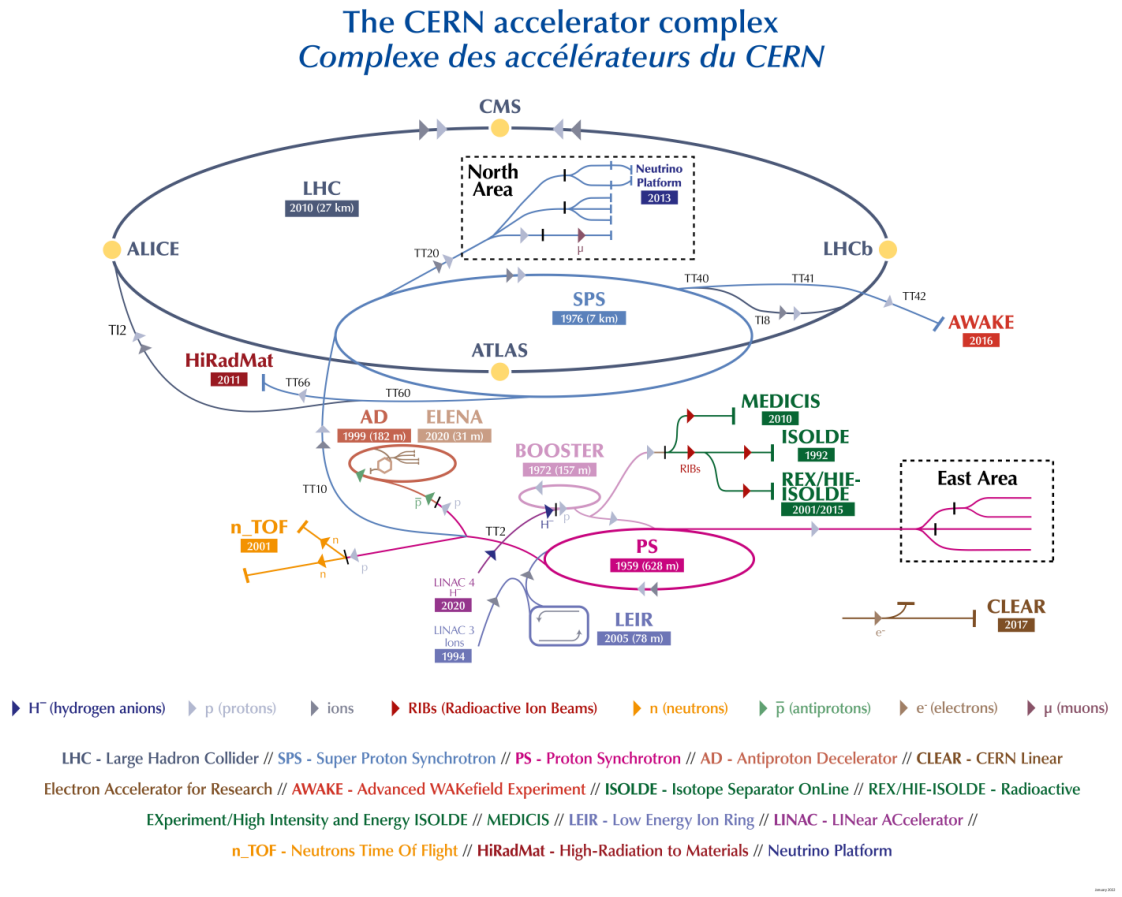


Figure 13: The CERN accelerator complex, showing the LHC, its pre-accelerators which accelerate the proton beams to LHC injection energy, and the numerous other experiments which utilise these pre-accelerators when they are not being used to provide LHC beams. Source: [20]

addition to proton-proton (pp) collisions, the LHC also collides heavy ions such as lead and xenon, which aim to allow the study of a very hot, dense state of matter called quark-gluon plasma, and performs other runs such as lower energy pp collisions to be used as a reference for these measurements. The LHC has since been upgraded again, with pp collisions starting at $\sqrt{s} = 13.6$ TeV in 2022, and over the coming years a lot more data will be taken to allow the LHC to continue its study of the SM and searches for BSM phenomena.

4.1.1 Luminosity and Pile-up

The rate at which a particular physics process occurs is:

$$\frac{dN}{dt} = \frac{f N_1 N_2 \sigma}{A} = L\sigma \tag{12}$$

where f is the frequency at which bunches collide, N_1 and N_2 are the number of protons in each beam, A is the area in which the beams collide, and σ is the cross section for the process. One can therefore increase the rate for all processes by having more particles in the beam, or by focusing them so they collide in a narrower area. These beam parameters are all absorbed into a single value, called the luminosity, L. To allow easy predictions of the number of different types of events for periods of data-taking, one often considers the integrated luminosity, $L_{\text{Int}} = \int L dt$, which has units of inverse area, most commonly measured in fb^{-1} ($1\text{fb}^{-1} = 10^{43}\text{m}^2$).

Since the high energy processes of interest have a much lower cross section than the total cross section for all proton-proton interactions, a very high luminosity is employed, which leads to multiple interactions per bunch crossing. Therefore for the events studied the high-energy collision of interest is accompanied by a number of "pile-up" events, which are lower energy QCD scatterings. Increasing the instantaneous luminosity, and hence the average number of pile-up interactions, increases the rate at which interesting processes are produced, but at the cost that the increased number of additional collisions can make the high energy scattering harder to analyse.

4.2 The CMS Experiment

The bulk of the work performed in this thesis used data collected by the CMS experiment. This experiment has been built and maintained by a collaboration of scientists from around the world [23] and aims to study a range of high energy phenomena. To achieve this, the experiment is constructed as a cylinder along the LHC beam, centered on the point at which the two beams collide. The experiment consists of a set of different types of detectors which aim to measure the different types of particles. These are generally structured as a barrel, which runs the length of the cylinder, and two endcaps, which are discs at the ends of the cylinder, covering almost the entire region into which particles could travel, with the exception of a two small regions where the beams enter and leave the experiment. To aid orientation, a coordinate system is established for the experiment, with the z axis along the direction of one of the beams, the x axis towards the center of the LHC ring, and the y axis away from the center of the earth. Given the cylindrical nature of the detector, it is often useful to use angular coordinates. The first two of these are inherited from cylindrical polar coordinates: the radius in the xy-plane, r, and the azimuthal angle, ϕ. However, instead of the angle with respect to the xy-plane in the z direction, θ, the third coordinate is chosen to be the pseudorapidity:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{13}$$

which has the advantage that differences in pseudorapidity are invariant under Lorentz boosts, which is important since the individual quarks and gluons in the protons which collide can have different energies, and hence the center of mass frame of the collision is

not necessarily the same as the frame of the detector. The CMS experiment is shown in figure 14, and the various detector components will be described in the following sections.

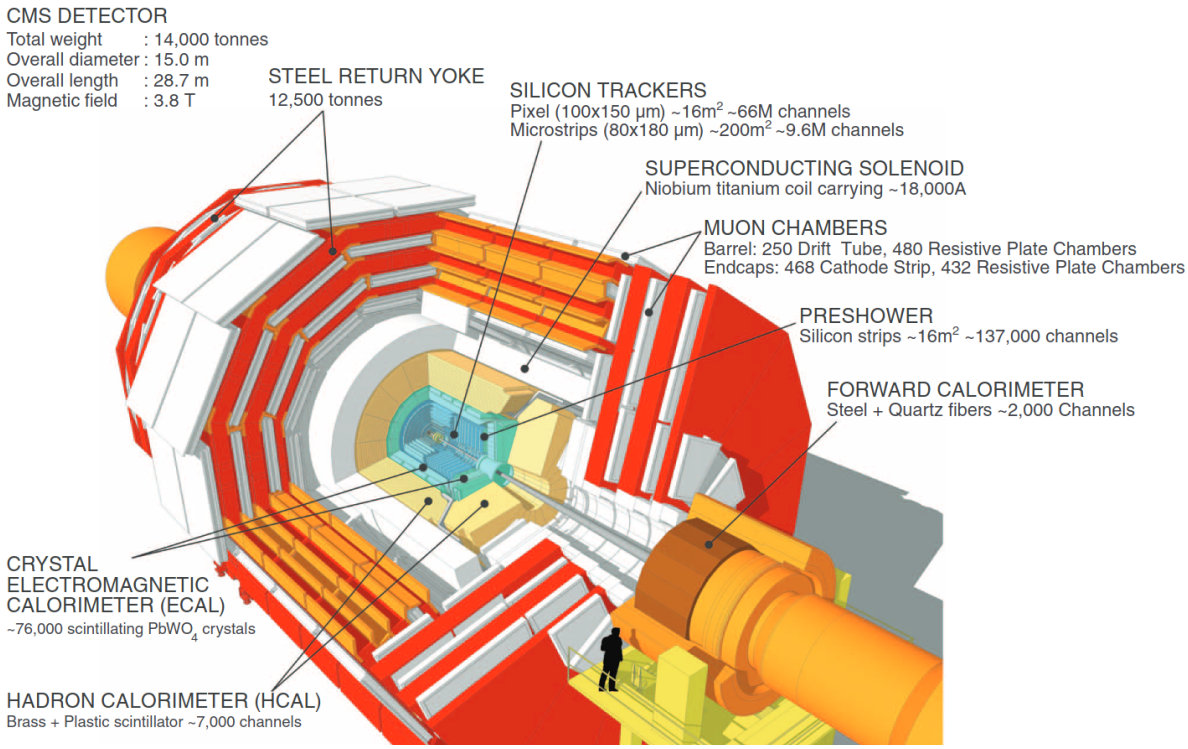


Figure 14: A schematic view of the CMS experiment, highlighting the different detector componenets. Source: [24]

4.2.1 Solenoid Magnet

One of the best ways to measure the momentum of charged particles is to measure their trajectory as they are deflected by a magnetic field, as the higher the momentum of the particle, the less it is deflected. One of the guiding decisions when designing the CMS experiment was that the experiment should have as large a magnetic field as possible, which leads to particles being deflected more, making the momentum measurement more accurate. This is achieved using the solenoid superconducting magnet in the experiment's name, which generates a field of 3.8 T along the length of the detector. In order to achieve this high field within technological and budget constraints, the magnet is relatively small for this type of experiment, with an inner diameter of 5.9m and a length of 12.9m, hence the experiment being described as "compact". The solenoid is surrounded by an iron return yoke, which both supports the detector and keeps the majority of the magnetic field contained within the experiment.

4.2.2 Tracker

The closest detector component to the interaction point is the tracker, which is a set of small silicon detectors which detect individual hits from charged particles. As the name suggests, these hits can then be combined to form "tracks" of charged particles. These are useful for two purposes: as discussed, measuring the degree to which the particles are deflected gives a measure of their momentum, and these tracks can also be traced back to identify which vertex the particles originated from. For most particles this is a collision vertex, which can be used to determine whether particles come from the high energy process of interest or one of the pile-up collisions. However if the particles come from the decay of a charm or bottom hadron, the tracks will originate from the point at which this hadron decayed, which, as discussed in section 2.5, can be displaced from the collisions, allowing identification of these hadrons.

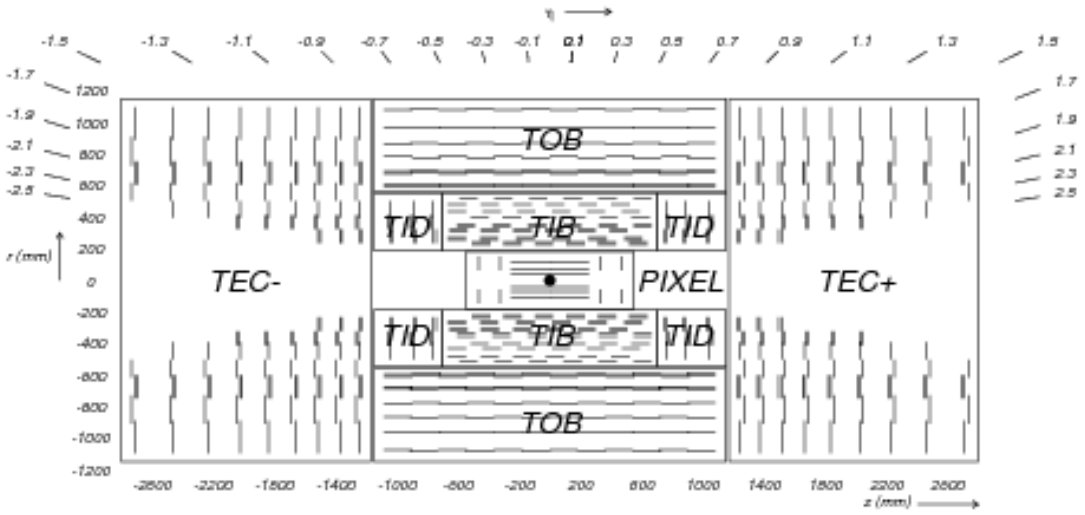


Figure 15: A diagram of the CMS tracker, showing the different silicon layers. Source: [25]

To achieve these aims, the tracker is formed of two different types of sensors: a pixel tracker close to the interaction point, with three layers in the barrel and two discs at each end, which is mainly focused on measuring the initial tracks to allow identification of the vertices, and a tracker constructed of silicon strips, which is designed to meausre the curvature of the tracks as they are bent by the magnetic field. This is divided into several components: the tracker inner barrel (TIB), tracker outer barrel (TOB), tracker inner discs (TID) and tracker endcap (TEC), the arrangement of which is shown in figure 15. This design is optimised to have sufficient layers of silicon detectors to give an accurate measurement of the particle tracks, while keeping the amount of material as low as possible, reducing the chance that the particles lose energy in the tracker, which would affect measurements in subsequent detectors. These detectors cover the entire azimuthal angle, and a pseudorapidity range of $|\eta| < 2.4$, with larger values of $|\eta|$ not

instrumented due to the very high radiation in this forward region.

4.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is designed to measure the energy of electrons and photons. When traversing matter, electrons will tend to emit energy in the form of photons (bremstrahlung), while high energy photons will convert in to electron-positron pairs, which can of course undergo further bremstrahlung, leading to a electromagnetic cascade. The CMS ECAL consists of lead tungstenate crystals ($PbWO_4$) which are opaque to x-ray light but transparent to visible light. Hence a photon or electron incident on the ECAL will produce an electromagnetic cascade leading to a flux of visible photons proportional to the momentum of the incident particle, which are measured by photodetectors at the back of the calorimeters.

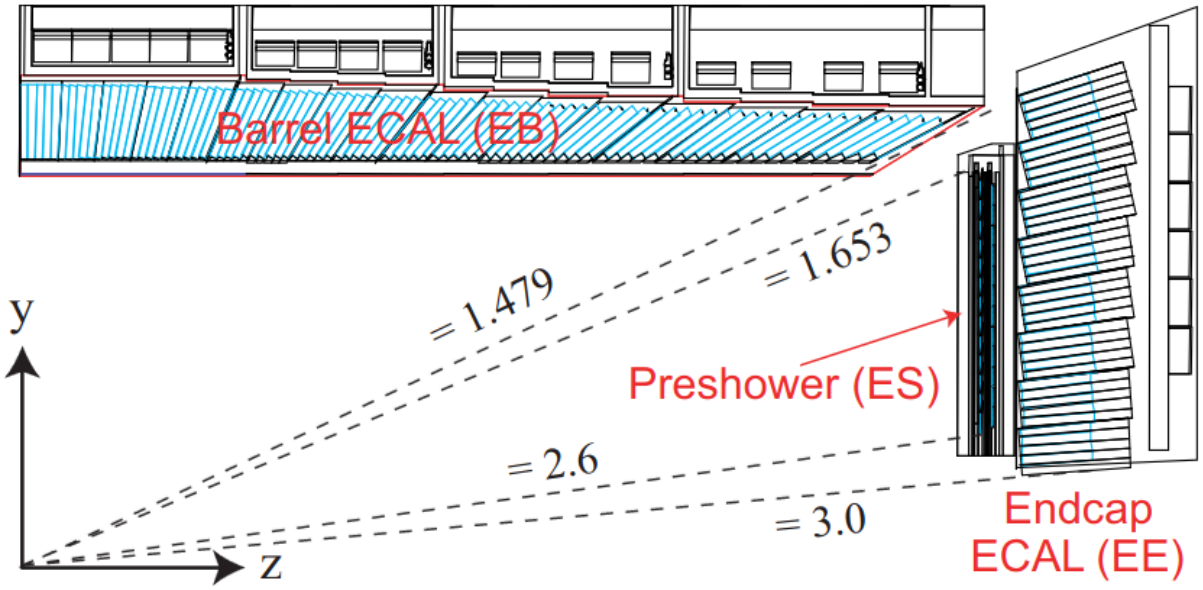


Figure 16: A longitudinal cross-section of the ECAL, showing the orientation of the crystals. Source: [26]

The ECAL layout is shown in figure 16. It is also organised into a barrel and endcap, with the crystals oriented to point towards the interaction point. These cover the range $|\eta| < 3.0$, however there is a small gap between the barrel and the endcap for the power, cooling and read-out electronics for the tracker, which leads to a narrow region of $1.4442 < |\eta| < 1.566$ not being fully instrumented. In front of the endcap there is a "preshower" covering the range $1.653 < |\eta| < 2.6$, composed of two alternating layers of lead and silicon detectors. This is designed to aid discrimination of photons from neutral pions, π^0, which are produced in very large numbers in these forward regions, and decay to two photons, and hence produce a very similar signal to genuine photons.

4.2.4 Hadronic Calorimeter

The energy of the hadrons is also measured by calorimetry. Since hadrons are significantly more massive than electrons or photons, they are much more penetrating, and hence only deposit a small amount of energy in the ECAL. The hadronic calorimeter (HCAL) sits outside of the ECAL, and is required to fit within the relatively small bore of the magnet. Hence a very dense material is needed so that the hadrons deposit all of their energy within this reasonably short distance. This leads to the choice of a sampling calorimeter, with layers of a brass or stainless steel absorber interspersed with thin layers of plastic scintilator. The absorber layers absorb the energy of the hadrons via a number of interactions and emit it as high energy photons, which are then converted to visible light by the scintilator, which can be measured by photodetectors.

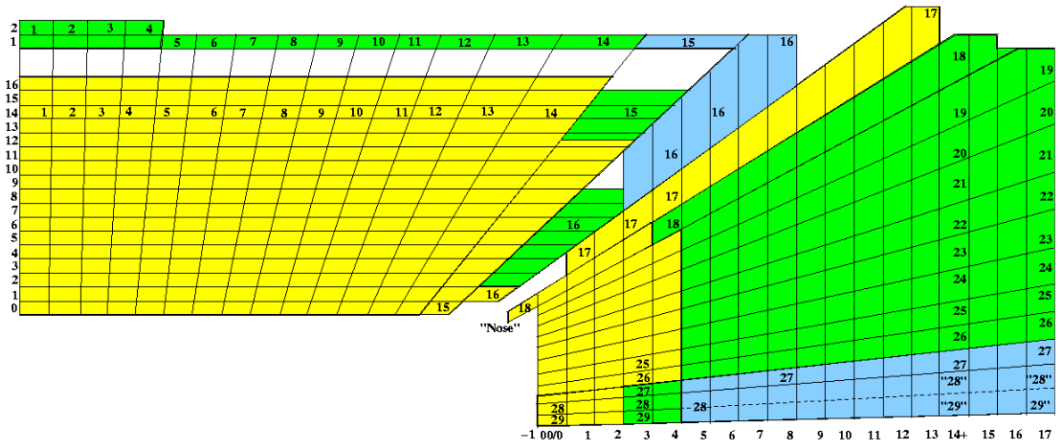


Figure 17: A longitudinal cross-section of the HCAL, showing the different towers. The readings from the layers of the same colour within each tower are combined for the level-1 trigger (see section 4.3.1. Source: [26]

The HCAL is segmented into "towers" in η and ϕ (figure 17). The barrel comprises towers 1-16 ($|\eta| < 1.3$), while the endcap comprises towers 17-29 ($1.3 < |\eta| < 3.0$). Although the design manages to absorb and measure most of the energy, some high energy hadrons are able to penetrate all the way through the central region of the calorimeter. There is therefore one further layer of scintilator outside the magnet in the range ($|\eta| < 1.26$), which measures energy deposited in the magnet. In addition, there is a forward calorimeter, covering the pseudorapidity range ($3.0 < |\eta| < 5.0$), as hadron colliders tend to produce a large amount of jets in the forward region.

4.2.5 Muon System

Muons are the most penetrating of all the particles which can be detected in the CMS experiment, and hence will tend to pass through the calorimeters and the magnet with relatively little loss of energy. As charged particles, one can measure their momentum

in the tracker, however to help identify them from other tracks and to improve the momentum measurement, a set of muon detectors are embedded in the iron return yoke of the magnet. The essentially perform the same role as the tracker, measuring the position of the muon tracks as they are deflected by the magnetic field in the iron return coil. However, since the area to be covered is much larger, it is not practical to instrument the entire region with silicon. Instead three technologies are used. Firstly, drift tubes (DTs) are used in the barrel region, $|\eta| < 1.2$, which are able to give a precise measurement of both the position and direction of the muons. In the endcap region ($1.2 < |\eta| < 2.5$) cathode strip chambers (CSCs) are instead employed, which have slightly worse position and direction resolution, but have much better response time, which is important in this region which has both a higher muon rate and higher fake backgrounds induced by the high rate of hadrons. In addition, these detectors are augmented across the entire range by resistive plate chambers (RPCs), which have the lower position resolution, but much faster response times than even the CSCs, which is important as these can be read out in time for use in the trigger system (see section 4.3.1). The arrangement of these various detectors are shown in 18

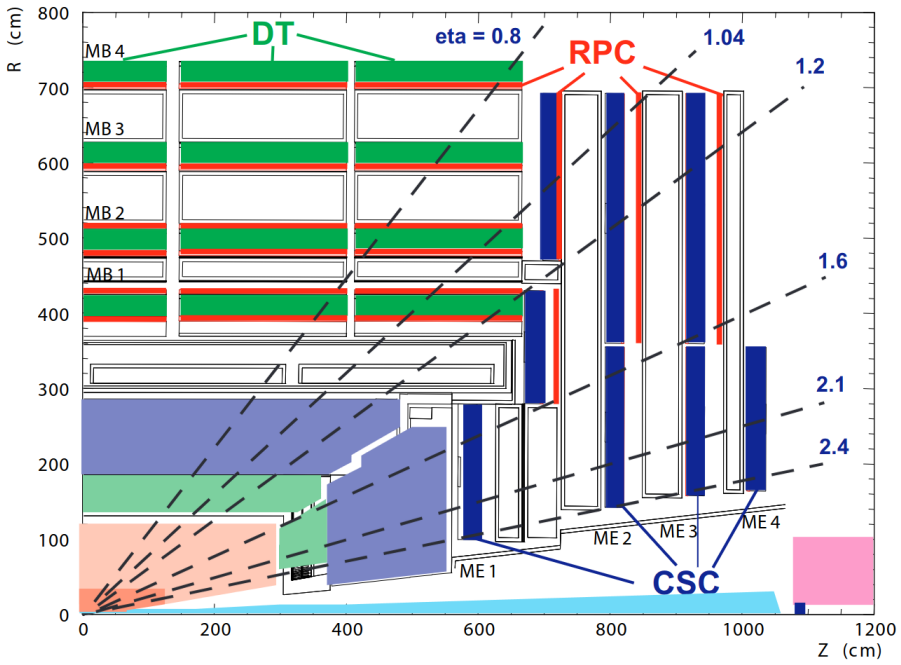


Figure 18: A longitudinal cross-section of the magnet return yoke, showing the positions of the various Muon detectors. Source: [26]

4.3 Event Reconstruction

4.3.1 Trigger

The LHC collides proton beams in the CMS experiment every 25 ns, which means that data is produced much faster than it is possible to read it out and store it. Furthermore, as has been noted, the processes of interest at the LHC happen much less frequently than the overall rate of pp scatterings, so it does not make sense to save all of the information, at considerable cost in computer storage, when much of it is of little interest to analysers. CMS therefore has a two tier trigger system, which selects which events should be recorded. The first tier is the level 1 (L1) trigger, which only receives information from the detector components which can be read out quickly (the ECAL, HCAL and muon RPCs) and uses field programmable gate arrays (FPGAs) to make a very quick decision on whether an event is potentially interesting, for instance if it appears to have one or more objects over a particular transverse momentum (p_T) threshold. This reduces the event rate from 40 MHz to less than 100 kHz. If an event passes the L1 trigger, the full detector information is read out from the buffers in which it is stored and passed to the high level trigger (HLT), which performs a full reconstruction of the event in software, and makes a decision on whether the events should be saved for analyses. Events are saved if they pass any of the possible selections on a "trigger menu", which is designed to select the events relevant to physics analyses, while keeping the overall number of events which need to be saved reasonably low. This "menu" therefore includes both selections for single objects with reasonably high p_T thresholds, as well as selections for events with multiple objects at lower thresholds, targeting specific physics processes. The rate of events selected by the HLT is about 1 kHz, and the full detector read-out for these events is then saved for later analysis.

4.3.2 Object Reconstruction

Once the full detector read out has been saved, the events are then "reconstructed" into objects corresponding to physical particles, which can be used in physics analyses. This is performed using the particle flow (PF) algorithm, which combines information from all detectors to get the best possible measurement of each particle. This is generally based on the track curvature for the charged particles, and calorimetry for the neutral particles which cannot be measured by track curvature. This reconstruction allows analyses to work not with individual tracks and calorimeter deposits, but with objects which correspond to the physics objects being studied, such as electrons, muons, jets and p_T^{miss}. The definitions of these objects in the PF algorithm are discussed here, while the exact cuts on the object used in the analysis will be discussed in section 6.5.

The momentum of these objects is usually measured in terms of the transverse momentum, p_T, azimuthal angle, ϕ, and pseudorapidity, η. The distance between objects, which is particularly important for determining if an object is well-isolated in the detector, is usually measured in terms of these angular coordinates, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

Electrons are usually identified as an energy deposit or set of deposits in the electromagnetic calorimeter which correspond to a track in the inner and outer trackers. Since electrons can lose a significant amount of energy in the tracker due to radiation of photons through bremstrahlung, nearby ECAL deposits are also clustered into a "supercluster", which is used to measure the total energy of the electrons including these emissions. The momenta is reasonably well measured, with an error between 1.7% (for electrons with relatively little bremstrahlung in the barrel) and 4.5% (for electrons in the endcaps with a lot of bremstahlung, or which fall between two ECAL towers) [27].

Muons are identified as a track in the tracker corresponding to hits in the muon system. Muons with sufficient hits in the muon system to form a track (typically the higher momentum ones) are referred to as "global" muons, with the others being "tracker" muons. Since their momentum can be measured from a charged track with relatively little bremstrahlung, muons are the best measured objects in CMS, with a momentum resolution of 1-3% [28].

Jets are formed using all of the available physics objects: electrons, muons, photons (ECAL deposits not matched to a track) and hadrons (HCAL deposits, which can be either charged (if matched to a track) or neutral). These are clustered using the anti-k_T algorithm [29], which iteratively clusters the nearby particles within a certain radius, with harder particles clustered earlier. This algorithm has the advantage of being infrared safe and having a shape which does not vary with soft radiation. A jet radius of 0.4 in the $\eta - \phi$ plane is selected, which is expected to be large enough to capture almost all of the partonic jet, while small enough to resolve jets coming from different partons in the hard process from one another. To mitigate the additional energy added to jets from pileup events, any charged hadrons with tracks originating from a primary vertex other than the main one are subtracted from the total energy of the jet [30]. Jet momentum is significantly harder to measure than electron or muon momentum, as there are many components close together, and neutral hadrons in particular can be difficult to measure as they deposit energy quite slowly and irregularly in the calorimeter. The momentum resolution is typically 10-20%, with better resolution for higher p_T jets [31].

The missing transverse momentum is the negative vector sum of the transverse momentum of all particle flow candidates in the event. As with jets, charged hadrons associated with pile-up vertices are omitted. This is the most difficult quantity to measure, as it is affected by the limited resolution of every other object (of which the worst measured are typically jets and the unclustered energy, which is the sum of the PF objects not associated with a jet), and by neutral hadrons from pile-up vertices, which cannot be distinguished from those coming from the primary vertex. The resolution is typically about 15-30 GeV, increasing with the total sum of hadronic energy in the event [32].

5 Monte Carlo Simulation

In order to interpret the large amount data collected from particle colliders, it is necessary to have some theoretical predictions of the various Standard Model processes, and the BSM models which are being searched for. For the complex collisions at the LHC where many particles are produced in each collision, it is not possible to predict the entire spectrum of particles which will be produced analytically due to the huge complexity of the final states, and instead Monte Carlo (MC) techniques are used. This approach approximates the integrals over the possible momenta of all final state particles in a probabilistic manner by producing individual events with the momenta of the different particles determined probabilistically. Not only is this the fastest numerical method for integrating over the large phase-space of all possible final state momenta, having predictions as individual events is useful for analysers, as when investigating new analysis variables or multivariate techniques, one does not need to obtain a new theoretical prediction for the variable of interest, but can instead simply create a histogram of this distribution for both the data and MC events, which serve as the theoretical prediction.

Since LHC collisions represent physics over a large range of scales, the generation of events is factorised into a number of steps. Parton distribution functions are used to determine which components of the protons interact; matrix element generators calculate the high energy interactions, general purpose event generators then account for additional emitted particles and other lower energy effects from QCD, and finally the interaction of the particles with the detector is simulated. These various steps are described in more detail in the following sections.

5.1 Parton Distribution Functions

As has already been discussed in section 2.5, the protons collided in the LHC are not fundamental particles, but bound states of two up quarks and a down quark. These three quarks (the valence quarks) are surrounded by a "sea" of low energy gluons which bind them together. Additionally, these gluons can split into quark-antiquark pairs, and hence other quark flavours than the valence quarks can exist on short timescales, which are referred to as the sea quarks. Therefore when protons are collided at the 13 TeV, which is much larger than the confinement scale of around 1 GeV, it makes sense to consider the collisions between these individual quarks and gluons, referred to as partons, rather than the entire protons. More formally, one can write the cross section for pp scattering in terms of a sum over all possible partons in the protons:

$$\sigma_{pp\to X}(p_1,p_2) = \sum_{i.j\in\{g,q,\bar{q}\}} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1,\mu_F) f_j(x_2,\mu_F) \sigma_{ij\to X}(x_1 p_1, x_2 p_2, \mu_F) \quad (14)$$

where $\sigma_{ij\to X}$ is the cross section for the partons to scatter to the given final state, x_1

and x_2 are fraction of momentum carried by the respective partons, and $f_i(x_1, \mu_F)$ is the parton distribution function (PDF) - the probability of finding a parton of flavour i with momentum fraction x_i. This splitting involves introducing a factorisation scale, μ_F, which is the energy at which the partons are sampled from the PDF. The partonic cross-section also has a dependence on this scale, as corrections from virtual partons or real emissions below this energy should not be considered in the calculation, since these are described by the PDF.

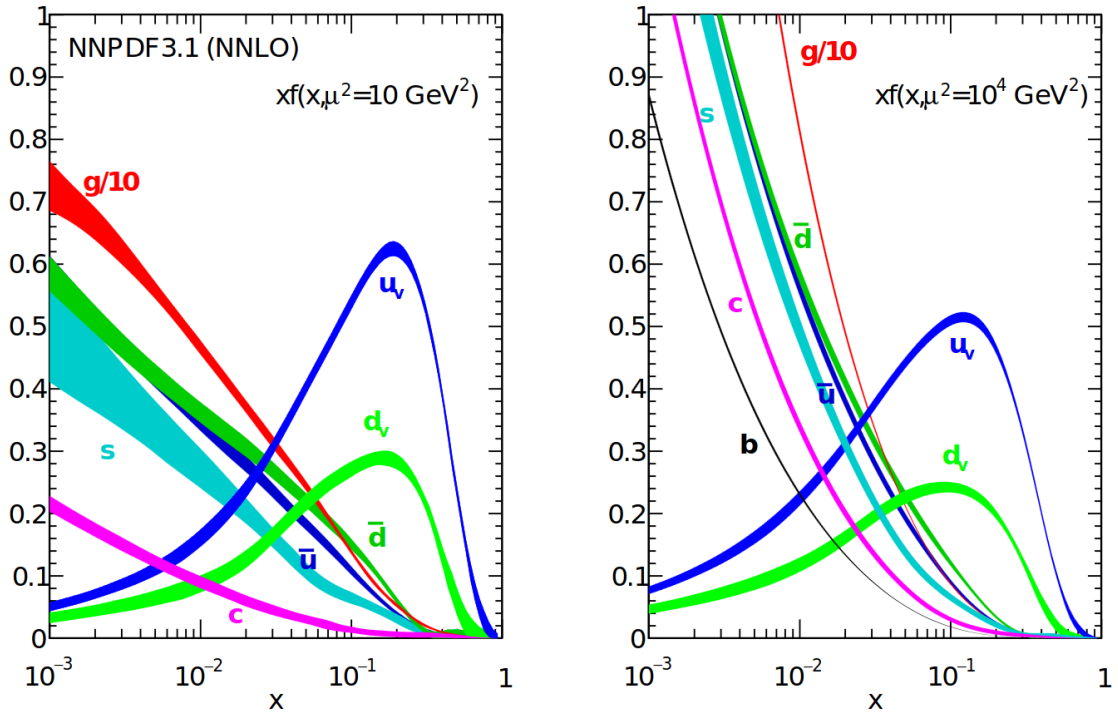


Figure 19: The values of the NNPDF v3.1 NNLO dataset as a function of momentum fraction, x. The width of the lines represents the uncertainty on each value. The lines for the strange, charm and bottom quarks also represent the corresponding anti-quarks, while the total pdf for up and down quarks is the sum of their valence contributions (u_v and d_v) and their sea contributions ($\bar{u}$ and $\bar{d}$). Source [33]

Since protons are highly non-perturbative systems, it is not known how to calculate the PDFs, which are instead fitted based on measurements from previous experiments, such as the H1 and ZEUS experiments at the HERA electron-proton collider, which was designed to study the structure of the proton, and more recent measurements from CMS and ATLAS. In this analysis, PDFs provided by the NNPDF collaboration are employed - version 3.0 [34] for 2016 data-taking, and 3.1 [33] for 2017 and 2018. The NNLO PDF from version 3.1 is shown for two different values of the factorisation scale in figure 19. These coincide with what one would expect from the structure of the proton - the valence quarks (u_v and d_v) are peaked at slightly less than 0.3, since these each carry

roughly 1/3 of the energy of the proton, with the up quark PDF approximately twice that of the down quark. The gluon and sea quark distributions are peaked towards low energies, and are much higher for the higher factorisation scale, as one expects to find more virtual particles at these higher energies. Note also that b quarks are only observed in PDFs at higher factorisation scales, due to their relatively high mass.

5.2 Matrix Element Generators

To calculate partonic cross sections, the first step is to calculate the probability of the high energy scattering, which is often referred to as the "hard process", in contrast to "soft" low energy corrections. This is performed by matrix element (ME) generators, which randomly generate events with momenta of the final state particles sampled from the possible final phase space, then give each event a weight, w, equal to squared amplitude of all possible Feynmann diagrams calculated for these momenta. An issue with this approach is that it gives the same statistical weight to the bulk of the distributions as the tails, when it is more useful to have higher statistical power (more events) in the bulk of the distribution. To resolve this, one can find the highest event weight, $w_{\max}$, and randomly keep each event with probability $w/w_{\max}$, then give each event an equal weight. This procedure is referred to as "unweighting".

There are a number of programs which can generate events at LO at ME level, but the main one used in this work is MadGraph5_aMC@NLO [35], which for any process can generate all possible Feynmann diagrams, then write code to calculate the squared amplitude. To match the precision at which many observables can be measured at the LHC, however, it is necessary to go beyond LO, and so many processes are generated at NLO in QCD. The MadGraph5_aMC@NLO program is also able to generate all of the NLO diagrams and calculate the corresponding amplitudes squared. The challenge when going to NLO however, is that the real emission diagrams can lead to double counting with emissions added by the parton shower (PS, see section 5.3.1), and hence possible emissions must be "matched". In MadGraph5_aMC@NLO this is accounted for by the automatic MC at NLO (aMC@NLO) method [36], which adds some negatively weighted events to cancel-out this double counting. These negatively weighted events do reduce the statistical power of the simulated samples, so an alternative code, the positive-weight hardest-emission generator (POWHEG) [37–40], was also developed, which, as its name suggests, uses an alternative matching scheme where all events have positive weight, though this does not automatically generate diagrams, so each process must be manually added. Generating and matching to NNLO predictions is an ongoing area of development, but will not be discussed further here as no NNLO samples were used in this work.

Whilst it is very challenging to produce events to more than NLO, it is possible to include part of these higher orders by considering diagrams where additional partons are produced, which is particularly useful for observables which are sensitive to these real emissions, such as the p_T of the entire system, X for any $pp \to X$ scattering. This

again results in an overlap with the emisions described by the parton shower, which can be accounted for by a procedure known as merging. In MadGraph5_aMC@NLO this can be done for LO simulation using MLM merging [41], while for NLO samples with additional partons the FxFx procedure both matches and merges the ME predictions to the PS.

5.3 General Purpose Event Generators

In addition to the process generated in the ME, there are numerous lower energy effects which mostly occur as a result of the strong force which are described by general purpose event generators: the additional radiation from coloured or charged particles is described by a parton shower, hadronisation describes the confinement of colour charged particles into hadrons, and finally the effects of additional soft interactions are added. This section will discuss two general purpose programs which implement all of these steps: Pythia 8 [42] and Herwig 7 [43, 44].

5.3.1 Parton Showers

The parton shower (PS) generates the additional low energy particles emitted from the particles involved in the hard process. Generally the probability to emit an additional particle is suppressed by a power of the coupling constant α, however additional radiation can be enhanced either if it has very low energy (soft radiation), or if it is emitted at a very small angle (collinear radiation), and in fact the probability of this radiation diverges as either of these variables tend to zero. However, the radiation in these divergent regions cannot be resolved, as a particle cannot be detected if it has zero energy or is parallel to another particle, and these divergences can therefore cancel with the ultra-violet divergences in loop diagrams in a full higher-order calculation. In parton showers, the probabilities of emissions are expanded around these poles, with "cut-offs" at a scale below the experimental resolution to avoid the divergences. This radiation occurs both for particles charged under QCD and EM, however it is much more significant for QCD since as discussed α_s is much larger than α_{EM}, and increases as one goes to the lower energies where this soft and collinear radiation occurs, so this description will focus on QCD.

Parton shower radiation is generally divided into two types: initial state radiation (ISR) and final state radiation (FSR). FSR is the emission of additional particles from the particles produced by the hard scattering, which leads to colour charged particles forming "jets", as the collinear radiation produces a number of particles travelling approximately parallel to each other. There are also some emissions at large angles, but these are only probable if the emitted particle is soft, and so generally do not have enough energy to form a new jet.

ISR, on the other hand, is emission from the particles which initiate the hard process, and

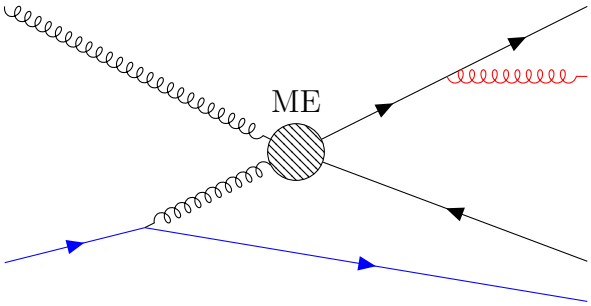


Figure 20: Two possible parton shower emisions from a $gg \to q\bar{q}$ scattering (where the hard process is shown as an effective vertex): ISR radiation of a quark (blue), and FSR radiation of a gluon (red).

considers whether the initial parton could have come from a splitting of an earlier parton with more energy, which must then be sampled from the pdf at a lower factorisation scale. Since the incoming partons generally have very high energies, the "soft" emissions can still have enough energy to form new jets, while collinear emissions result in a very high rate of jets in the high pseudo-rapidity regions. It is generally these emissions which can be double-counted with the additional hard partons added in the ME, necessitating matching and merging.

These two possible types of radiation can occur simultaneously, and the radiated particles can then emit further radiation, so it is important consider all possible emissions simultaneously. This is achieved via an evolution variable, which runs from the high scale of the hard process down to the lowest scale which can be described by the parton shower (the cut-off), with all particles able to radiate according to the probability for all possible emissions as a function of the evolution variable. In the default Pythia shower the evolution variable is the p_T of the emission, so higher transverse momentum emissions are generated before softer ones, while in the default Herwig shower the emissions are ordered by the angle of the emitted particle to the emitter, so large angle (generally soft) radiation is produced before the collinear emissions.

The probability of each splitting is given by a splitting kernel, which can be computed from the DGLAP equations [45–47]. To compute the next value of the evolution variable at which a splitting occurs, one can generate a uniformly distributed random variable, and convert this to a value of the evolution variable using the inverse of the splitting kernel. However these typically have complex functional forms, so instead an over-estimate which is simpler to invert is used, and then the splitting is retained by comparing a second random variable to the ratio of the actual splitting kernel to the over-estimate.

5.3.2 Hadronisation and Decays

The next step in generating an event is simulating the process by which the coloured quarks and gluons are confined into hadrons. This is inherently challenging, since QCD becomes non-perturbative in this region, and there is no fundamental theory which completely describes the process of hadronisation. Instead semi-empirical models are used, which are based on the expected structure of QCD in this energy range.

The main hadronisation model used in Pythia is the Lund string model [48, 49]. In this approach, quark - anti-quark ($q\bar{q}$) pairs are joined by QCD field lines, which form a very narrow "string" since the field lines interact with each other, unlike EM field lines, which spread out across space. If the relative momentum of the $q\bar{q}$ system is reasonably low, the string will be able to pull them back together as they move apart, and hence they will "yoyo" back and forth as a bound state, which can be interpreted as a meson. If however the particles are moving apart with high momentum, they will cause the string to break, resulting in a $q\bar{q}$ pair being produced, which then form two strings with the original $q\bar{q}$ pair, which, depending on the relative momentum of these new pairs, can either form bound states or fission again, leading to a line of mesons between the original quark and anti-quark. Gluons are then treated as kinks on the strings between the $q\bar{q}$ pairs. This model gives very good predictions of the momenta of the produced hadrons, however it has more difficulty predicting the flavours of the produced hadrons, since it is unclear how the bound "yoyo" states correspond to mesons, and hence there are a large number of parameters which need to be tuned to fit the data.

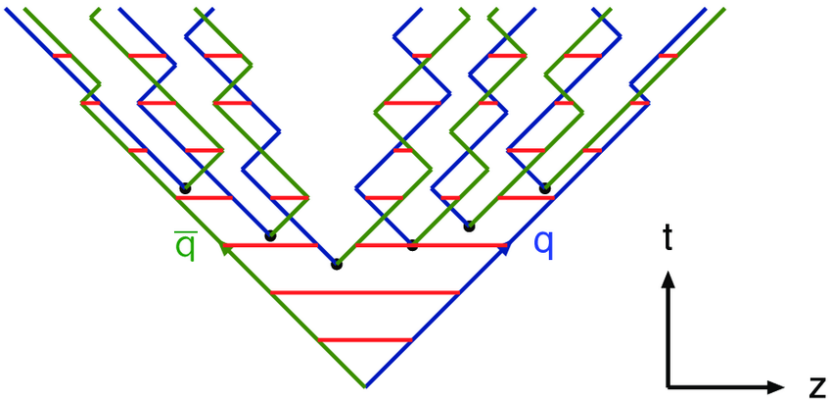


Figure 21: A sketch of the Lund string model. Quarks are shown in blue, anti-quarks in green; the red lines represent the string connecting these at various points in time. As the $q\bar{q}$ pair move apart they cause this string to split a number of times, until the $q\bar{q}$ pairs no longer have sufficient energy to split the strings, and instead "yoyo" around their common center.

Herwig, by contrast, uses the cluster model [50], which splits all the gluons into $q\bar{q}$ pairs, then forms clusters of colour connected $q\bar{q}$ pairs, which represent semi-bound excited hadron states. Each cluster then decays to two hadrons by producing either an

additional $q\bar{q}$ pair (for decays to two mesons) or an additional two quarks and two anti-quarks (for decays to a baryon - anti-baryon pair). All possible kinematically allowed hadrons are considered in this decay, weighted by the number of available spin states. A small number of clusters are too light to decay to two hadrons, and are instead decayed to a single hadron. This approach is effective at predicting the spectrum of hadron flavours with few empirical parameters. However, it has issues when when a $q\bar{q}$ pair is very separated in momentum space, leading to a cluster with an unphysically large momentum, which can occur if no gluons, which would break up the colour dipole, happen to have been emitted. These clusters are fissioned to lighter clusters, however the parameters for this fissioning have to be tuned to data to reproduce the observed hadron momentum distributions.

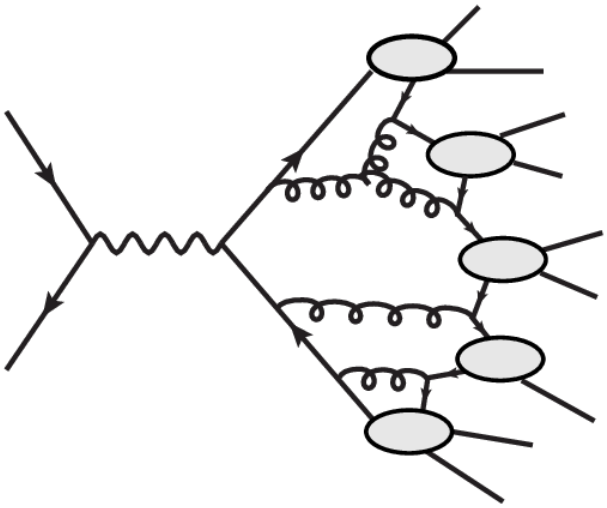


Figure 22: A sketch of the cluster hadronisation model. The gluons produced in the PS are split into $q\bar{q}$ pairs, which, together with the progenitor quarks, form clusters (represented as ovals), which each decay into two hadrons, except for the second cluster from the bottom, which decays to a single hadron.

Most of the hadrons produced in hadronisation are unstable, and can decay either to other hadrons, photons or leptons. Most of the lighter hadrons have already been well studied at lower energy experiments, so their decays are well known, and the event generators randomly decay them according to the measured branching ratios and models which describe the kinematics for the decay products (which can be non-trivial for three body decays). For some heavier resonances, notably $c\bar{c}$ and $b\bar{b}$ bound states (quarkonium), not all of the decay modes have been measured, so instead these states are allowed to decay to $q\bar{q}$ pairs or final states with gluons, which are treated again with the parton shower and re-hadronised to produce the light decay products.

5.3.3 Underlying Event and Pile-up

When protons collide, it is possible for more than one set of partons to interact, and in fact at high scales many of the PDFs, especially that for the gluon, reach values much higher than one for low momentum fractions, x, (figure 19), which means that one should expect any high energy collision to be accompanied by multiple lower energy scatterings, referred to as the underlying event. These are simulated inside the QCD event generators using dedicated MEs for QCD scatterings (which dominate over all other processes at these low energies), and the outgoing particles are evolved using the PS and hadronised as for the outgoing particles from the hard-process. Generally the outgoing particles from the underlying event are not of sufficiently high energy to produce jets which pass analysis selection criteria, but can contribute energy to existing jets.

In addition, as has already been discussed, any given bunch crossing includes numerous pile-up collisions. Each of these is simulated as a collision without a hard process, so only the underlying event scatterings are generated. The number of PU vertices and their positions are randomly generated according to the expected PU profile for the corresponding data taking period. Unfortunately it is very hard to predict the PU profile exactly, so in the analysis MC events are scaled in bins of the number of collisions to match the pileup in data (which is calculated from the measured instantaneous luminosity and the total $p\bar{p}$ cross section, which is found by dedicated studies to be 69.2 ± 3.2 mb [51]).

5.4 Detector Simulation

The particles produced in LHC collisions are unfortunately not recorded perfectly by the CMS detector: a particle might fly in a gap between two sensors, multiple particles might enter into the same detector, hits from a number of low energy particles in the tracker might conspire to fake a track which looks like a high energy particle, a particle might scatter in the tracker or against one of the support components, affecting measurements in later detectors, etc. To ensure all such possible effects are considered, the interactions of all the outgoing particles from the collision with the various detector components are simulated using the Geant4 program [52–54]. This provides events in the form of the same detector read-outs as the observed data, which can then be reconstructed into the same physics objects, as described in section 6.5. Some residual discrepancy with objects in data is observed due to factors which are difficult to simulate such as degradation of detector components over time. Corrections known as scale factors are derived based on studies of these differences in performance, and applied to the MC events to get as a close a prediction to the data objects as possible.

6 Analysis Design

6.1 Introduction

Searches for Dark Matter in collider experiments typically rely on identifying events with a set of visible objects which are expected to be produced in association with the DM, as well as a significant amount of missing transverse momentum, p_T^{miss}, due to the unobserved DM. In the next four chapters a search for the spin-0 mediator DM model introduced in section 3.4 is presented, targeting final states with two leptons (electrons and muons).

The first search for this model performed by CMS at $\sqrt{s} = 13$ TeV used only the data collected in 2016, and only considered the $t\bar{t}+\text{DM}$ signal [55]. This did not see a statistically significant excess, and follow-up analyses using the same dataset demonstrated that one could obtain better sensitivity by considering the $t+\text{DM}$ processes [56], and multi-variate analysis techniques in the dilepton channel [57]. More recently, a CMS search for stop quarks, the supersymmetric partners of top quarks, using the full Run II dataset published a reinterpretation in terms of this simplified DM model, which again only considered the $t\bar{t}+\text{DM}$ process [58], while the ATLAS collaboration published a search which considered both the $t\bar{t}+\text{DM}$ and $t+\text{DM}$ signals, but was only optimised to target the $t\bar{t}+\text{DM}$ final state [59]. Neither of these analyses observed a signal excess, and represent the most stringent limits on this model. This analysis will be part of a combined CMS search for both the $t\bar{t}+\text{DM}$ and $t+\text{DM}$ processes across all top decay final states with the full Run II dataset, which aims to be the most sensitive search for this model.

The two lepton final state corresponds to diagrams where dark matter is produced in association with either a top quark pair ($t\bar{t}$) or single top quark and W boson (tW), and all the top and W boson decays are leptonic. As discussed in section 2.6, the probability for a single top quark (or W boson) to decay to electrons or muons (including via tau decays) is about 26%, so the probability for this to happen for both top quarks (or the top and W boson) is about 6.7%. These final states also have either two ($t\bar{t}$) or one (tW) jets originating from b quarks. A challenge for this analysis is that in addition to these visible decay products, two invisible neutrinos are also produced, which means that the Standard Model (SM) $t\bar{t}$ and tW processes already tend to have a large amount of p_T^{miss}, so this variable alone is not very effective at distinguishing the signal from the SM background, and additional variables and multivariate techniques are required to improve sensitivity.

The analysis is described as follows; this chapter introduces the analysis, describing the software and datasets used, the triggers used to record the data, the definitions of the physics objects and the corrections applied to these objects and the events to remediate detector effects and mismodelling in the simulation. Additionally the selection cuts used to define the signal region (SR) are documented, and the neural network used in the final signal extraction is described, including the sensitive variables used and the

optimisation of the NN architecture. During the optimisation phase of the analysis this signal region was "blinded", i.e. only a small fraction of the data (in this case 1 fb^{-1} per year) was used to reduce the risk of being biased by the data distributions when introducing corrections. To ensure good modelling of the data by the MC simulation a further $t\bar{t}$ validation region (VR) was introduced, which has a significantly lower signal contribution relative to the backgrounds, making it possible to check the background modelling without potentially being biased by the signal. This is described in chapter 7, together with two control regions (CRs) which are used in the final fit to improve the estimation of the Drell-Yan and $t\bar{t}Z$ backgrounds. The systematic uncertainties which affect this measurement and the process of estimating these is described in chapter 8, and chapter 9 describes the statistical fit used to determine the presence or absence of the signal, and presents the results of this fit.

6.2 Analysis Tools

The data recorded from the detector (as described in chapter 4) or produced by Monte Carlo (MC) simulation (see chapter 5) is made available to analysers in three data formats: Analysis Object Data (AOD), MiniAOD and NanoAOD, with MiniAOD and NanoAOD being succesively skimmed collections of AOD, dropping a large amount of the low level information, such as tracks and calorimeter clusters, in favour of high level objects such as electron, muon and jet four-momenta. Since this analysis can be performed with these standard physics objects, it was performed with the most lightweight data tier, NanoAOD, which stores only about 1kB information per event. This has three main advantages: firstly that the event files are sufficiently small that all the files necessary for this analysis can be stored on-site at the DESY computing facility, secondly that the information is stored in standard ROOT TTrees (as opposed to the other data-tiers, which require dedicated CMS software to read), and finally that a number of the most common corrections calculated centrally in CMS are already applied.

To extract meaningful physics results from the very large input datasets produced from the LHC a number of code packages are needed, but the final physics analysis is typically conducted with an "analysis framework", which implements the analysis selection, applies any corrections and systematic uncertainties to the data, calculates derived variables and produces histograms or a dataframe of information for further processing, e.g. Neural Network training. Traditionally, frameworks have followed an event-loop paradigm, i.e. the code performs all of the previous steps for a single event, then proceeds to the next one. This approach works well in a compiled language such as C++ on traditional CPUs. However in Python iterating over events in this way is very slow, and instead it is recommended to use packages such as numpy [60], which work with arrays of values rather than looping over them. This "Single Instruction, Multiple Data" approach is well suited to modern computer architechtures with multiple cores, and enables Python programs written in this way to reach comparable speeds to compiled C++ code. It is also well suited to the ROOT file format, which allows data to be read in chunks

from each "column", rather than giving every column for a given event. However, there is a problem as numpy and similar packages require that the arrays are regular, i.e., every event should have columns of the same dimension, which is not true for particle physics events, which contain variable numbers of particles and hence columns of the variables relating to these particles are "jagged". To address this problem the awkward package [61] was developed, which provides similar functionality to numpy, but for these "jagged" arrays. The coffea (COlumnar Framework For Effective Analysis) package [62] extends this by offering a number of utilities for analysing NanoAOD in Python using this "columnar" style.

For this analysis a new framework, called "Pepper" (ParticlE Physics ProcEssoR) [63], was developed using this columnar analysis technique. The design philosophy was to create a user interface that was simple without being restrictive, as well as using the readability of the Python language to make it transparent what analysis steps were being performed, even without full knowledge of the code. To this end the code takes a JSON format configuration input, in which many analysis parameters, as well as the histograms to be produced and the per-event output information can be specified. The basic event structure is also defined in a single function, which should call other functions to implement cuts and add derived variables. Many such functions are already included in the basic processor for common cuts and variables, and the user can easilly add more functions specific to their analysis. The more technical details, such as the code to allow the framework to be run on the DESY batch cluster, or for creating and filling histograms for each of the cuts added to the selector, are designed to not require intervention from the user. These design choices have had a positive effect, with the Pepper framework being used for a number of analyses in the DESY group, and users commenting that it makes it simple to implement their physics analyses, even with limited experience.

6.3 Datasets

This search was performed using data collected in Run 2 of LHC operation. It is divided by the years in which the data was taken, and corresponds to 36.33 fb^{-1} in 2016 [64], 41.53 fb^{-1} in 2017 [65] and 59.74 fb^{-1} in 2018 [66] for a total of 137.6 fb^{-1}. Since the signal processes contain leptons, the datasets selected using the single and double lepton triggers are used, as will be discussed in section 6.4.

This data is compared to simulated events produced by MC event generators, as discussed in Chapter 5. The predominant backgrounds for this search are the SM $t\bar{t}$ and tW processes, which are calculated at NLO using POWHEG [67]. Another major background is the Drell-Yan process (production of a lepton anti-lepton pair via either a photon or a Z-boson mediator), with additional jets from initial-state radiation. This process has a high cross-section, but does not contain any invisible particles which will give a source of p_T^{miss}, so in order for these events to form a background for the high p_T^{miss} signal, the events must have p_T^{miss} due to neutrinos produced from hadron decays or mismeasurement of the jet energy. Since these processes are not better described by

a more accurate matrix element calculation, it was decided to use LO samples of this process produced by MadGraph5_aMC@NLO with up to four additional jets in the ME to maximise the statistical power of the sample. Further backgrounds considered were diboson processes, where one or both of the W and Z bosons can decay to visible leptons and/or neutrinos to produce signal-like events. These were produced at NLO using either POWHEG [68, 69] or MadGraph5_aMC@NLO[35]. A final background considered was $t\bar{t}X$, where X is a W, Z or Higgs boson, which are generated at NLO using either MadGraph5_aMC@NLO or POWHEG [70]. These processes are mostly minor backgrounds due to their low cross-section, however for the case that X is a Z boson decaying to two neutrinos this can be a significant irreducible background since it has a near identical signature to the signal with the top decay products and additional p_T^{miss} due to neutrinos, as will be discussed in section 7.2. All samples use Pythia 8 [42] for parton showering and GEANT4 [52–54] for detector simulation. For 2016 simulation the NNPDF3.0 [34] PDF set was used, and the Pythia underlying event parameters were set using the CUETP8M1 tune [71] (which was produced using CMS data from Run 1, i.e. $\sqrt{s} = 7$ and 8 TeV collisions, as well as data from previous experiments), while for 2017 and 2018 the NNPDF3.1 [33] PDF set and CP5 tune [72] (which also includes measurements on 2015 and 2016 CMS data) were used. A list of the processes considered, and their cross-sections, is given in table 1.

Since the search is not highly sensitive to the exact modelling of the signal shapes, the signal samples were only generated at LO, though with one additional jet included in the matrix element to allow an easier future comparison with the NLO process. The cross-sections for the different signal samples are given in table 2.

6.4 Trigger

As already discussed in section 4.3, only events which pass CMS's two-tier trigger system are stored, reconstructed and made available for analyses. Since the signal for this analysis includes two leptons, dileptonic triggers were used to select events. These have the advantage of reasonably low p_T thresholds, which are possible since the rate of two lepton production is relatively low compared to the overall pp cross section. However, due to the limited information in the online reconstruction, sometimes one of the leptons is not correctly reconstructed, and hence only 92% of events passing the two lepton requirement of this analysis pass the trigger. To recover some of the additional events which do not satisfy the dilepton trigger requirement, single lepton triggers were also included. Although these have higher p_T thresholds, they are still able to capture many of the events failing the dilepton triggers, giving an overall trigger efficiency of 98%. The full list of triggers used in this analysis is shown in table 3. These involve both cuts on the p_T of the leptons as well as quality cuts- for muons only cuts on the isolation and the distance from the primary vertex, Δ_z, are employed, for electrons a set of ID cuts similar to those used in the selection for this analysis, described in section 6.5.1 are employed - the exact values are given in reference [73]. The exact values of the cuts

Process	Sample	Cross section [pb]
Drell-Yan	Z+jets, $Z \to l\bar{l}$ $(10 < m_{ll} < 50\text{GeV})$	18610.0
	Z+jets, $Z \to l\bar{l}$ $(m_{ll} > 50\text{GeV})$	6077.22
	$Z\gamma$, $Z \to l\bar{l}$	55.48
$t\bar{t}$, dileptonic	$t\bar{t}$, dileptonic	87.310
	$t\bar{t}\gamma$	3.697
Single top	$\bar{t}W^-$	35.85
	tW^+	35.85
	$t/\bar{t}$ s-channel	3.360
	$\bar{t}$ t-channel	80.95
	t t-channel	136.02
	tWZ	0.0112
	$t/\bar{t}\gamma$	3.697
$t\bar{t}$, semileptonic	$t\bar{t}$, semileptonic	365.35
ttV	$t\bar{t}Z$, $Z \to l\bar{l}/\nu\bar{\nu}$ $(m_{ll} > 10\text{GeV})$	0.2814
	$t\bar{t}Z$, $Z \to l\bar{l}$ $(1 < m_{ll} < 10\text{GeV})$	0.0822
	$t\bar{t}Z$, $Z \to q\bar{q}$	0.5297
	$t\bar{t}W$, $W \to l\nu$	0.2043
	$t\bar{t}W$, $W \to q\bar{q}'$	0.4062
	$t\bar{t}H$, $H \not\to b\bar{b}$	0.2118
	$t\bar{t}H$, $H \to b\bar{b}$	0.0328
VV	$WW \to l\bar{l}\nu\bar{\nu}$	12.178
	$WZ \to lll\nu$	4.42965
	$WZ \to l\bar{l}q\bar{q}'$	5.595
	$ZZ \to l\bar{l}\nu\bar{\nu}$	0.5640
	$ZZ \to l\bar{l}q\bar{q}$	3.22
	$ZZ \to l\bar{l}l\bar{l}$	1.382
Rare	tWZ, tW dileptonic, $Z \to l\bar{l}/\nu\bar{\nu}$	0.0030
	tHW	0.147
	tHq, $Z \to l\bar{l}$	0.071
	tZq, $Z \to l\bar{l}$	0.0942

Table 1: Monte Carlo backgrounds and cross-sections

varied slightly between the years to fit within the rate requirements for these triggers while selecting as many events as possible; however in general the cuts are significantly looser than those described in section 6.5, so these changes have minimal impact.

In order to keep the treatment of data and MC simulation as comparable as possible, the same trigger algorithms applied to data are used on MC events. However the efficiency of these algorithms is found to differ slightly between data and simulation, with the

	Cross section [fb]			
Mediator Mass [GeV]	$tW + \phi$	$tW + a$	$t\bar{t} + \phi$	$t\bar{t} + a$
50	13.2	4.95	34.4	340
100	5.63	3.66	21.6	80.3
150	3.4	2.75	14.1	26.7
200	2.36	2.1	9.77	11.6
250	1.74	1.62	6.75	6.02
300	1.33	1.25	4.81	3.58
350	1.02	0.785	2.74	2.38
400	0.672	0.437	1.41	1.44
450	0.454	0.31	0.93	0.902
500	0.321	0.232	0.662	0.62

Table 2: Cross-sections for signal processes including a scalar (ϕ) and pseudoscalar (A) mediator. All cross-sections are for dileptonic decays.

Lepton flavour	Leading p_T (GeV)	Sub-leading p_T (GeV)	Other requirements
ee	23	12	Loose Calorimeter and tracker IDs, Very loose isolation
$e\mu$	23	8 (2016), 12 (2017, 2018)	Electrons: Loose Calorimeter and tracker IDs, Very loose isolation, Muons: Very very loose tracker isolation, $\Delta_z < 0.2$ cm from primary vertex (2016 era H, 2017, 2018)
μe	23	12	
$\mu\mu$	17	8	Very very loose tracker isolation, $\Delta_z < 0.2$ cm from primary vertex (2016 era H, 2017, 2018), $m_{ll} > 8$ GeV (2017) / 3.8 GeV (2018)
e	27 (2016), 35 (2017), 32 (2018)	-	Tight cut
μ	24 (2016, 2018), 27 (2017)	-	Loose isolation cut

Table 3: Triggers used to select events for this analysis. Electron quality cuts are described in reference [73].

efficiency for data generally slightly lower due to detector effects. In order to correct for this difference, scale factors were calculated following the method described in reference [74]. In this procedure, the trigger efficiency was computed for events which passed the lepton selection for this analysis, and a scale factor calculated as ratio of the trigger efficiency in data divided by that in MC simulation. This was done separately for each year, for each combination of lepton flavours (ee, $e\mu$ and $\mu\mu$) and in bins of the p_T of the two leptons. Of course, for the data events used in this analysis the trigger requirement is necessary for them to be recorded, so it was necessary to find a sample of data events with a high fraction of two lepton events which had been recorded with a different trigger. Fortunately, dileptonic $t\bar{t}$ events tend to have a fairly large amount of p_T^{miss}, so one can use events recorded by the trigger for high p_T^{miss} events to get a sample sufficiently enriched in dileptonic events to use for this purpose. The trigger efficiencies for 2018 data-taking are shown in figure 23.

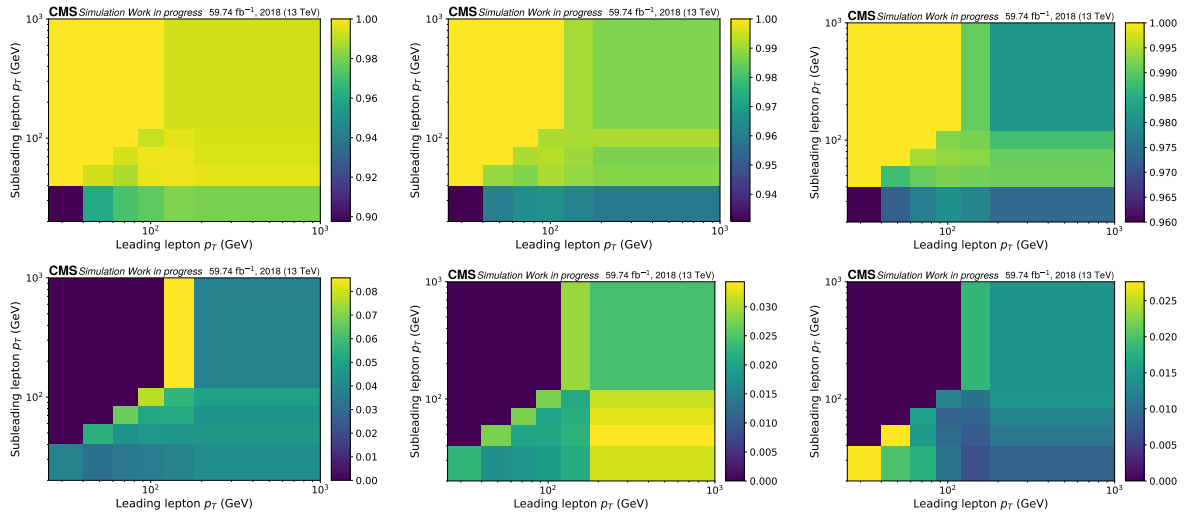


Figure 23: Trigger efficiencies (top row) and uncertainties (bottom row) in the dileptonic channel for 2018 data taking as a function of leading and subleading lepton p_T, in the ee (left), $e\mu$ (center) and $\mu\mu$ (right) channels. Note the upper left region (subleading p_T > leading p_T) is filled with dummy values of 1 (for efficiencies) and 0 (for uncertainties).

These scale factors were then calculated for each separate era of data taking within the year, with the maximum of the difference of these from the values calculated for the whole year taken as a systematic uncertainty due to potential degradation of detector response throughout the year. Similarly, scale factors were calculated in regions of jets ($<$ or $\geq$ 3), MET ($<$ or $\geq$ 150 GeV) and the number of primary vertices ($<$ or $\geq$ 30) with the difference from the nominal value taken as a systematic uncertainty on the correlation of the SFs with each of these variables. Finally, potential correlations between the dileptonic triggers and the p_T^{miss} cross-triggers were estimated from MC simulation as:

$$1 - \frac{\epsilon(\text{lep})\epsilon(p_T^{\text{miss}})}{\epsilon(\text{lep}, p_T^{\text{miss}})} \tag{15}$$

where ϵ(lep) is the efficiency of the combination of dilepton and single lepton triggers used in the analysis, $\epsilon(p_T^{\text{miss}})$ is that for the p_T^{miss} triggers, and $\epsilon(\text{lep}, p_T^{\text{miss}})$ is the efficiency for both together. These different sources of uncertainty were then added in quadrature to give an overall systematic uncertainty for the scale factors, though this was not found to be significantly higher than the statistical uncertainty for any values. The scale factors for 2018 are shown in figure 24, together with the total (statistical + systematic) uncertainties on these values.

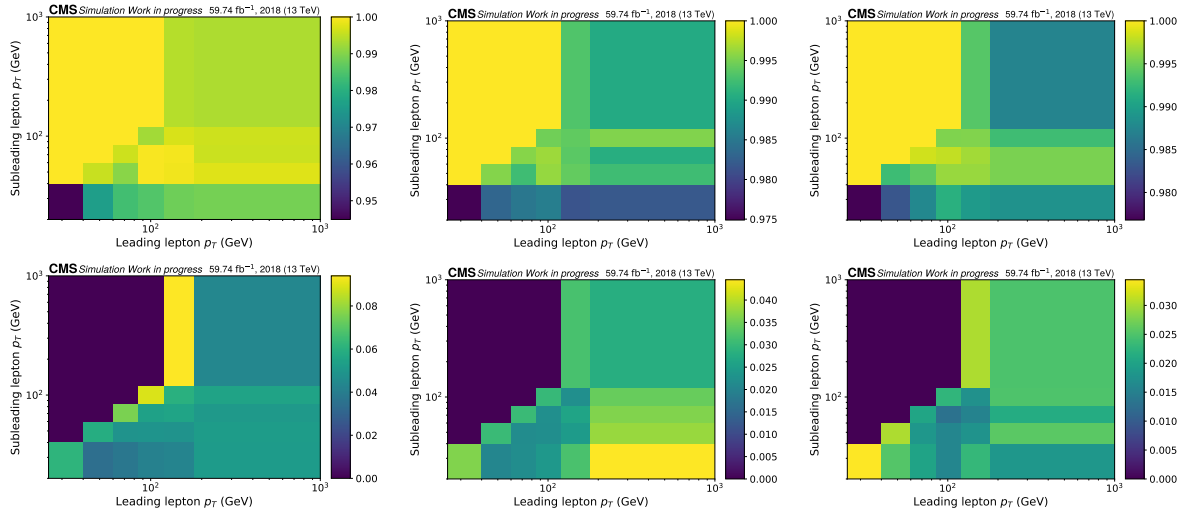


Figure 24: Trigger scale factors (top row) and uncertainties (bottom row) in the dileptonic channel for 2018 data taking as a function of leading and subleading lepton p_T, in the ee (left), $e\mu$ (center) and $\mu\mu$ (right) channels. Note the upper left region (subleading $p_T >$ leading p_T) is filled with dummy values of 1 (for scale factors) and 0 (for uncertainties).

6.5 Object Selection and Corrections

As discussed in chapter 4, the various visible particles which impact on the detector can be observed and identified by their different interactions with different detector components. For this analysis four types of objects are relevant: electrons, muons, hadronic jets (some of which can be "tagged" as originating from a b quark) and p_T^{miss}. The defintions used for each of these objects have been selected from a range of options developed by dedicated groups within the CMS experiment, and are designed to supress fakes while maintaining a good rate of acceptance for genuine objects.

6.5.1 Electrons

Electrons identified by the particle flow algorithm are required to fall within the coverage of both the tracker and the ECAL, i.e. $|\eta| < 2.4$, and to have $p_T > 10$ GeV. Additionally, the corresponding supercluster should not fall in the transition region between the ECAL barrel and endcap, $1.4442 < |\eta| < 1.566$

To help reject fake electrons, which can come from either misidentified hadronic jets or pile-up events, further sets of cuts, referred to as "cut based IDs" are developed. These consist of cuts on the distance between the track and the calorimeter; the ratio between the activity in the HCAL and ECAL (electrons are expected to deposit the majority of their energy in the ECAL); the difference between the energy, as measured in the calorimeter, and the momentum, measured from track curvature; the number of hits

missing in the track, and the isolation from other energy deposits. Of the available IDs, the medium ID is used in this analysis, which has an efficiency of around 80%, as this was found to provide good background rejection while still having a reasonably high efficiency.

The probability for a electron to be reconstructed in the detector, and for it to pass the ID requirements, differs slightly between observed data and MC simulation. To correct for this effect, scale factors, binned in p_T and $|\eta|$, are applied to the MC to correct for the different reconstruction and identification efficiencies. The reconstruction SFs vary between 95% and 99%, with uncertainties $< 1\%$, while the ID SFs vary between 95% and 102% , with uncertainties varying between 0.01% and 7% [75].

6.5.2 Muons

Muons are required to fall within the tracker and muon systems, $|\eta| < 2.4$, and have $p_T > 10$ GeV. Cut based IDs based on the goodness-of fit of the tracks in the tracker and (if present) muon system are again developed to reject fakes. However, when identifying muons the main problem is not fakes, but "non-prompt" muons coming from processes other than the hard process, typically decays of B-mesons. To reject these, additional isolation requirements, rejecting events with sizable energy deposits in the calorimeters within $\Delta R < 0.4$ of the muon are used. For this analysis, the medium ID [28] was used, which has an efficiency of about 98% since the muons are generally easy to identify and have few backgrounds. For the isolation requirement the tight working point [28] is used, which has an efficiency of about 95%. Scale factors are applied to correct for differences in the efficiencies of the identification and isolation cuts [76] - due to the excellent muon identification of CMS the ID SFs vary between 99% and 97% with uncertainties <0.1%, while the ISO SFs vary between 97% and 102% with uncertainties <1%.

6.5.3 Jets

As discussed in section 4.3.2, jets are clustered using the anti-k_T alorithm with a radius of 0.4. The energy of the measured jets is found to not be very accurate, and therefore corrections are applied to the measured energy in both data and simulation [77]: the first corrects for additional energy added to jets by pileup, while the second set corrects for the inconsistent detector response. A small further correction is applied to data to correct for the differences in actual detector response compared to the response in simulation. Additionally, the detector response is found to be more variable in data than simulation, so simulated jets are "smeared" to match the lower resolution of data. This is done in one of two ways: for jets within $\Delta R < 0.4$ of a generator level jet, the difference in p_T between the measured and generator-level jet is scaled by a stochastic factor, otherwise the p_T of the jet overall is simply scaled by a stochastic factor, with the random factor in either case sampled from a normal distribution with standard deviation based on the uncertainties of the jet measurement for different regions of the detector.

For this analysis, the "tight-lep-veto" cut-based ID [78] is applied to reject fake jets arising from detector noise, which operates by vetoing events which are dominated by activity ($> 90\%$) in a single calorimeter without corresponding charged tracks, as well as jets containing lepton candidates. Jets dominated by contributions from pile-up interactions are vetoed using a tight PU ID [79] produced by a boosted decision tree (BDT) trained on variables such as the consistency of the charged tracks in the jet with the primary vertex and the shape of the jet in the calorimeter (which will be different for a jet formed by the overlap of several pileup jets and a genuine jet from the hard interaction). This PU ID has slightly different performance in data and MC, however since each event can contain multiple jets (including some fakes from PU) and the efficiency of the ID can vary according to the position and momentum of the jet, a slightly more complicated procedure is used than for the lepton SFs. Firstly a general scale factor was calculated as the probability for a jet to pass the ID in data divided by that in simulation, then the efficiencies, ε_i of a jet passing this ID were calculated after the lepton selection of this analysis, before the first cuts on jets (as described in section 6.6), both as a function of p_T and η. This was done both for genuine jets (identified as jets within $\Delta R < 0.4$ of a generator jet, i.e. a jet coming from the primary collision after the parton shower, before PU and detector simulation) and for the mis-ID SFs and efficiencies of PU jets. Each event is then weighted by the ratio of the probabilities for the given combination of passing and failing the IDs in data and MC:

$$w = \frac{P(\text{data})}{P(\text{MC})} = \frac{\prod_{i\in\text{jets, pass PU ID}} \text{SF}_i\varepsilon_i \prod_{i\in\text{jets, fail PU ID}}(1-\text{SF}_i\varepsilon_i)}{\prod_{i\in\text{jets, pass PU ID}} \varepsilon_i \prod_{i\in\text{jets, fail PU ID}}(1-\varepsilon_i)} \tag{16}$$

Jets within $\Delta R < 0.4$ of electrons or muons passing the selections for this analysis are vetoed to avoid double counting these objects (since the energy deposits from leptons can also be clustered into a jet). Finally jets are required to have a $p_T > 20$GeV, and fall within the tracker and calorimeters, $|\eta| < 2.4$.

6.5.4 B-tagged Jets

As has already been discussed, the t + DM and $t\bar{t}$ + DM signal processes contain b quarks from the top decays. Since b quarks are quite heavy and form relatively long lived mesons, they can give rise to jets whose tracks do not seem to originate from the main interaction vertex, but from a "secondary vertex" a few mm from the primary vertex. In this analysis the CSVv2 (Combined Secondary Vertex, version 2) algorithm [80], which is a BDT trained on a number of variables derived from the tracks in the jet and the corresponding secondary vertex, is used to discriminate b-jets from jets coming from lighter quarks and gluons. A working point is selected at which the algorithm rejects 99% of "light flavour" (up, down and strange quark, plus gluon initiated) jets; the rejection probablity for charm initiated jets is much lower (about 85%) as charm mesons are also relatively long lived and hence can also give displaced secondary vertices, however none of the major backgrounds for this analysis feature charm quarks, so this

is not a major issue. This working point has an efficiency of about 60% for genuine b jets, and jets which pass this working point will henceforth be referred to as "b-tagged". The efficiency of this algorithm, as well as the mistagging rate for non-b jets, is found to not be consistent between data and simulation, and therefore scale factors [81], which are calculated using a sample of jets enriched in soft muons, are applied to simulated events in the same way as for the PU ID SFs (see equation 16), with (mis-)tagging efficiencies calculated after all the steps of the event selection before the b-tag cut have been applied.

6.5.5 Missing Transverse Momentum

The p_T^{miss} is also taken from the PF algorithm; all of the energy corrections applied to jets in section 6.5.3 are propagated to the p_T^{miss}, as well as a dedicated correction for the unclusterred energy. Further "xy-shift" corrections are then applied, which correct for the fact that the p_T^{miss} distribution is not homogeneous in ϕ, but rather shifted to differing degrees in data and MC due to the fact the beamspot is off-center and the detector response is not entirely homogeneous.

In 2017 data-taking additional mis-modelling of the p_T^{miss} was observed due to excessive noise in the forward ECAL endcap (EE) [82]. Since this issue predominantly affected lower energy jets and unclustered energy deposits, which only give a relatively small contribution to the p_T^{miss}, it was decided to exclude all jets with $p_T < 50$ GeV in the region $2.65 < |\eta| < 3.139$, as well as unclusterred calorimeter deposits in the same region, from the p_T^{miss} computation. This noise was also observed in 2018 data-taking, but accounted for by scaling down the energy measured in this region of the EE when computing the jet energy in the particle flow algorithm, so did not cause the same data-MC disagreement.

6.6 Event Selection

After applying the triggers described in section 6.4, the dataset is dominated by events with quite different topologies to the signal process, some of which, such as QCD jet production including a "fake" lepton, are not very well-modelled. Therefore a series of cuts are applied which dramatically reduce such backgrounds while having high efficiency for the signal process, giving a well-modelled dataset on which machine learning methods can be used to extract the most signal-like events.

The first step of this selection is to require that all events have two leptons (electrons or muons; hadronically decaying τ leptons are not considered since they are much harder to identify from the QCD jet background). Additionally the leptons are required to have opposite charge (as is required for leptons coming from a $t\bar{t}$ or tW decay), and the highest p_T lepton is required to have $p_T > 25$ GeV, while the subleading one is required to have $p_T > 20$ GeV. These cuts help to dramatically supress a number of backgrounds,

including QCD jet, W boson and semileptonic $t\bar{t}$ production, which are all reduced to the point they will not be discussed further. Events with more than two leptons are vetoed to reduce minor backgrounds such as diboson and ttV (where V is a W or Z boson) production.

After these cuts the main remaining backgrounds are Drell-Yan production and production of a lepton pair by low mass resonances. To suppress the latter a cut is applied on the invariant mass of the dilepton system, $m_{ll} > 20$ GeV, while for the former the region $76 < m_{ll} < 106$ GeV (i.e., within 15 GeV of the mass of the Z boson) is cut away for events with two electrons or muons (since Drell-Yan production only contributes to events with one electron and one muon in the case the Z boson first decays to a tau/anti-tau pair, one of which decays to an electron and the other to a muon, which is a process with a fairly low branching ratio). To further supress these processes, events are required to have a jet with $p_T > 30$ GeV and a b-tagged jet, as both signal processes contain at least one b jet, while these backgrounds do not (though they can have a b jet as an additional emission).

These cuts are sufficient to give a sample of events which are well modelled (consisting mainly of SM $t\bar{t}$ and tW production). However a further cut is applied for a different reason to the preceding ones; this analysis is designed to be combined with a search across all possible decay modes of $t\bar{t}$ and single top production, and in the search for semileptonic $t\bar{t}$+ DM a major background is SM dileptonic $t\bar{t}$, where one lepton is either not reconstructed or falls out of acceptance, giving a process which appears to have two b jets, one lepton and a large amount of p_T^{miss} from the two neutrinos, which if there are also two additional jets from additional emissions corresponds to the signature of semileptonic $t\bar{t}$ + DM. It is therefore useful for the semileptonic channel to have a control region enriched in SM dileptonic $t\bar{t}$ (and depleted in dileptonic $t\bar{t}$+ DM) to estimate the rate of this process as a function of p_T^{miss}, and this should be orthogonal to the regions used in the dileptonic channel presented here to avoid double counting of data. A further cut is therefore introduced on a "stransverse mass" variable (as introduced in reference [83]) defined as:

$$M_{T2}(l,\bar{l}) = \min_{\vec{p}_{T,\bar{\nu}}+\vec{p}_{T,\nu}=\vec{p}_T^{\text{miss}}} [\max\{M_T(m_l, m_{\bar{\nu}}, \vec{p}_{T,l}, \vec{p}_{T,\bar{\nu}}), M_T(m_{\bar{l}}, m_\nu, \vec{p}_{T,\bar{l}}, \vec{p}_{T,\nu})\}], \quad (17)$$

with

$$M_T(m_1, m_2, \vec{p}_{T,1}, \vec{p}_{T,2}) = \sqrt{m_1^2 + m_2^2 + 2(E_{T,1}E_{T,2} - \vec{p}_{T,1} \cdot \vec{p}_{T,2})}, \quad (18)$$

where m_i, $\vec{p}_{T,i}$, and $E_{T,1}$ respectively correspond to the mass, transverse momentum vector, and transverse energy of the particle i, and $\vec{p}_T^{\text{miss}}$ is the missing transverse momentum vector. This variable is an extension of the concept of transverse mass, the

mass of a multi-particle system computed only using the x and y components of the momenta (which is useful when one of the components is p_T^{miss}, where the z component of the momentum is not measured). By performing the minimisation over partitions of the p_T^{miss} into two neutrinos one gets the result that for SM $t\bar{t}$ (and tW) processes, where the $p_T^{\text{miss}}$is entirely due to the two neutrinos, the $M_{\text{T2}}^{\ell\ell}$ variable should have an upper bound of the transverse mass of the W bosons, which are themselves bounded by the W boson mass $m_W = 80.379$ [84]. In practise this bound is not exact as the p_T^{miss} is difficult to measure accurately due to issues such as pile-up subtraction and inhomogenous calorimeter response, however applying a cut of $M_{\text{T2}}^{\ell\ell} > 80\text{GeV}$ removes 95% of SM $t\bar{t}$ events, but only 60% of signal events (the W mass bound discussed for SM $t\bar{t}$ does not apply to the signal processes since the dark matter also contributes to the p_T^{miss}). The $M_{\text{T2}}^{\ell\ell} \leq 80\text{GeV}$ region can then be used as a control region by the semileptonic search channel, and for validation of this channel (as will be discussed in section 7.1).

After these cuts the predominant backgrounds are SM $t\bar{t}$, $tW$and Drell-Yan production. Other backgrounds are diboson production, typically processes with two leptons, though processes with three leptons also contribute as it is fairly likely that one of these will fall out of detector acceptance or otherwise not be reconstructed, and associated production of a $t\bar{t}$ pair and a vector or Higgs boson. One background which is particularly notable is the $t\bar{t}Z$ process, where the $t\bar{t}$ pair decays dileptonically and the Z boson decays to a neutrino anti-neutrino pair- this gives an almost identical signature to the $t\bar{t}$+ DM, as will be discussed later in 7.2.

6.7 Event Corrections

In addition to the per-object corrections described in section 6.5, there are a number of corrections which are applied to the entire event to account for mismodelling or detector effects.

6.7.1 Top Quark p_T mismodelling

While the SM $t\bar{t}$ background is mostly well described by MC simulation, a significant discrepancy has been observed in the p_T distribution of the top quarks [85], which leads to worse modelling of other variables, such as the p_T of the leptons. This effect is believed to be mainly due to missing higher order terms in the ME calculation, as it has been shown that analytical NNLO QCD + NLO EW predicitions [86] give a better fit to the data than the NLO QCD MC prediction. It was therefore decided to reweight the produced events by the ratio of this higher order prediction to the NLO prediction, as a function of the ME level top p_T. The effects of this can best be seen in the $t\bar{t}$ validation region (section 7.1) since this has much better statistics than the SRs; the top p_T and leading lepton p_T in this region before and after applying this reweighting are shown in figure 25.

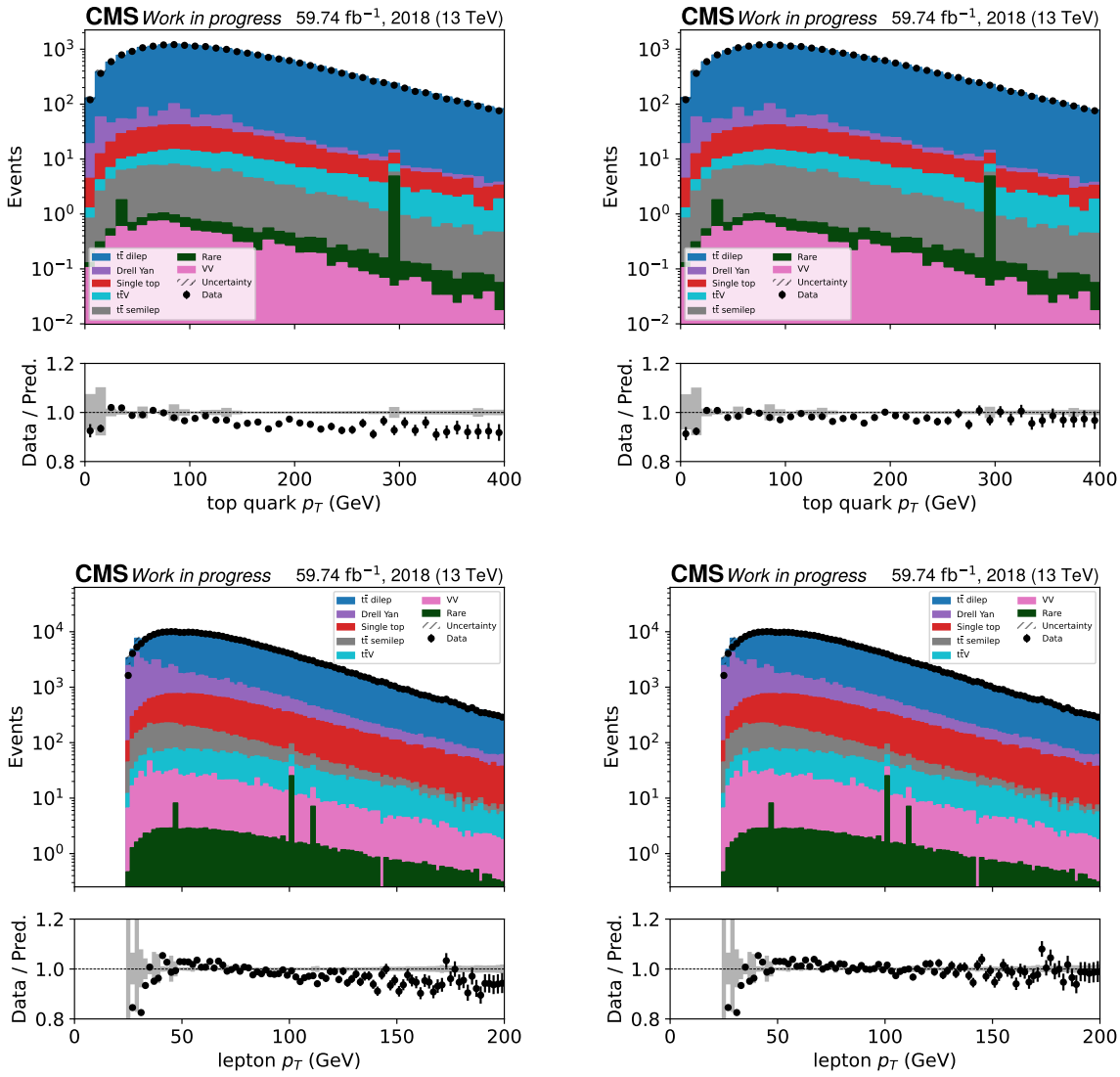


Figure 25: The comparison of top p_T and leading lepton p_T in the $t\bar{t}$ VR before (left) and after (right) applying top p_T reweighting. The shaded bands indicate the statistical uncertainty on the MC simulation.

6.7.2 Level 1 Trigger Prefiring

As the crystals of the ECAL are exposed to radiation, they slowly start to become opaque, which slows down the rate at which the signals can be read out from them. Unfortunately this effect was not properly accounted for in the read-out for the L1 trigger in the 2016 and 2017 data-taking periods, so high p_T electrons and photons could be assigned to the previous bunch crossing. This effect was highly pseudorapidity dependent, being much more prominent in the forwards regions where the particle flux and hence rate of degradation of the ECAL crystals was higher. While this trigger

information not being available for the correct bunch crossing in the L1 trigger is not very relevant for this analysis, which only uses trigger objects up to $|\eta| < 2.4$, avoiding the significantly affected region, the real issue is that the L1 trigger rules do not allow two consecutive bunches to be read out, as many detector components cannot be read fast enough. Hence any event with a large deposit in this forward ECAL region, either from an electron, photon, or jets containing these, could potentially "self-veto", by triggering on the bunch before. To account for this each MC event is scaled by a weight equal to the probability for each relevant object to prefire, based on their p_T and η:

$$w = 1 - P(\text{prefiring}) = \prod_{i=\text{photons,electrons,jets}} (1 - \epsilon_i^{pref}(\eta, p_T)) \tag{19}$$

6.7.3 Failure of HEM Modules

During 2018 data-taking, two sectors of the negative HCAL endcap, covering the region $-3 < \eta < -1.5$, $-1.57 < \phi < -0.87$ failed due to a power supply issue before the final run of era 2018 B, and hence provided no data in this run or any of eras C and D. This significantly affected the measurement of jets in this region, since there was no measurement of the neutral hadrons, but could also potentially lead to more fake electrons, since these rely on the measurement in the HCAL to reject jets which could otherwise fake electrons. Since this analysis is sensitive to events with mismeasured jets leading to mismeasured p_T^{miss}, and to processes with one genuine and one fake lepton, it was decided to veto any events with either electrons or jets in this region.

6.8 Kinematic Reconstruction

One can construct a number of useful physics observables given the momenta of the top quarks, as will be discussed in section 6.10. The standard way to obtain the four momenta of such heavy resonances is to sum the four-momenta of their decay products, however to do this one must assign the decay products in the event to their parent top quarks. For leptons this is trivial: top and anti-top quarks decay to anti-leptons and leptons respectively, and charge identification for both muons and electrons in CMS is excellent. For the bottom quarks this is more challenging; while charge identification algorithms do exist for b-jets, these are typically only about 60% efficient, since jets contain a large shower of both positively and negatively charged hadrons. For this analysis the following approach is taken: only jets which are b-tagged are considered as possible decay products of the top quarks (as will be described in section 6.9, only events with at least two b-tagged jets are considered for kinematic reconstruction), and of these the b-jet pair which has the most probable invariant mass when combined with the respective leptons, m_{lb}, are selected (where the expected probabilities of the m_{lb} distribution are taken from the generator level information of simulated $t\bar{t}$ events).

The most difficult objects to reconstruct are the neutrinos, since these are not observed. For a Standard Model $t\bar{t}$ event, the usual solution to this problem is to observe that the four momenta of the two neutrinos collectively have eight free parameters, and one can apply eight constraints to these: that both neutrinos, W bosons and top quarks must be on-shell, and that the transverse (x- and y-) components of the neutrino momenta must be equal to the measured p_T^{miss}. Using these constraints, one can construct a fourth order polynomial in one of the neutrino momenta, and solve it algebraically, as described in reference [88]. In the event that the p_T^{miss} or any of the particle momenta are mis-measured it is possible that no real algebraic solutions will exist, so this process is repeated multiple times with the inputs varied within their resolutions, and a weighted average taken of all successful solutions. However for the $t\bar{t}+\text{DM}$ signal this procedure does not work since the assumption that the p_T^{miss} is equal to the transverse momentum of the neutrinos no longer holds, as the dark matter also contributes. The solution is now undefined since there are not enough parameters to constrain the extra degrees of freedom of the DM, however one can get an approximate solution using the approach described in reference [87]. This first constructs ellipses of possible solutions in terms of the neutrino transverse momenta, using the on-shell conditions, then relates the position

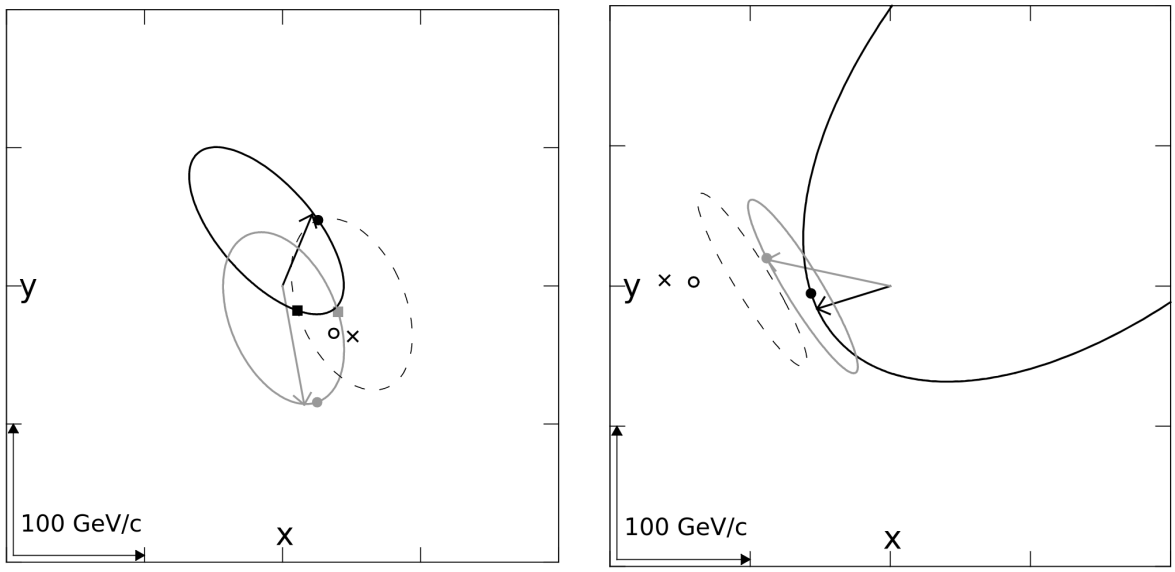


Figure 26: Determination of neutrino momenta in the kinematic reconstruction algorithm. The solid black ellipses show the possible neutrino momenta, p_ν, based on the on-shell constraints, while the grey ellipses show the possible anti-neutrino momenta, $p_{\bar{\nu}}$. The dashed ellipse represents p_T^{miss}-$p_{\bar{\nu}}$, which should intersect with p_ν if all of the p_T^{miss} is due to the neutrinos. The left hand plot shows an event where these ellipses do intersect, so the two points of intersection give possible solutions for the neutrino momenta. The right hand plot shows an event where the ellipses do not intersect, so the point of closest approach, represented as a dot, is taken as the solution. Source: [87].

of the ellipses to each other using the fact that the neutrino momenta sum to the p_T^{miss}. One can then take the intersections of the ellipses as solutions, and assuming the ellipses intersect these solutions will be identical to those from the previous method. However if the ellipses do not intersect (which corresponds to the case where no algebraic solution for the neutrino momenta exists) one can still take the point of closest approach of the ellipses as a solution. For a Standard Model event, the distance between the two ellipses is simply due to mismeasurement of the p_T^{miss} and other input momenta, however for $t\bar{t}$ + DM events this can be interpreted as a minimum amount of p_T^{miss} which must be attributed to the dark matter, this variable was therefore christened "p_T^{dark}". This is expected to be larger for $t\bar{t}$ + DM events than $t\bar{t}$ since it comes from an additional particle rather than just mismeasurement, and is used in the NN training, as described in section 6.10.

6.9 Signal Regions

As discussed previously, the t + DM and $t\bar{t}$ + DM signals have very similar signatures, differing by only one b-jet, and in practise these signatures overlap at reconstruction level, as often one of the b-jets in the $t\bar{t}$ + DM process will not be tagged, while the t + DM process can sometimes include additional b-jets due to initial state radiation. Nevertheless, it is useful to target these signals separately, since for $t\bar{t}$ + DM one can perform the kinematic reconstruction discussed in the previous section to gain access to further variables, while for t + DM it is useful to use variables designed to differentiate single top processes from the dominant $t\bar{t}$ background.

The signal region described in section 6.6 was therefore split in two, with one region targeting each of the two signals. Various methods were tried for this splitting, including using a neural network to differentiate the signals, however while these could achieve slightly better discrimination of the two signals from one another than a simple cut on the number of b-jets, optimising this classification did not notably improve the final signal/background discrimination. It was therefore decided to use the simple splitting, with events with exactly 1 b-tagged jet entering in the t + DM region, and events with ≥ 2 b-tagged jets in the $t\bar{t}$ + DM region.

Furthermore, it was found that the Drell-Yan process was not very well modelled in the signal regions due to the $M_{\text{T2}}^{\ell\ell}$ and number of b-tagged jet (n_{b-tag}) requirements. To mitigate this, it was decided to split each signal region into same-flavour (SF) and opposite-flavour (OF) lepton regions, i.e. a region containing ee and $\mu\mu$ events, where Drell-Yan is a more significant background, and a region containing $e\mu$ events, where Drell-Yan is significantly suppressed, only contributing via the $Z \to \tau\tau \to e\mu$ channel, which has a very low branching ratio. The SF t+DM region was found to be particularly enriched in the Drell-Yan process, and therefore a data-driven correction was derived in a control region, as described in section 7.3. For the SF $t\bar{t}$ + DM region it was not possible to define a similarly enriched control region since a significantly lower fraction of Drell-Yan events are produced in association with two b jets, however for the same

reason the fraction of Drell-Yan events in this region was much lower than in the SF t + DM region, so no data-driven correction was deemed necessary.

6.10 Sensitive Variables

It has previously been shown [57] that one can achieve a significant increase in sensitivity to $t\bar{t}$ + DM by using multivariate analysis techniques rather than just fitting on p_T^{miss}, which as has been discussed in section 6.1, is not very effective at discriminating against the dominant $t\bar{t}$ background. A set of high level variables were therefore identified, which help discriminating both the $t\bar{t}$+DM and t+DM signals from the SM backgrounds. The first two of these have already been discussed: first p_T^{miss}, which, although it only provides moderate discrimination against $t\bar{t}$, is very useful for suppressing other backgrounds such as Drell-Yan production, and secondly $M_{\text{T2}}^{\ell\ell}$, which as discussed in section 6.6, is designed to suppress the SM $t\bar{t}$ and tW processes. The distributions of these variables are shown in figure 27.

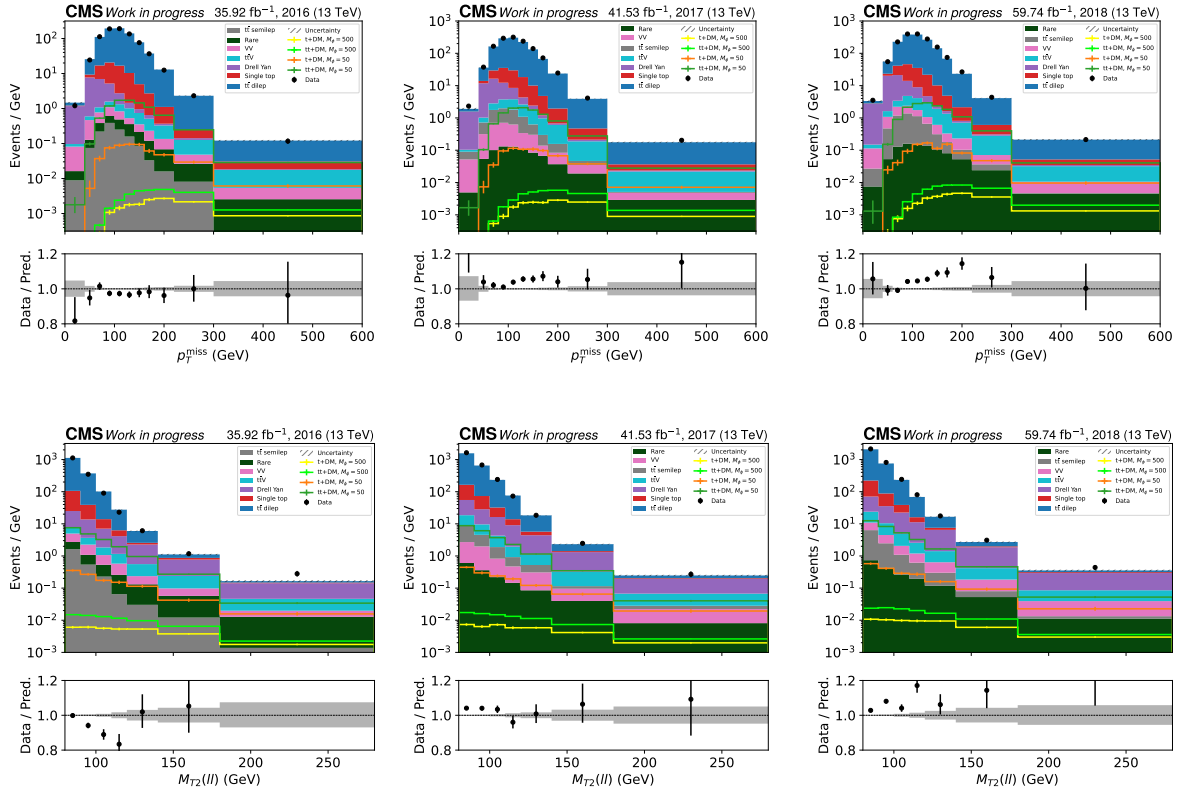


Figure 27: p_T^{miss}(top) and $M_{T2}(ll)$ (bottom) distributions in the signal region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

For events in the $t\bar{t}$+DM region, it is possible to approximately reconstruct the original $t\bar{t}$ system, as described in section 6.8. This gives access to a number of further variables, as

shown in figure 28. The first of these is the "p_T^{dark}", the amount of p_T^{miss} which cannot be assigned to neutrinos in the kinematic reconstruction, so is either due to mismeasurement or the presence of additional invisible particles such as dark matter. The signal process, which contains high p_T invisible dark matter, is expected to have higher p_T^{dark} than the $t\bar{t}$ and other background processes, where this is just due to mismeasurement.

The next of these is c_{hel}[89], which is the cosine of the angle between the lepton and anti-lepton, when boosted back in to the rest frame of the their parent (anti-)top quarks. This is sensitive to the spin correlations of the top quark pair, since the tops decay to a W boson and b quark before forming a hadron, and hence the spin is transferred to the W boson and the lepton is thus emitted in the direction of the spin of the parent top quark. This variable is most sensitive to the process with a pseudoscalar mediator, since this has a γ_5 term in the matrix element, which flips the spin of the top quark from which it is radiated. However the emission of the scalar mediator also has an impact on this variable, since the higher center of mass energy required to produce the additional final state particles leads to different contributions for the different initial states ($q\bar{q}$ relative to gluon fusion), which changes the spin correlations of the $t\bar{t}$ pair.

The final variable based on the $t\bar{t}$ system is $|\Delta\phi(t,\bar{t})|$ the azimuthal angle between the top quarks. This is sensitive to the signal as for the dominant $t\bar{t}$ background the top quarks should be back-to-back in the azimuthal plane by momentum conservation (aside from slight boosts of the system due to the emission of additional jets, but these are generally lower in energy than the top quarks), while for the $t\bar{t}+\text{DM}$ process the $t\bar{t}$ system is significantly boosted by the emission of the dark matter, and hence the azimuthal angle between the top quarks is smaller.

Finally there are two variables designed to provide sensitivity to the $t+\text{DM}$ process, shown in figure 29. The first of these is $|\Delta\phi(l,\bar{l})|$, the angle between the two leptons in the azimuthal plane. This is sensitive to the same spin correlation effects as c_{hel}, though less so since the leptons are not boosted into the rest frames of their parent top quarks. Nevertheless it can be helpful for identifying the spin correlations in $t+\text{DM}$ events, or $t\bar{t}+\text{DM}$ events where one b-jet is not tagged. The other variable is $|\Delta\phi(p_T^{\text{miss}},l\bar{l}b)|$, the azimuthal angle between the p_T^{miss} and the sum of the lepton, anti-lepton and b-jet momenta. This variable is expected to be close to π radians for both SM tW and $t+\text{DM}$, since the visible decay products are expected to recoil against the neutrinos for SM tW, and the neutrinos and DM for $t+\text{DM}$ (with small deviations due to additional radiation), whilst for $t\bar{t}$, which is the dominant background even in the $t+\text{DM}$ SR, there is an additional (untagged) b-jet, which means the $l\bar{l}b$ system is not back-to-back with the p_T^{miss}.

Although these distributions were designed to target the $t+\text{DM}$ signal, they were also found to have some sensitivity in the $t\bar{t}+\text{DM}$ region, likely because, as discussed in section 6.9, there is a notable amount of $t+\text{DM}$ signal in the $t\bar{t}+\text{DM}$ region. These were therefore used as inputs to the NN in both regions. To compute $|\Delta\phi(p_T^{\text{miss}},l\bar{l}b)|$ in the $t\bar{t}+\text{DM}$ region, the highest p_T b-tagged jet was taken as the b-jet.

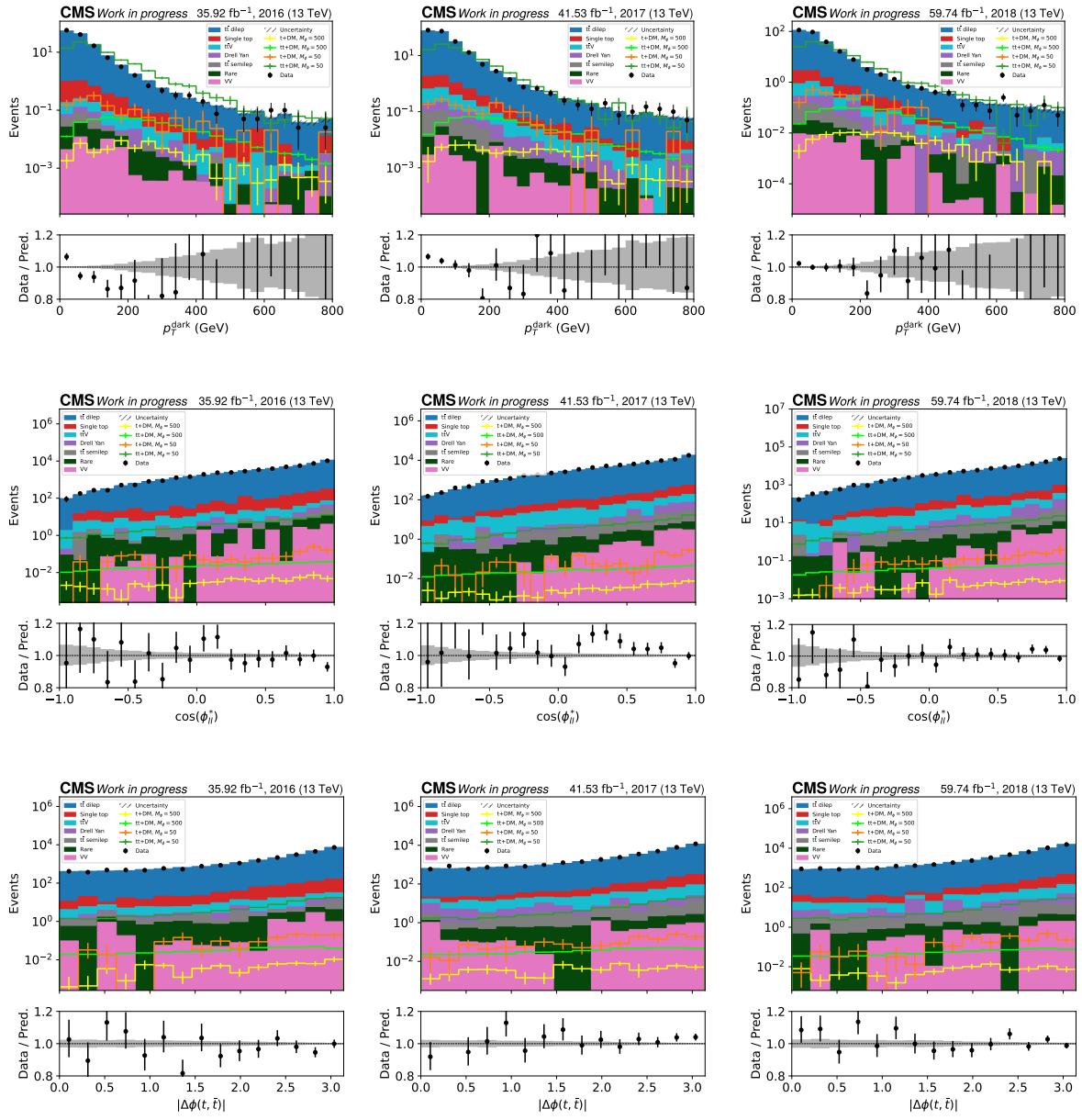


Figure 28: p_T^{dark} (top), c_{hel} (middle) and $|\Delta\phi(t,\bar{t})|$ (bottom) distributions in the signal region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

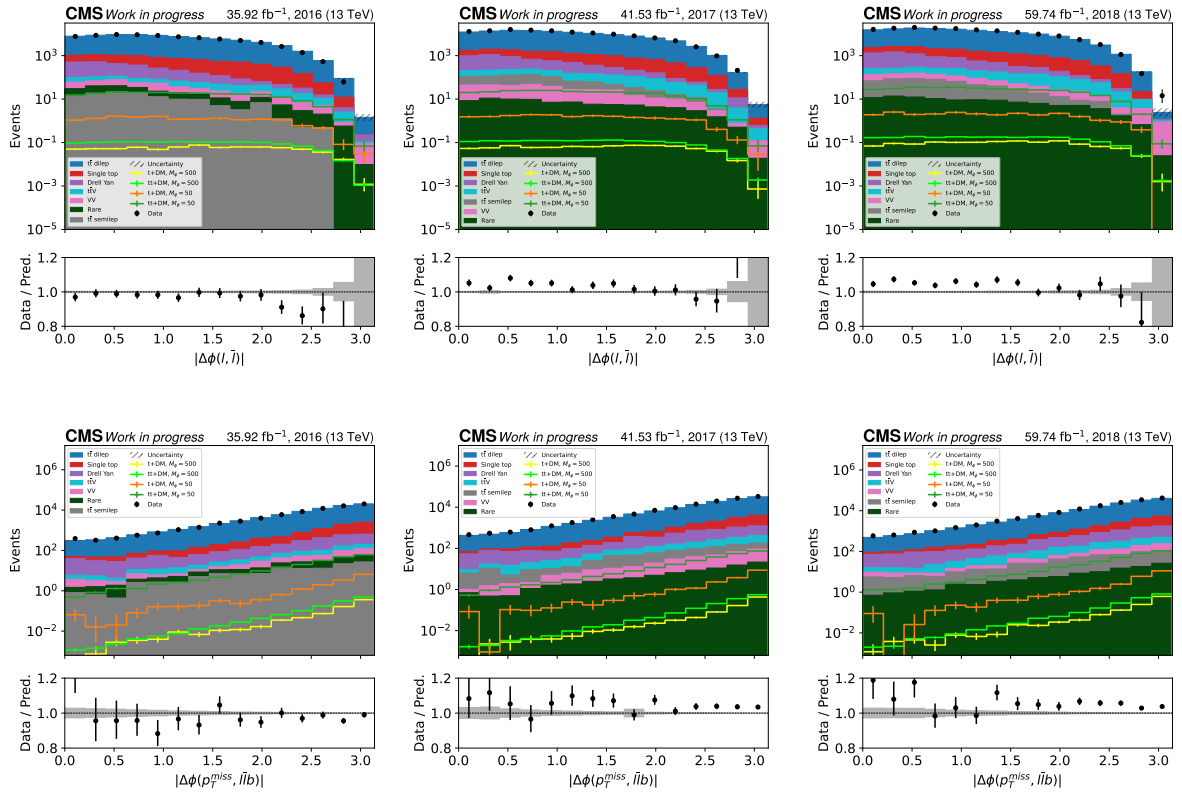


Figure 29: $|\Delta\phi(l,\bar{l})|$ (top) and $|\Delta\phi(p_T^{\text{miss}}, l\bar{l}b)|$ (bottom) distributions in the signal region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

6.11 Neural Network Set-up

A neural network was implemented in each of the signal regions to extract the signals from the background using the sensitive variables discussed in the previous section. This was implemented using the Keras [90] Python package, and the selected architecture consisted of three hidden layers, with 40, 40 and 30 nodes respectively, using "leaky" Rectified Linear Unit (ReLU) activation functions, then two output nodes corresponding to the "score" for an event to be signal or background respectively, using a "softmax" activation function. This has the property that both scores lie between 0 and 1 and sum to 1, allowing them to be interpreted as probabilities. Since there is some overlap in signals between the two regions, both types of signal were included in the training in both regions. A "Sparse Categorical Cross-entropy" function was used to evaluate the loss, since this is optimised for this type of classification problem, and the Adam optimiser [91], which uses a variable learning rate and has been shown to achieve rapid, stable training even on noisy data, was used for optimisation. 50 epochs were used for the training as the loss function was not found to decrease significantly with further epochs after this point. A batch size of 100 events was used - varying this was not found to have a notable impact on the learning rate or final performance, so this value was

selected as it gave a notable speed up in training time while not exceeding the memory constraints.

A common problem for neural networks is that they can "overtrain", learning features which do not correspond to the structure of the problem in question, but instead arise due to statistical fluctuations in the training set, which can prevent the neural network from generalising well to unseen data. To mitigate this, the MC events were divided into three sets, with 40% used in training the NN, 20% for validation and 40% for testing. The purpose of separate validation and test datasets is that the validation set can be used while optimising the neural network to ensure the neural network does not overtrain, then the test dataset can be used once the optimisation is complete to check that the hyperparamters and architecture selected do indeed not cause overtraining, and that the agreement between the training and validation sets is not due to an accident of hyperparameter tuning. During the training the loss function was monitored for the validation dataset to ensure this stayed close to the training loss, since if the validation loss no longer decreased as rapidly as the training loss, or started to increase, this would be a sign of over-training. As can be seen in figure 30, no signs of overtraining were observed (likely due to the relatively low dimensionality of the inputs), and this was confirmed by comparing the final output distributions of the test and training datasets (figure 31).

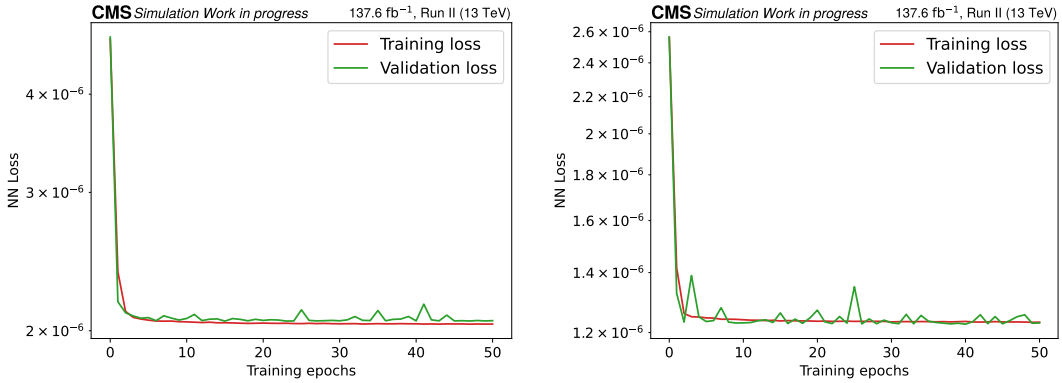


Figure 30: Cross-entropy loss of the NN targeting the scalar mediator in the $t\bar{t}$ + DM region (left) and the t + DM region (right).

Separate neural networks were initially trained for all of the different signal mass point hypotheses considered. However, since the signatures are very similar for all mass points, it was observed one could achieve very similar discrimination on the individual mass points using a neural network trained on a combination of all mass points to one trained on that specific hypothesis (figure 32). It was therefore decided to train a single neural network in each region for all of the mass points combined to have a consistent background distribution for all hypotheses. However, separate neural networks were still trained for the scalar and pseudoscalar mediators, since these have notably different distributions, especially for c_{hel}, as discussed in section 6.10.

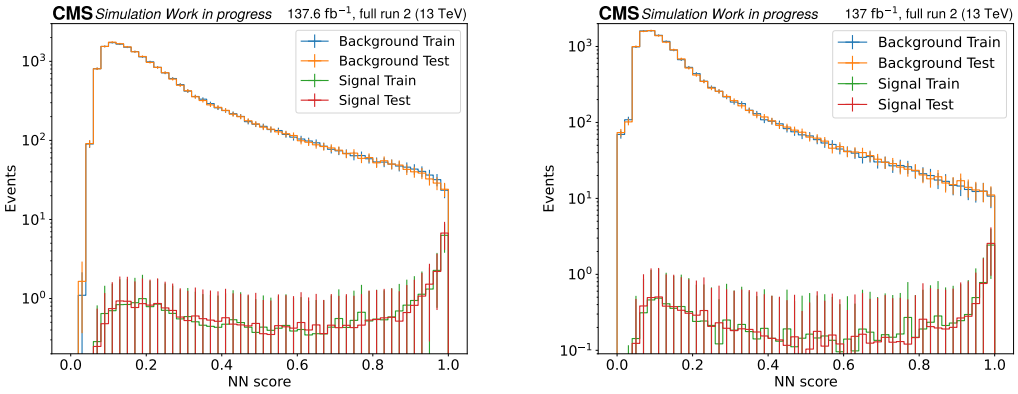


Figure 31: A comparison of the NN output between the training and test datasets for the 100 GeV scalar mediator signal and the combined SM backgrounds for the NNs in the $t\bar{t}$ + DM region (left) and the t + DM region (right).

In order to use the NN output for further signal extraction, it is necessary to bin it into a histogram. This requires a trade-off, as particularly for the events with a NN score close to 1, selecting a finer binning can increase sensitivity, but at the risk that the low MC statistics used in training such a small region may lead to fluctuations which give false positive peaks. It was therefore decided to use a binning in each year which gives 40, 20 and 10 expected events in the last three bins, respectively, in the $t\bar{t}$ + DM and t + DM regions, which correspond to 20, 10 and 5 events after splitting by lepton flavour, which is believed to be enough to avoid issues with statistical fluctuations, while giving good sensitivity to the model. The remainder of the distribution has very little sensitivity since the signal is sharply peaked at high NN values, but is useful for constraining uncertainties on the backgrounds, and so is divided into nine equal bins in NN score. The NN output distributions with this binning are shown in figures 33 and 34.

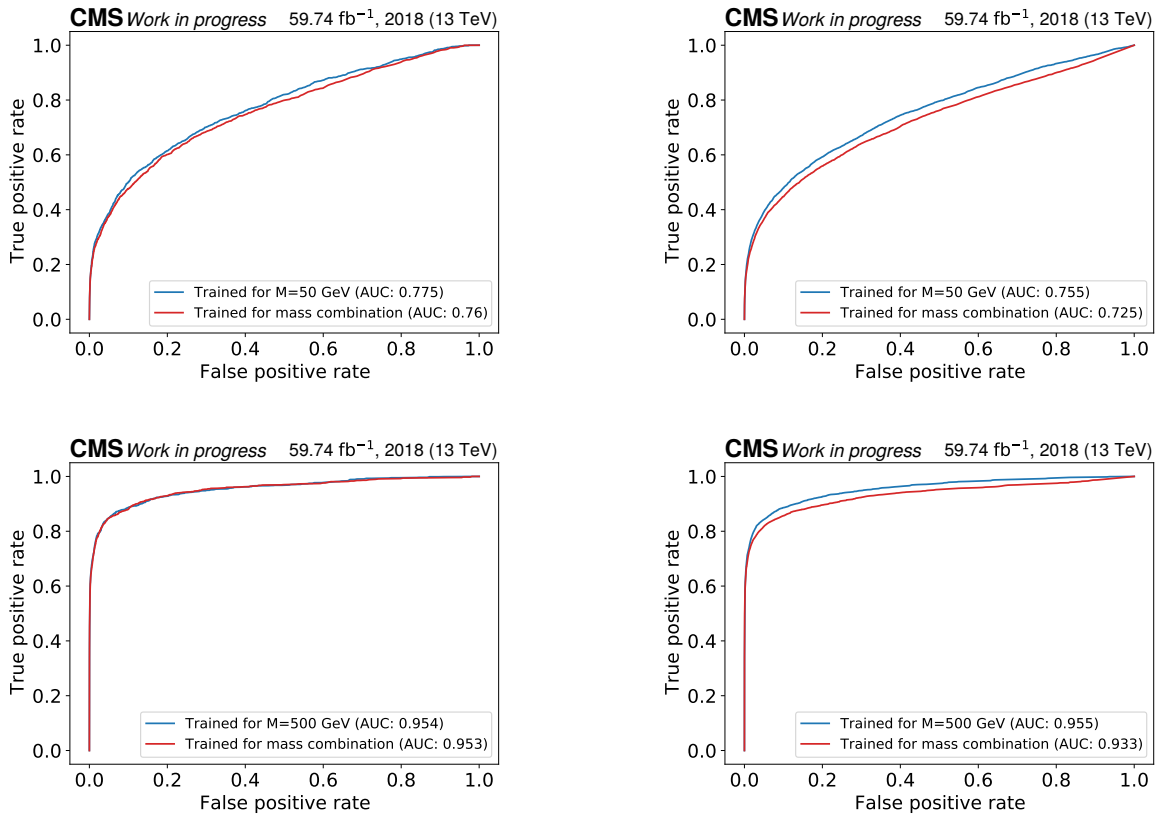


Figure 32: A comparison of the receiver operating characteristic (ROC) of Neural Networks trained on the mass point under consideration (blue) versus those trained on a combination of all mass points (red) for the scalar mediator. The area under the curve (AUC) is also given to aid comparison. The top row is for $m_\phi = 50$GeV, the bottom $m_\phi = 500$GeV; the left column is the $t + \text{DM}$ region, right $t\bar{t} + \text{DM}$. The similarity of the curve for these extremal values of the mass hypothesis indicates good performance of the NNs trained on the combination of signals.

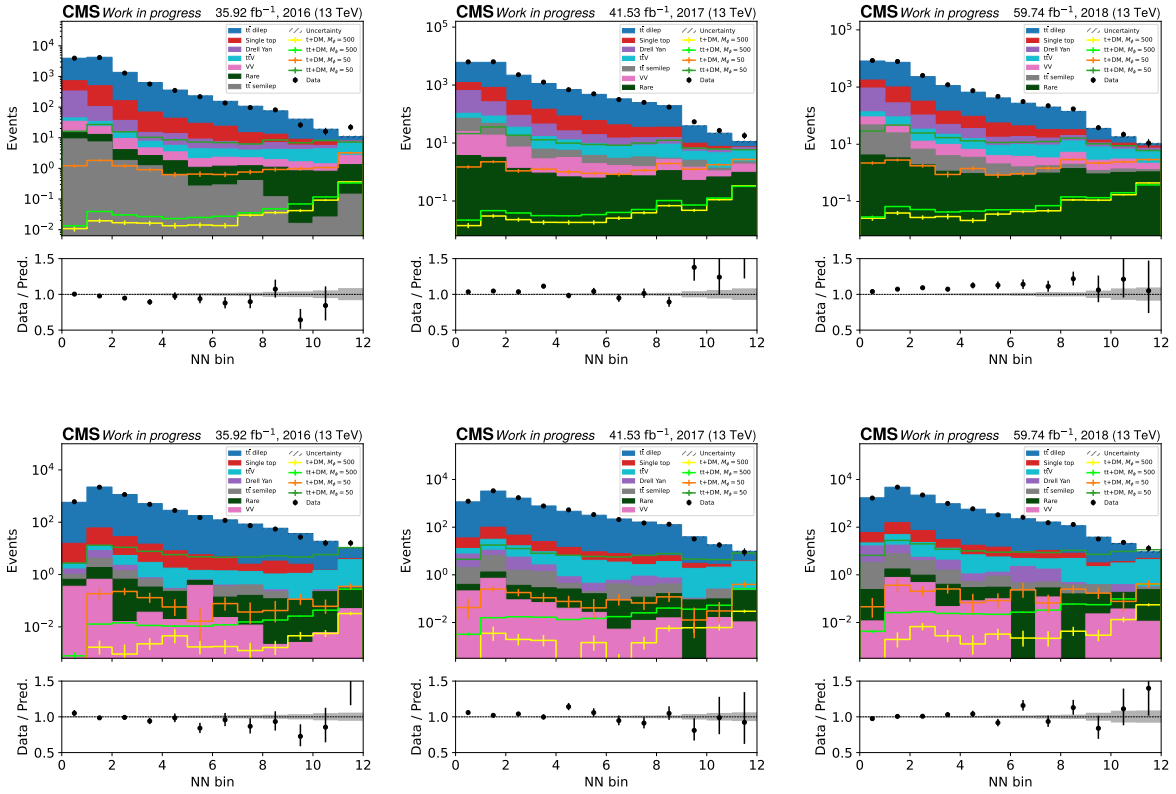


Figure 33: Scalar mediator NN output distributions in the t + DM (top) and $t\bar{t}$ + DM (bottom) signal region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

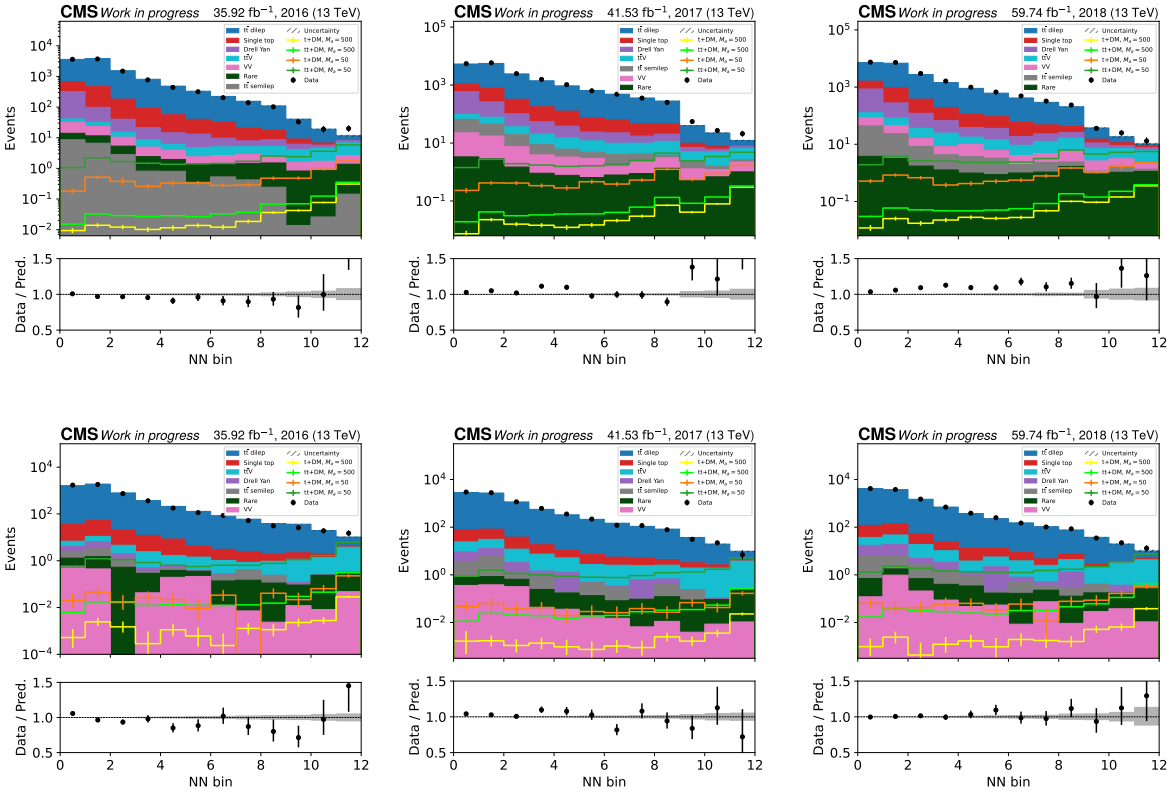


Figure 34: Pseudoscalar mediator NN output distributions in the t + DM (top) and $t\bar{t}$+DM (bottom) signal region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

7 Control and Validation Regions

In addition to the signal regions defined in 6.9, a number of further regions were defined to ensure good modelling of the background processes. These can be split into two types: validation regions, which are simply used to check the background is well modelled in MC simulation, and control regions, which are used to derive a correction to the MC background based on the data in this region, which can be useful if the process is not well-modelled by the simulation, or if it has large theoretical uncertainties, which can be constrained by the data. These regions were particularly useful while setting up the analysis, as the signal region was "blinded", but one could still look at the full data in these regions as they have a significantly lower signal contribution relative to the backgrounds.

7.1 $t\bar{t}$ Validation Region

The first region is used to validate the dominant $t\bar{t}$ and single top backgrounds. This is done using a region with an identical selection to the signal region described in section 6.6, except requiring $M_{\mathrm{T2}}^{\ell\ell} \leq 80$ GeV. Since SM $t\bar{t}$ and single top are expected to satisfy this bound (assuming no quantities are mismeasured), this region is significantly enriched in these processes, giving a signal-to-background ratio 6 to 10 times lower than the signal region, depending on the signal point considered. This allowed this region to be used for monitoring the kinematic distributions of these backgrounds, which were found to be well modelled, meaning no data-driven corrections were required. This region was also used to ensure the variables used as inputs for the NN were well modelled; these are shown in figures 35 to 37.

7.2 $t\bar{t}Z$ Control Region

As discussed in section 6.3, $t\bar{t}Z$ ($Z \to \nu\bar{\nu}$) is a challenging background for this analysis since it has a signature of two top quarks + p_T^{miss}, almost identical to the $t\bar{t}$+DM signal. It is therefore important to ensure good modelling of this background. Moreover, previous measurements have found the cross section for this process to be in tension with the SM value [92]. A dedicated control region was therefore set up to constrain the cross section of this process. This was done by targeting the $Z \to ll$ decay, since this is the easiest of the Z boson final states to identify. Ideally one could also require dileptonic $t\bar{t}$ decays, then one could study the same analysis variables as the signal region by adding the p_T of the leptons coming from the Z boson to the p_T^{miss}. However this channel has very low statistics due to the small $t\bar{t}Z$ cross section and $t\bar{t}$ dileptonic branching ratio, so instead the semileptonic $t\bar{t}$ decay was targeted, since it has the highest branching ratio.

The selection to target this process therefore required three leptons with $p_T > 20$ GeV, two of which were required to be of the same flavour and opposite sign (i.e., an e^+e^- or $\mu^+\mu^-$ pair from the Z boson decay). This pair was required to have a reasonably tight

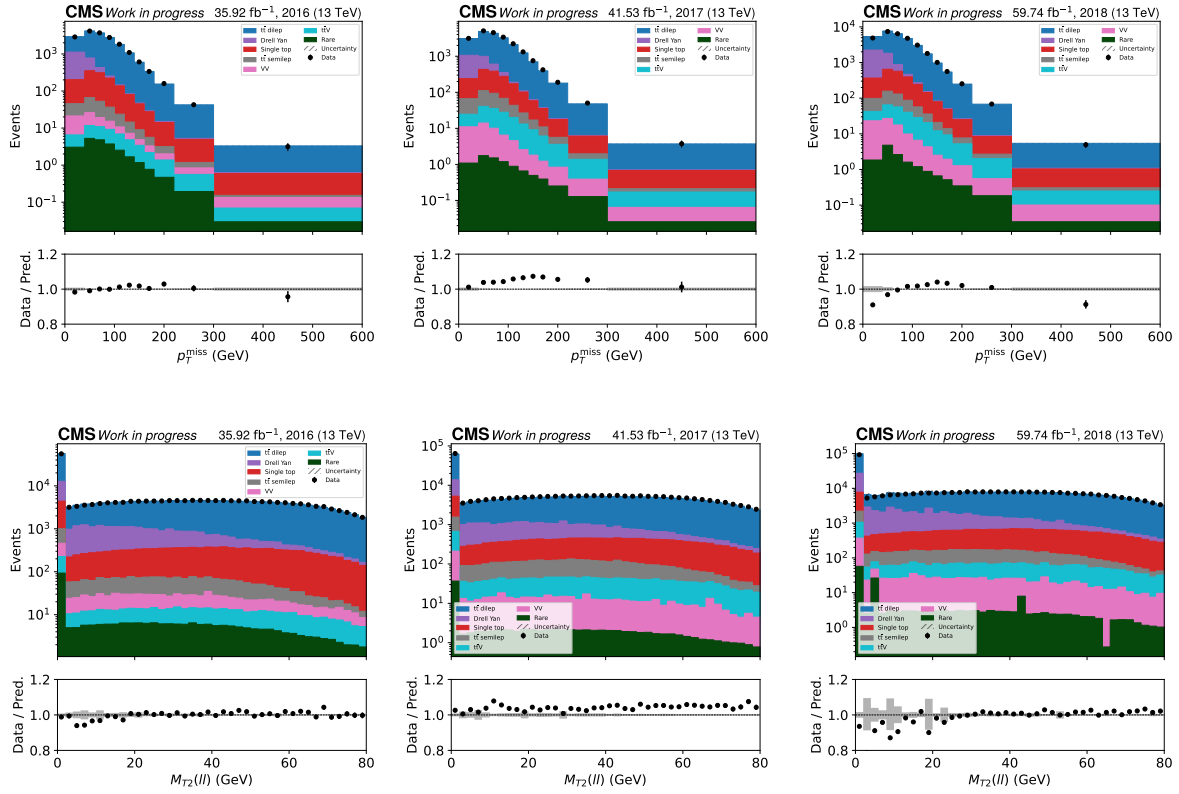


Figure 35: p_T^{miss}(top) and $M_{T2}(ll)$ (bottom) distributions in the $t\bar{t}$ control region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

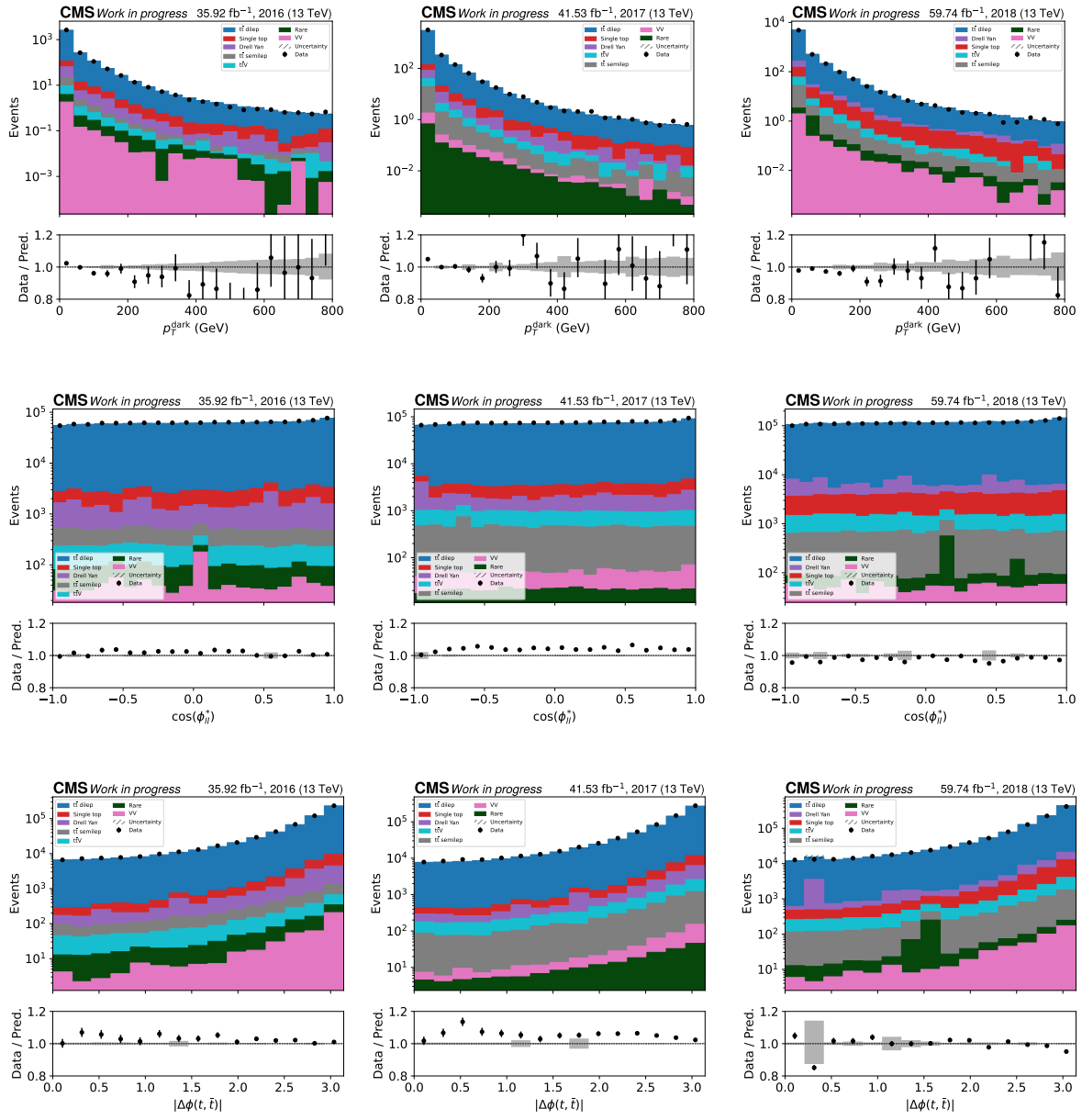


Figure 36: p_T^{dark} (top), c_{hel} (middle) and $|\Delta\phi(t,\bar{t})|$ (bottom) distributions in the $t\bar{t}$ control region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

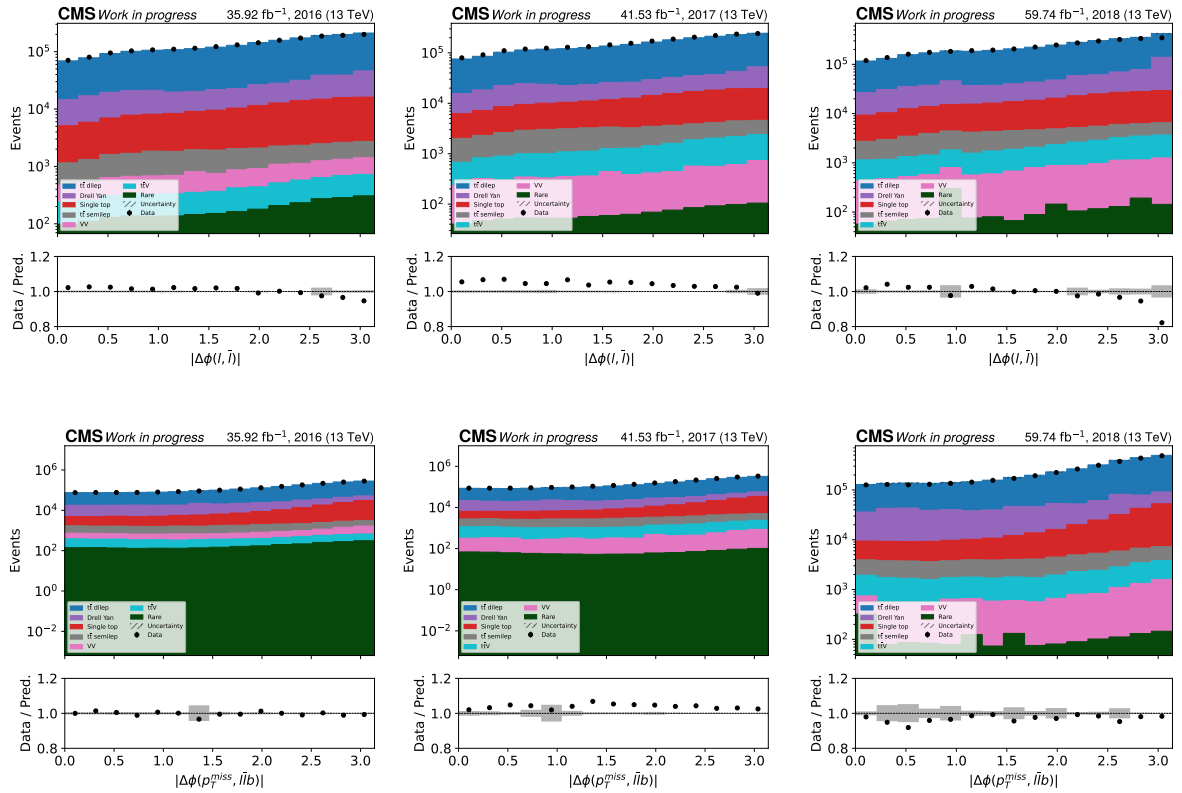


Figure 37: $|\Delta\phi(l,\bar{l})|$ (top) and $|\Delta\phi(p_T^{\text{miss}}, l\bar{l}b)|$ (bottom) distributions in the $t\bar{t}$ control region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

cut of $|m_{ll} - m_Z| < 10$ GeV to increase the probability they came from a Z boson decay rather than another combination of leptons; if there were two possible same flavour, opposite sign pairs, the pair closer to the Z boson mass was taken as the pair coming from the Z boson. Since a major background for this region was Drell-Yan, where there is a genuine same flavour, opposite sign lepton pair from the Z decay, and the other lepton is a fake with generally low p_T, a cut of $p_T > 35$ GeV was applied on the lepton not considered to come from the Z, which helped to supress this and other background with only two genuine leptons.

The semileptonic decay of the $t\bar{t}$ pair is expected to produce four jets, two of which are b jets, however since these jets can fall outside of detector acceptance or the b jets not be tagged, the events were only required to have three jets (with $p_T > 30$ GeV for the leading jet and $p_T > 20$ GeV for the others), with one of these b-tagged.

This selection gives a region with good statistics moderately enriched in the $t\bar{t}Z$ process, which represents about 45% of the expected yield. The main backgrounds are diboson production, mainly the WZ ($W \to l\nu$, $Z \to ll$) process, which has the same leptons, though generally has fewer jets than $t\bar{t}Z$, and processes with two genuine leptons and a "fake" lepton, mainly $t\bar{t}$, Drell-Yan and tW production. Other backgrounds are other $t\bar{t}V$ processes, mainly $t\bar{t}W$, where both tops and the W decay leptonically, which has the right number of leptons and b jets, though the leptons will tend not to satisfy the requirements to have come from a Z boson, and tZq, which can give two leptons from a Z decay plus one from a top, but has fewer jets and b jets, as well as a small inclusive cross-section. The p_T distributions for the leptons and the first two jets are shown in figures 38 and 39, respectively.

To extract the best possible $t\bar{t}Z$ cross section, one should fit on a variable which helps discriminate $t\bar{t}Z$ from the relatively large backgrounds in this region. This was done using a two dimensional histogram of jet and b jet multiplicities, since $t\bar{t}Z$ tends to have 2 b-tagged jets and more jets than the backgrounds. These distributions are shown in figure 40.

7.3 DY Control Region

Unlike most of the backgrounds used in this analysis, the Drell-Yan process was found to be poorly modelled by the MC simulation, which is not unexpected since, as discussed in section 6.3, it is difficult to model p_T^{miss} for this process. It was therefore decided to create a control region to scale the predictions for this background on a bin-by-bin basis in the SF t + DM region, which, as discussed in section 6.9, is the region in which this background is most present.

In order for this CR to provide reasonable predictions of the yield for the Drell-Yan background on a per-bin basis, it is important that the control region is as close to the signal region as possible, especially in the variables which are known to be difficult to

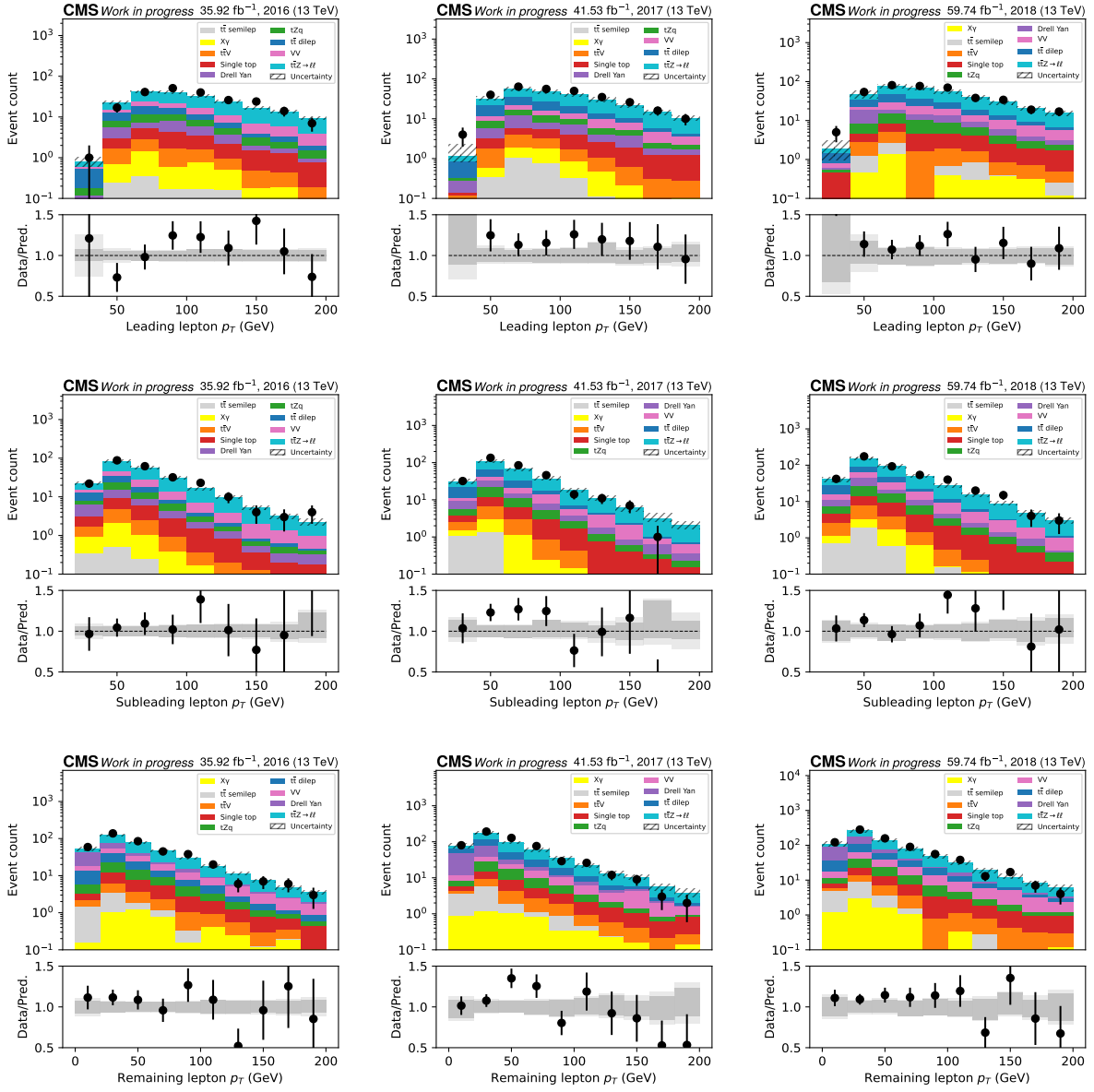


Figure 38: p_T of the leading (top) and subleading (middle) lepton from the pair identified as coming from the Z boson, and of the lepton identified as coming from the top decay (bottom) in the $t\bar{t}Z$ control region for the 2016 (left), 2017 (center) and 2018 (right) data-taking periods. Errors are systematic and statistical added in quadrature.

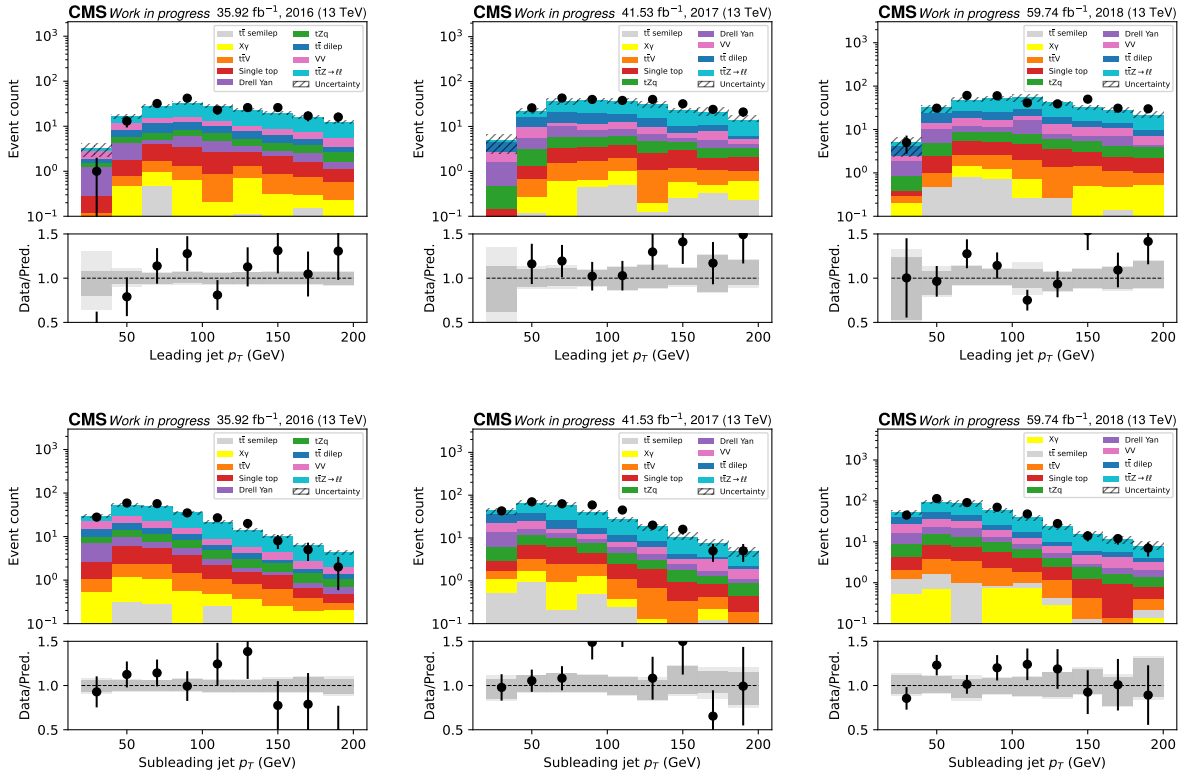


Figure 39: p_T of the leading (top) and subleading (bottom) jet in the $t\bar{t}Z$ control region for the 2016 (left), 2017 (center) and 2018 (right) data-taking periods. Errors are systematic and statistical added in quadrature.

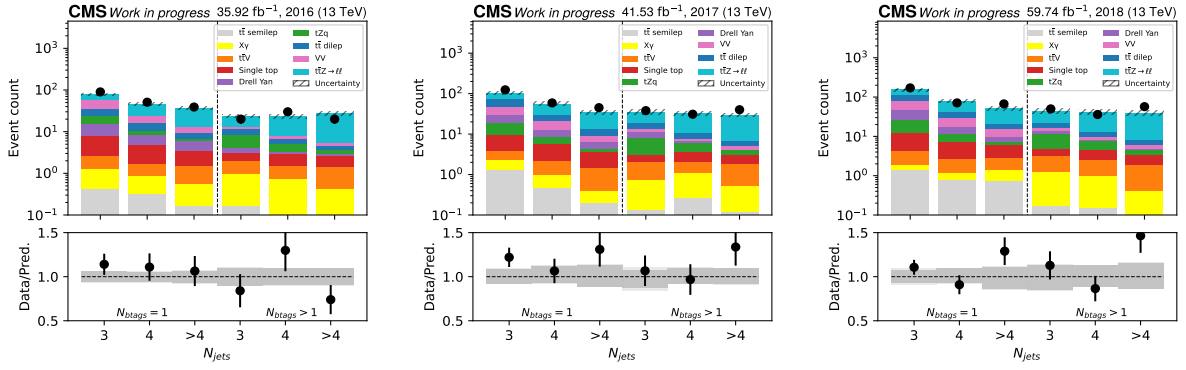


Figure 40: Two dimensional n_{jet}/n_{b-tag} distributions in the $t\bar{t}Z$ control region in 2016 (left), 2017 (center) and 2018 (right). Errors are systematic and statistical added in quadrature.

model, $M_{\mathrm{T2}}^{\ell\ell}$ (since this is derived from p_T^{miss}) and the b jet multiplicity (which is hard to model since these come from additional emissions in the initial state). The selection used is therefore identical to that for the $t + \mathrm{DM}$ signal region, i.e. all of the cuts in section 6.6 are applied, except the Z window cut, which is inverted, so events are required to have a same flavour lepton pair with $76 < m_{ll} < 106$ GeV. Additionally all events are required to have exactly 1 b jet, to keep the selection as close as possible to the $t + \mathrm{DM}$ SR. This selection is naturally enriched in Drell-Yan, with about 80% of events coming from this process. Due to its proximity to the signal region, it naturally also contains some signal events, but due to the large number of DY events the signal/background ratio is roughly 6 times lower than in the $t + \mathrm{DM}$ SR.

The NN input and output distributions in this region are shown in figures 41 and 42, respectively. These distributions are generally reasonably well modelled by the MC simulation, however there are some disagreements in regions which are particularly sensitive to the $t\bar{t} + \mathrm{DM}/t + \mathrm{DM}$ signal, such as the tail of the p_T^{miss} distribution in 2016 and 2017, where there is a deficit of MC with respected to the observed data, which leads to a corresponding deficit in the final bins of the NN distribution, which, if left uncorrected, could lead to a corresponding deficit of Drell-Yan in the $t + \mathrm{DM}$ signal region, potentially leading to a spurious excess of signal-like events.

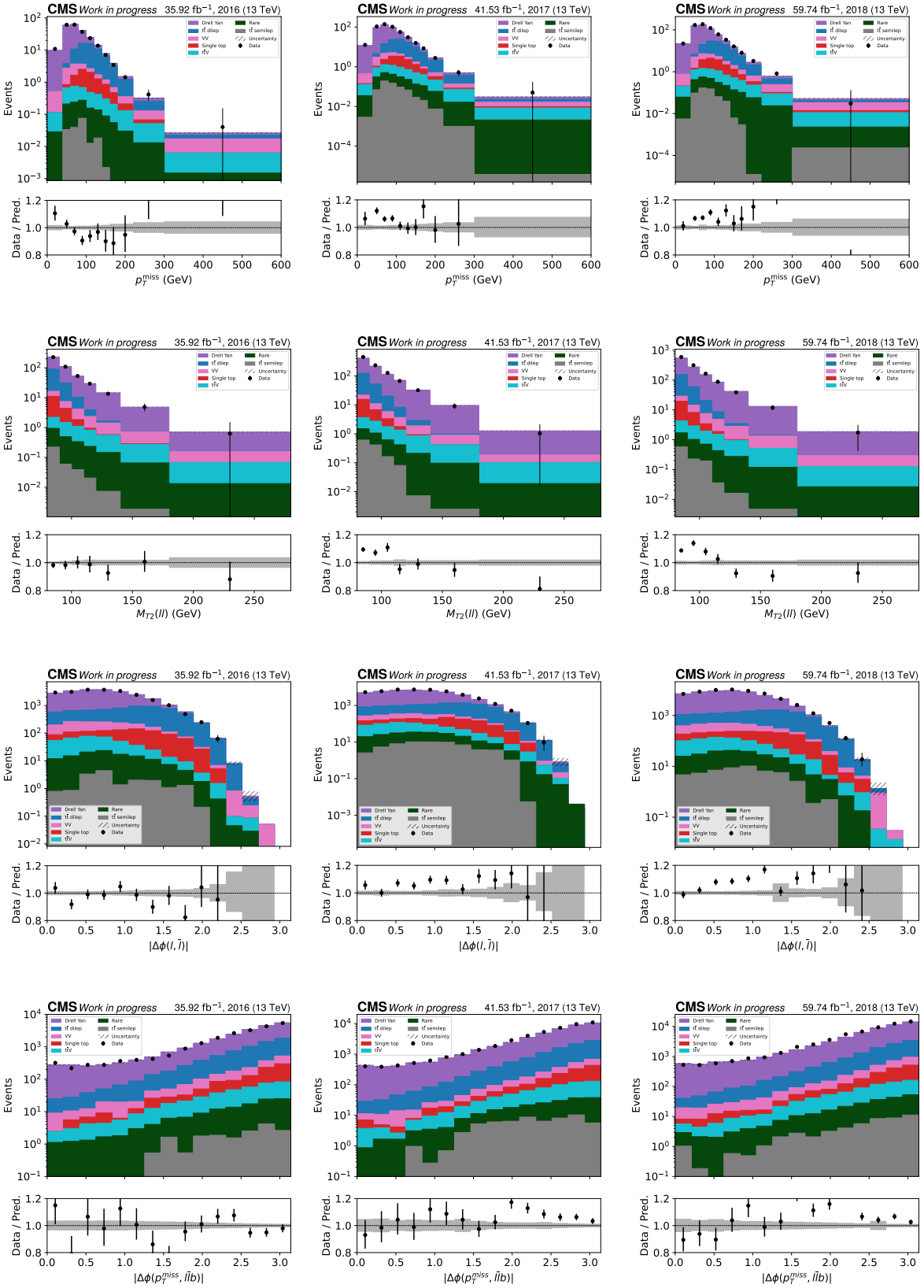


Figure 41: NN input distributions in the Drell-Yan control region for 2016 (left), 2017 (middle) and 2018 (right). Top to bottom: p_T^{miss}, $M_{\text{T2}}^{\ell\ell}$, $|\Delta\phi(l,\bar{l})|$ and $|\Delta\phi(p_T^{\text{miss}}, l\bar{l}b)|$. The shaded bands indicate the statistical uncertainty on the MC simulation.

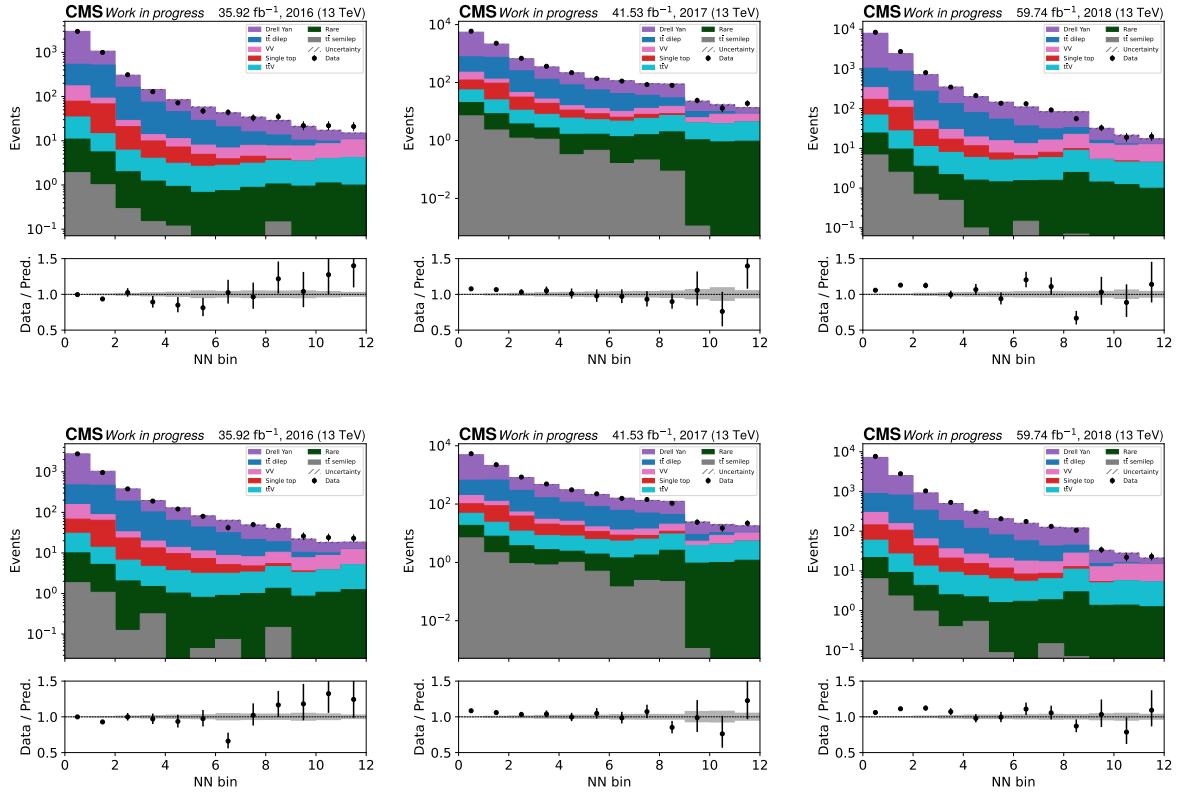


Figure 42: Scalar (top) and pseudoscalar (bottom) mediator NN output distributions in the Drell-Yan region for 2016 (left), 2017 (middle) and 2018 (right). The shaded bands indicate the statistical uncertainty on the MC simulation.

8 Systematic Uncertainty Estimation

There are number of systematic uncertainties associated with the modelling of the SM backgrounds for this search, which are estimated in various ways. In this chapter the various sources of systematic uncertainty are discussed, as well as how these uncertainties are estimated and treated.

8.1 Systematic Uncertainty Sources

The following uncertainty sources were considered in the final fit. All uncertainties are treated as correlated across all years and data taking periods, unless otherwise noted.

8.1.1 Theoretical Uncertainties

The first set of uncertainties are associated with the generation of MC events. These are factorised into separate uncertainties for the different steps of the generation. These uncertainties are not applied to the Drell Yan process, since this is estimated in a dedicated CR, while for the $t\bar{t}Z$ process, where the overall rate is estimated in a CR, the variations are normalised to the same value as the nominal template, so only the shape effect of these uncertainties is considered.

- *PDF uncertainties.* There are multiple sources of uncertainty on the PDF sets used in this analysis, coming from the uncertainties on the measurements used in deriving the sets, the functional form used to parameterise them and the theoretical uncertainties associated with evolving the sets between different scales. For NNPDF 3.0, the effect of these uncertainties is estimated by producing 100 separate MC replicas of the PDF set for variations of these uncertainties, then the 16^{th} and 84^{th} replicas are taken as the 1σ variation, as recommended in reference [93]. For NNPDF 3.1, 100 pseudo-hessian uncertainties are provided (which are produced from 1000 MC replicas using the mc2hessian method [94]), and these are then summed in quadrature. This set is also supplied with a separate uncertainty based on varying α_s within 0.002 of its central value of 0.118, which is added in quadrature to the other uncertainties.

- *Hard process uncertainty.* There are two arbitrary scales introduced in the ME calculation, the renormalisation and factorisation scales, and varying these scales gives terms which scale as one power of α_s higher than the highest power considered in the calculation. These scales are therefore independently scaled up by a factor of 2 and down by a factor of 0.5, giving two uncertainties which are taken as an estimate of possible higher order corrections to the ME. Since different processes occur at different scales and potentially have different corrections, this uncertainty is treated as uncorrelated between different processes.

- *Parton shower modeling.* The uncertainty on the parton shower can also be estimated by varying the (generator specific) scale associated with the branchings in the shower, as well as other parameters which affect the kernels used to give the probabilities for these branchings. These uncertainties are summed in quadrature, but treated separately for ISR and FSR.

- *Simulation statistical uncertainty.* The MC events also have a small statistical uncertainty, which is treated by allowing each bin of the distributions used in the signal extraction to fluctuate independently, using the Barlow-Beeston method [95].

- *Top p_T reweighting.* The effect of not applying the top p_T reweighting described in section 6.7.1 to the SM $t\bar{t}$ sample is taken as a systematic uncertainty on this higher order correction.

8.1.2 Collider uncertainties

The second set of uncertainties considered relate to the measurements of the colliding beams:

- *Luminosity.* The measurement of the integrated luminosity has an uncertainty of 1.2% in 2016 [64], 2.3% in 2017 [65] and 2.5% in 2018 [66]. These uncertainties are treated as uncorrelated between the different data-taking periods.

- *Pileup.* The uncertainty on the PU distribution is estimated by re-performing the PU reweighting for while varying the total $p\bar{p}$ cross section, which is used to estimate the number of PU collisions in data, within its uncertainty.

8.1.3 Detector uncertainties

The final set of uncertainties is related to how well different objects are reconstructed in the detector, estimated from the uncertainties on the scale factors used to correct for detector response:

- *Trigger.* The trigger scale factors are varied within the uncertainties described in section 6.4 to account for the uncertainty on the difference of these triggers in data and MC.

- *Level 1 trigger prefiring.* The uncertainty on the Level 1 trigger prefiring issue in 2016 and 2017 discussed in section 6.7.2 was accounted for by varying the probability for each object to prefire within its uncertainties.

- *Leptons.* The uncertainties on the electron reconstruction and ID SFs (described in section 6.5.1) and on the muon ID and isolation SFs (see section 6.5.2) are taken as systematic uncertainties on these objects.

- *Jet PU ID.* The weights for the jet PU ID described in section 6.5.3 are recomputed for variations of the SFs within their uncertainties, which come from the fit to data used in the derivation of the SFs, propagation of the jet energy scale uncertainties to this derivation, and a systematic term estimated from MC. This is performed separately for genuine jets (those with $\Delta R < 0.4$ of a generator jet) and PU jets.

- *b-tagging efficiency scale factors.* The uncertainties on the b-tagging SFs are estimated by varying different parameters in the computation of the SFs, such as details of the MC simulation used in this estimation, the cuts used to select the sample used for the estimation, and propagation of the jet energy scale and pile-up uncertainties [96]. These are split into a component which is correlated across the years, and an uncorrelated component for each year. The effect of varying these uncertainties is evaluated separately for heavy (bottom and charm) and light (up, down, strange and gluon) jets.

- *Jet energy scale.* The systematic uncertainties identified in the derivation of the jet energy corrections described in section 6.5.3 are split into 27 sources [97], which cover various effects, such as the uncertainty on the modelling of the parton showers used in the derivation of these corrections, and differences between estimating the impact of pileup via different methods. Seven of these uncertainties, relating to the forward endcap regions $|\eta| > 2.5$ and the forward hadronic calorimeter, are neglected since jets in this region are not considered and these jets tend to contribute relatively little to p_T^{miss}. For each of the remaining uncertainty sources the selection is re-performed with the jets varied according to their uncertainties, which can cause jets to fall in and out of acceptance, but also affects quantities derived from these, most notably the p_T^{miss}.

- *Jet energy resolution.* The selection is also reperformed with jets smeared to different degrees according to the uncertainty on the resolution of the jets, again including the propagation to derived variables such as p_T^{miss}.

- *Unclustered energy.* The majority of the uncertainty on the p_T^{miss}measurement comes from the energy of the jets, which is described by the previous two uncertainties. However there are also tracks and calorimeter deposists not associated with jets but which are included in the p_T^{miss}calculation. The selection is therefore also re-performed with this unclusterred energy varied within its uncertainties.

One could also consider the uncertainty on the measured energy of the electrons and muons, which might also affect the measurement of p_T^{miss}, however these are only about 1.5% and 0.3% respectively, compared to about 30% for jets, so these effects are neglected.

8.2 Smoothing of Systematic Uncertainties

Since the systematic uncertainties are estimated as variations of MC events, they have an associated statistical uncertainty. Generally this uncertainty is small and can be neglected, however in some cases these can be comparable to or even larger than the systematic effect being estimated. This can lead to the uncertainty having either an overly large or overly small impact, depending on the direction of the fluctuations, and potentially being spuriously constrained by the fit, as the statistical fluctuations do not fit to data. Three cases are identified in this analysis where the statistical uncertainties are large enough to be problematic. Firstly, the uncertainties which are evaluated by re-performing the selection with the energy of certain objects varied (i.e., jet energy scale, jet energy resolution and unclusttered energy) can have a large statistical uncertainty due to events randomly passing in and out of acceptance, or migrating between bins. Secondly, the parton shower uncertainties often have a large variance of the weights applied to each event, which can lead to large statistical uncertainties due to a few high weight events. Finally the hard process uncertainties are split into different backgrounds, so the statistical uncertainties are again problematic for many of the individual backgrounds.

In all these cases one would expect the systematic effect to be some sort of smooth distribution, which could be recovered by averaging out the statistical fluctuations. Since the expected functional form of the systematic variations is unknown, it does not make sense to fit to a polynomial of a particular order. Instead the LOWESS (LOcally WEighted Scatterplot Smoothing) algorithm is employed, as proposed in reference [98]. As the name suggests, this performs a weighted linear fit for each bin within a certain window, with the size of the window defined by a parameter called the bandwidth, b. Since all the uncertainties considered for this procedure are expected to be symmetric, this fit is performed for the difference between the "up" and "down" variations rather than the difference of these to the nominal to maximise the genuine systematic shape, then after smoothing the "up" and "down" templates are normalised to the same value as before smoothing.

The weight given to each bin in the linear fit is:

$$w = \frac{1}{\Delta\sigma^2}\left(1 - \left(\frac{|d|}{b}\right)^3\right)^3 . \tag{20}$$

Here $\Delta\sigma$ is the fractional uncertainty on the bin, and d is the distance of the bin being weighted to the bin for which the fit is being performed. Since the bandwidth both determines the overall size of the window of bins to be considered for each point and their weights, it has a considerable impact on the final fit; small bandwidths can fit most genuine shapes, but at the cost that it can also fail to remove the statistical noise, while larger bandwidths give smoother shapes, averaging out most statistical noise at the risk of also losing genuine features. To select the correct bandwidth for each systematic uncertainty distribution, a cross-validation procedure is employed: the

dataset is split into a "training" and "test" set, and a smoothed distribution is produced from the training set, and the χ^2 of fitting the test dataset is evaluated. A range of bandwidths increasing in steps of 0.1 from 0.2 to 1 are considered, and the cross-validation is performed 200 times (with different splittings into training and test sets for each iteration) for each bandwidth, and the bandwidth with the lowest average χ^2 is selected. For many of smaller uncertainties, however, the statistical noise completely dominates the original shape to the point it is hard to determine what the smooth shape should be. Therefore a further χ^2 fit is performed to a flat distribution, and if this gives a better fit it is used rather than the LOWESS smoothing.

An example of this smoothing is shown in figure 43 for the jet energy resolution uncertainty on the $t\bar{t}$ background across the different SRs and the DY CR (the distributions in the $t\bar{t}Z$ CR were not smoothed since there are reasonably good MC statistics in every bin). One can clearly see significant statistical fluctuations in the original distributions in the bins with lower MC statistics, particularly notably in the $t\bar{t}$ + DM OF SR. The smoothed distributions are mostly linear, but plateau in the final bins in many cases, even decreasing slightly in the final bins of the t + DM OF SR. These distributions indicate the smoothing algorithm is performing well - the smoothed distributions agree well with the original distributions in the bins with small statistical uncertainty, then in the later bins agree well with one's expectation of the physical shape before the fluctuations.

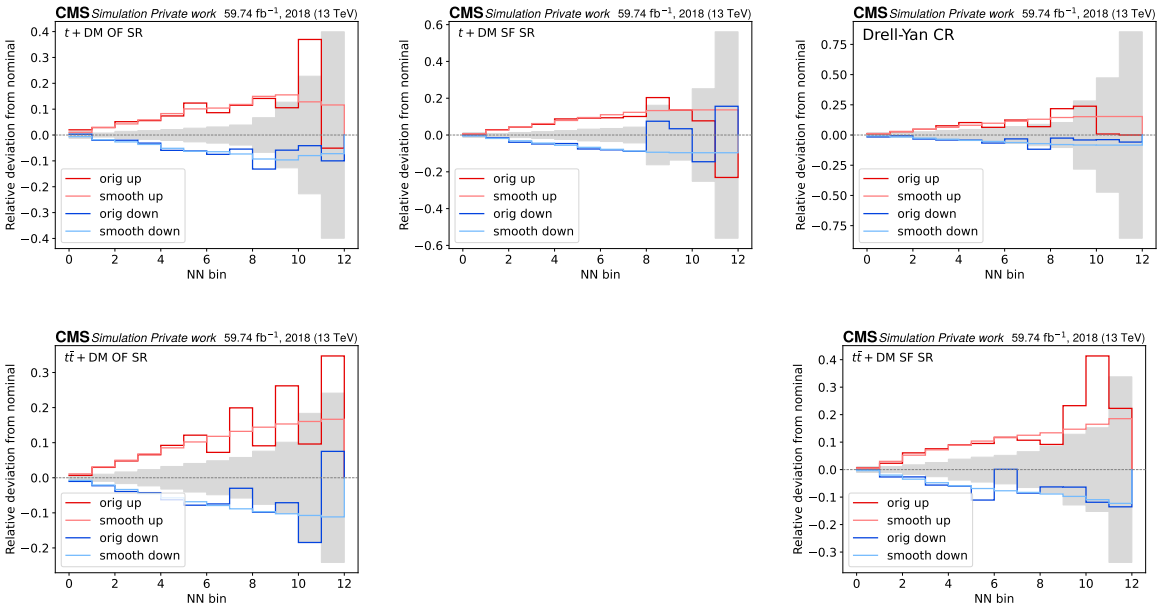


Figure 43: Original and smoothed relative deviation from nominal of the jet energy resolution uncertainty on the $t\bar{t}$ background for the t + DM OF SR (top left), t + DM SF SR (top center), DY CR (top right), $t\bar{t}$ + DM OF SR (bottom left) and $t\bar{t}$ + DM SF SR (bottom right). Statistical uncertainty on the nominal is shown in the grey band. NN bins are shown as being equal width in plot to improve visibility.

9 Fit and Results

9.1 Statistical Fit

To determine whether the observed data indicates the presence of the top + DM signal in addition to the expected backgrounds, given the expected statistical and systematic uncertainties, one must perform a statistical fit, to determine the probability of obtaining the observed data given some amount of signal compared to the background-only case. More concretely, for any given signal strength parameter, $r = \sigma/\sigma_{\text{th}}$ (i.e. the ratio of the signal cross section to that predicted by the signal model) one can evaluate the likelihood of observing the number of counts in each bin, n_i, which is proportional to the probability of observing the given number of counts from a Poisson distribution in each bin:

$$p(n_i|r) = \frac{(b_i + rs_i)^{n_i}}{n_i!} e^{b_i + rs_i}, \tag{21}$$

where s_i and b_i are the signal and background expectations in bin i, respectively.

Systematic uncertainties can then be accounted for by introducing a nuisance parameter for each uncertainty, which scales the expected distributions according to the templates for the systematic uncertainty; these will henceforth be denoted as a vector of nuisance parameters, $\vec{\theta}$. Since these uncertainties are expected to have effects that vary around their central value, an additional constraint term, $v(\vec{\theta})$, which gives the probability of the nuisance parameters taking a certain value, is included in the likelihood. In this analysis the constraints are taken to be Gaussian on all uncertainties except the MC statistical uncertainties, which are handled using the Barlow-Beeston approach [95], and rate parameters, which link the rate of a particular process between control and signal regions and have a uniform constraint, allowing the rate to be determined from the CR without a bias from the initial MC value. The likelihood is therefore:

$$L(n|r, \vec{\theta}) = \prod_i p(n_i|r, \vec{\theta}) v(\vec{\theta}), \tag{22}$$

where $p(n_i|r, \vec{\theta})$ is the Poisson probability of observing n_i events given the expected number of events for the given signal strength and nuisance parameter values. The nuisance parameters can then be removed again by taking the value of the nuisance parameters which maximises the likelihood, which is referred to as profiling:

$$L(n|r) = \max_{\vec{\theta}} L(n|r, \vec{\theta}). \tag{23}$$

The best fit value of the signal strength can be obtained by maximising this likelihood,

or more commonly minimising the negative log of the likelihood, which is numerically easier.

To perform more advanced statistical tests, a test statistic is constructed using the log-likelihood ratio:

$$q(n|r) = -2\ln\left(\frac{L(n|r)}{\max_{\tilde{r}\in[0,r]} L(n|\tilde{r})}\right). \tag{24}$$

If a best-fit signal strength notably larger than 0 is observed, the first test performed is to check the significance of the excess, which is done by computing the confidence level for the background:

$$\mathrm{CL}_B = \int_{q(n|0)}^{\infty} f(q_0|0)dq_0. \tag{25}$$

Here $f(q_0|0)$ is the probability distribution for q_0, which in this case is estimated using an Asimov dataset [99] (this approach relies on the asymptotic approximation that each bin has sufficient statistics that it can be approximated by a Gaussian distribution). CL_B is the p-value, or probability for such a distribution or a more extreme one to arise from fluctuations in the data in the absence of a signal. This is then converted to the corresponding number of standard deviations of a Gaussian distribution necessary to obtain such a fluctuation. Since the Standard Model has been very well studied, and the large number of searches carried out at the LHC means one would expect some measurements to see some signal-like excesses due to fluctuations in the data, rather stringent requirements are required for an excess to be considered significant: only fluctuations of at least 3σ ($\mathrm{CL}_B < 2.70 \times 10^{-3}$) are considered evidence for a process, while 5σ ($\mathrm{CL}_B < 2.87 \times 10^{-7}$) is required to claim a discovery.

If no significant excess is observed, it is common to set limits on the signal strength, i.e., determine for what ranges of r one can exclude the presence of signal. The naïve way to do this would be to compute the confidence level for the signal+background hypothesis at the signal strength under consideration, r:

$$\mathrm{CL}_{S+B} = \int_{q(n|r)}^{\infty} f(q_r|r)dq_r, \tag{26}$$

and exclude points with CL_{S+B} below a certain value. However for searches such as this one where the background is much larger than the signal, this can give counter-intuitive results since the consistency of the results with the background dominates this calculation. These issues can be mitigated by using the CL_S criterion [100, 101]:

$$\mathrm{CL}_S = \frac{\mathrm{CL}_{S+B}}{1 - \mathrm{CL}_B}. \tag{27}$$

A given value of the signal strength, r, is then said to be excluded if it has $\mathrm{CL}_S < 0.05$ - the less stringent requirement compared to the threshold for discovery reflects an a-prioi bias towards the SM on the basis of previous measurements.

For this analysis, the fit was performed simultaneously on the NN output histograms for the relevant mediator (scalar or pseudoscalar) in all SRs and the DY CR, as well as the $n_{\mathrm{jet}}/n_{\mathrm{b-jet}}$ histogram described in section 7.2 in the $t\bar{t}Z$ CR. This was performed using a program called "combine" [102], which was developed for this purpose in the CMS collaboration, based on the RooStats package [103].

9.2 Simplified Dark Matter Results

9.2.1 Analysis Optimisation

As has already been discussed, the analysis was "blinded" during optimisation, i.e. only a small fraction of the data was used in plots to avoid introducing a bias in the selection and corrections that were applied. For the fit a different approach was taken; no data was used and instead expected limits were set using an Asimov dataset constructed from the MC background with the full luminosity, as well as expected significances using an Asimov dataset constructed from the MC signal and background. These were used to optimise the search strategy, in particular the NN described in section 6.11 to maximise sensitivity to the signal. The expected limits for this analysis in terms of the signal strength modifier as a function of the mediator mass are shown in figure 44, which also show the breakdown of the improvements in this analysis with respect to the 2016 dilepton channel limits published in reference [55] (orange line). This can be directly compared with the limits for this analysis using only the 2016 data and only considering the $t\bar{t}$ + DM signal (pink line)- this alone demonstrates a 35% improvement compared to the previous result thanks to optimisations in this analysis, most notably the use of NNs for the signal extraction. Comparing with the 2016 limit considering both signals (blue line) demonstrates the improvement from also targeting the t + DM signal, which gives an improvement of between 15% for the low mediator masses and 40% for the highest mediator masses, where the t + DM cross section starts to be comparable to the $t\bar{t}$ + DM cross section (see figure 12). Finally the red line shows the improvement from also considering the 2017 and 2018 datasets, which again gives an improvement of approximately 40%.

Expected impact distributions were also produced under both the background only and signal+background hypotheses. These are plots showing the pulls (the distribution of the post-fit nuisance parameters relative to the pre-fit)[1], and the impact on the signal

[1]For the expected distributions the central value of the pulls will always be zero since this is the value

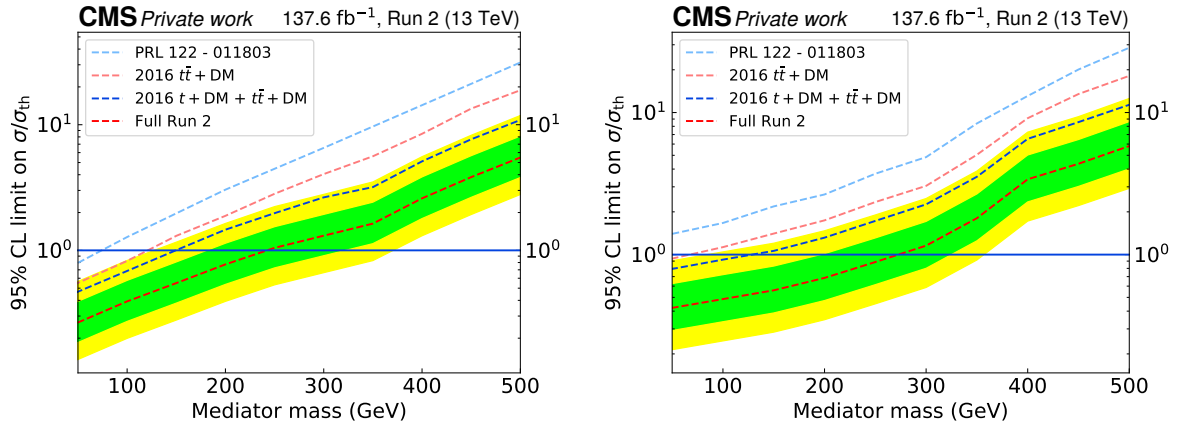


Figure 44: Median expected limits in terms of signal strength as a function of mediator mass for the scalar mediators (left) and pseudoscalar mediators (right). The light blue line shows the median expect limits for the last published CMS search in the dilepton channel [55], while the pink and dark blue lines show the median expected limit for the 2016 dataset in this search when considering only the $t\bar{t}$+DM signal and both the $t\bar{t}$+DM and t+DM signals, respectively, for comparison. The red line shows the median expected limit considering the full Run 2 dataset, while green and yellow bands indicate the range of possible limits for 1σ and 2σ variations of the background from the median value.

strength of varying these nuisances by 1σ of their uncertainty in either direction[2]. This allows one to see which uncertainties are most relevant in the final fit, and whether any nuisance parameters are constrained to have smaller uncertainties than the initial templates. For some uncertainties, for instance the theoretical uncertainties, it is expected that the data will contain additional information which may allow it to constrain the uncertainties, however if the data constrains an uncertainty to which it is not expected to be sensitive this may indicate a problem in the fit, such as the statistical fluctuations discussed in section 8.2. The expected impacts are shown in figures 45 and 46. These are generally as one would expect: the largest impact uncertainties are the renormalisation and factorisation scale uncertainties on the signal, as these would directly change the expected signal cross-section, changing the signal strength of any given excess. The importance of other uncertainties varies between the signal points, but other significant uncertainties include the jet energy resolution and unclusttered energy uncertainties, which impact p_T^{miss} and thus $M_{\text{T2}}^{\ell\ell}$, leading to a signal-like fluctuation in the background; the theoretical uncertainties, which can change both the shape of distributions and the relative rate of backgrounds, and the rate parameters for the $t\bar{t}Z$ process (which affects the rate of the very signal-like $t\bar{t}Z$ ($Z \to \nu\bar{\nu}$) process) and DY (where separate parame-

used to generate the Asimov dataset, though they can be constrained by the Asimov data

[2]The other uncertainties are profiled for each value of the nuisance being scanned, which allows accounting for the interaction of the various uncertainties with each other.

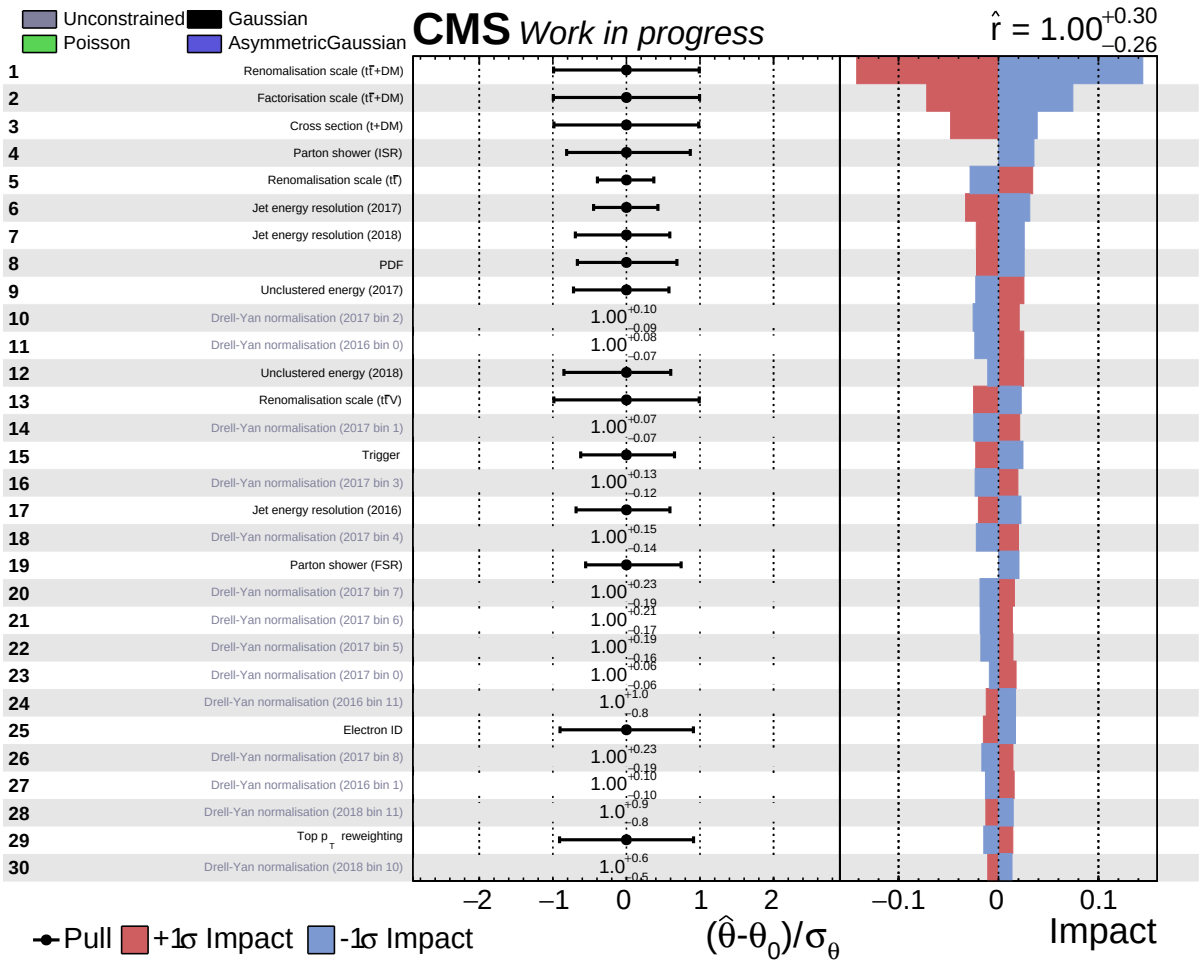


Figure 45: Expected impacts of the 30 highest ranked uncertainties for the 100 GeV scalar mediator, produced from a signal+background Asimov. The middle panel shows the pulls, while the right panel shows the impact on the signal strength of scaling the uncertainty 1σ up (red band) and down (blue band).

ters are used for each bin of the distribution, allowing for a wide range of possible shape effects, depending on the distribution in the DY CR). The theoretical uncertainties are often constrained, which is not a cause for concern as it is known experimental measurements are often more precise than these LO or even sometimes NLO predictions. The collider and detector uncertainties, on the other hand, are generally unconstrained, with the exceptions of the jet energy resolution and unclustered energy uncertainties, and the b-tagging uncertainties. This analysis is very sensitive to the jet energy resolution and unclustered energy uncertainties, and these have rather conservative uncertainty estimates due to some poorly understood detector effects in 2017 and 2018, so it is not surprising this fit is able to constrain these uncertainties relative to the initial estimates. Similarly, the b-tagging scale factors used in this search are not derived in regions enriched in $t\bar{t}$ processes to avoid unwanted correlations, however since $t\bar{t}$ processes provide

some of the best constraints on b-tagging efficiencies and this analysis includes a binning in the number of b jets, it is expected that the fit will be able to constrain these scale factors better than the original derivation.

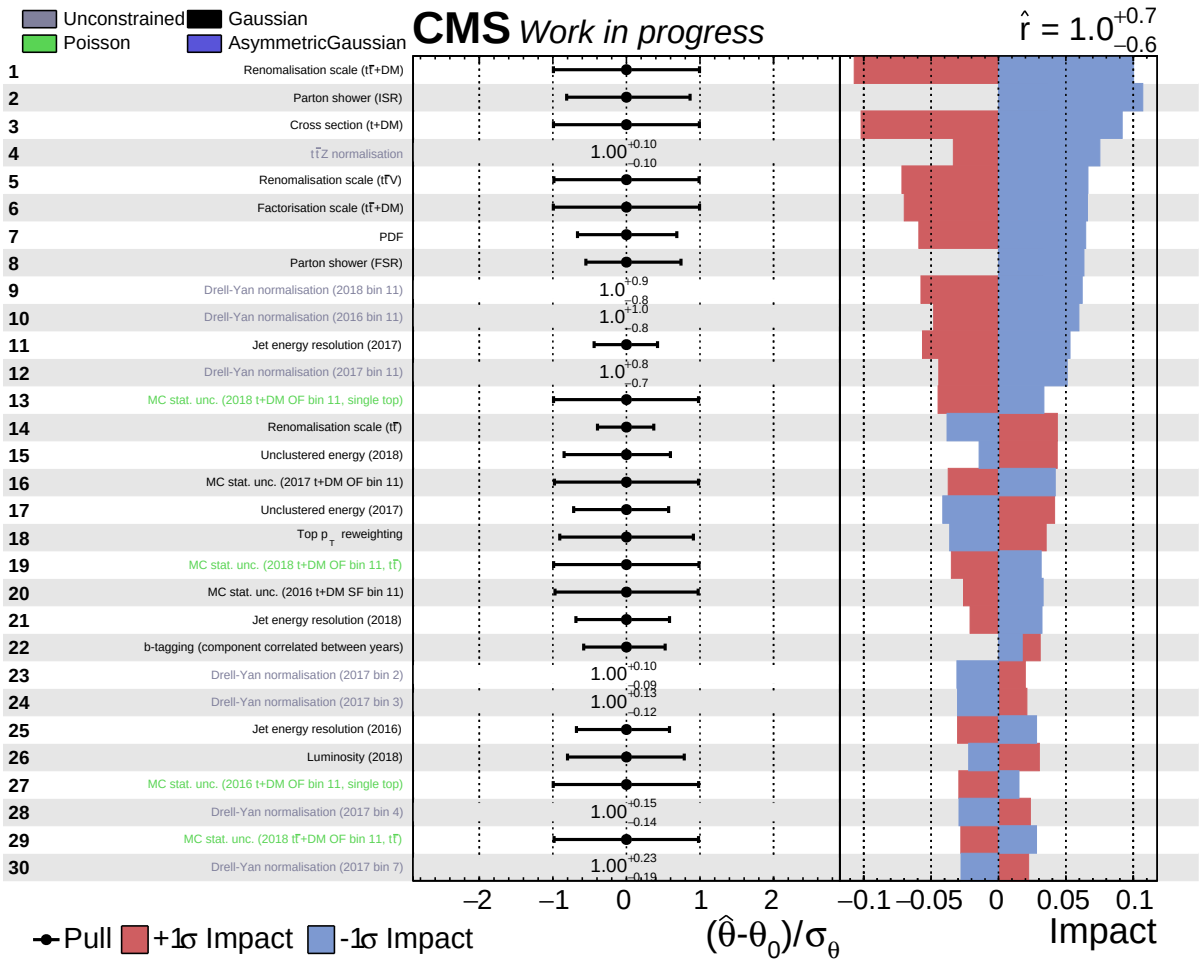


Figure 46: Expected impacts of the 30 highest ranked uncertainties for the 300 GeV scalar mediator, produced from a signal+background Asimov. The middle panel shows the pulls, while the right panel shows the impact on the signal strength of scaling the uncertainty 1σ up (red band) and down (blue band).

9.2.2 Unblinding Checks

Once the optimisation of this analysis was completed, it was "unblinded", i.e. the fit was performed including all the data. However to allow early identification of issues without the bias of seeing the full results, a multi-stage procedure was followed: first the systematic impacts were produced without viewing the measured signal strength, then plots of the goodness-of-fit to the data, and only if these did not indicate problems in the fit would the full unblinded distributions be produced, the signal strength be extracted

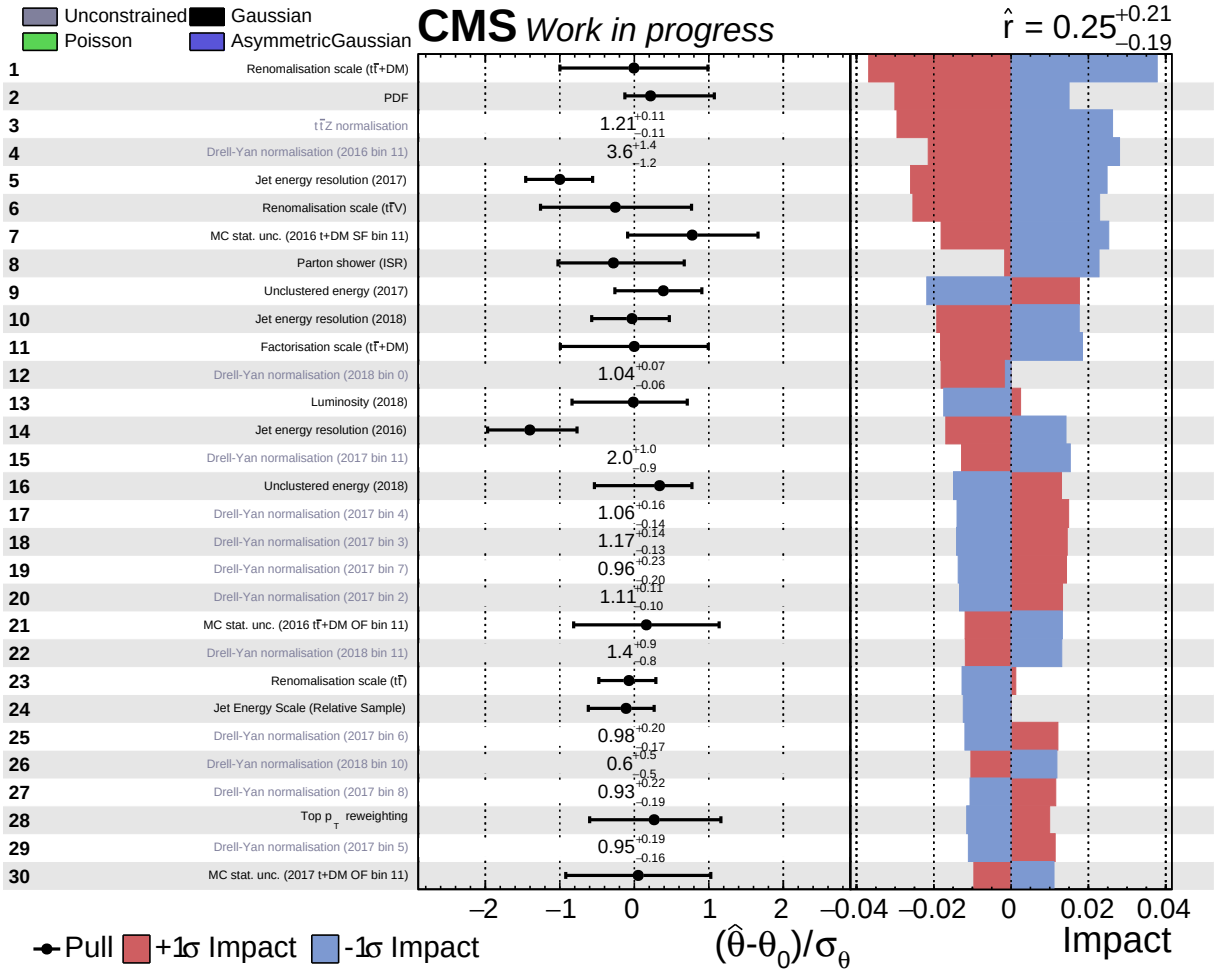


Figure 47: Unblinded systematic impacts of the 30 highest ranked uncertainties for the 100 GeV scalar mediator.

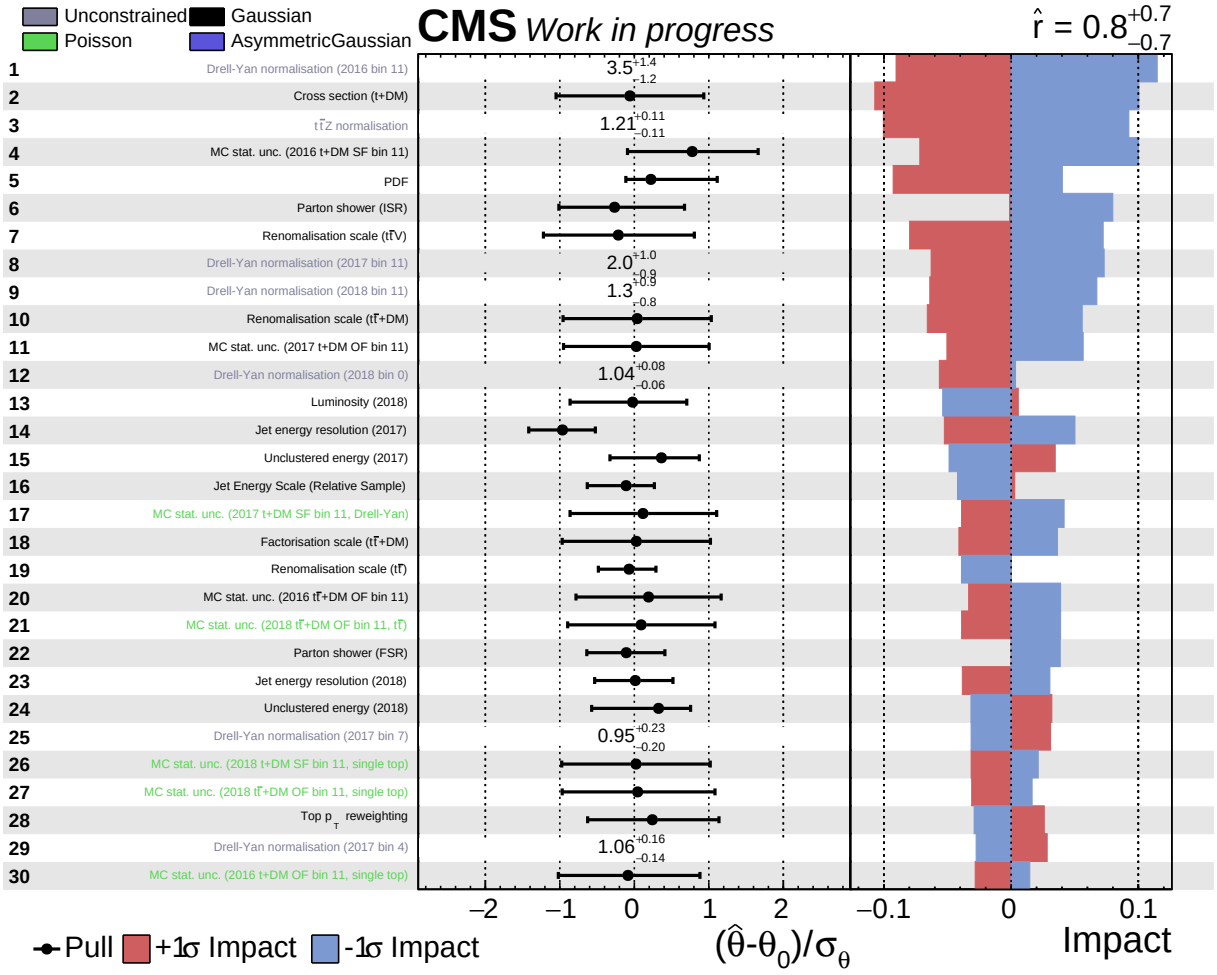


Figure 48: Unblinded systematic impacts of the 30 highest ranked uncertainties for the 300 GeV scalar mediator.

and significances and/or limits be set. The first distributions produced were the impacts, where the signal strength is allowed to float to its best-fit value (figures 47 and 48). These generally indicate good modelling; no uncertainties are significantly more constrained than in the expected impact distributions, and most nuisance parameters are within the 1σ uncertainties of the initial value. The exceptions are the jet energy resolution uncertainties, which are pulled down by slightly over 1σ in 2017, and about 1.5σ in 2016, however it was decided it was not necessary to investigate the pull of these uncertainties further since there are known to be some detector issues affecting these uncertainties, the uncertainties have relatively low impact on the signal strength, and the degree of the pull is not very dramatic. The rate parameters for the data-driven background estimation also show values notably different from 1 - firstly, the $t\bar{t}Z$ background is scaled by a factor of about 1.2, which is consistent with the value measured in reference [92]. The DY rate parameters also differ significantly from 1, indicating a significant shape effect coming from the mis-modelling of the DY p_T^{miss} distribution.

The next check was that the MC simulation could describe the data well when allowing the signal strength parameter to freely float. This was done using a "saturated" goodness-of-fit test [104], which computes the likelihood ratio of the best-fit value from the model to a "saturated" model which is equal to the data in every bin, $-2\ln\lambda$ (which is equal to the well-known χ^2 test for Gaussian distributed data). This was run for 1000 "toy" distributions for variations of the statistical and systematic uncertainties, and the result of this test for the data was compared to this distribution (figure 49). The result for data sits well within the range of the toy distributions, indicating the fit of the data to the MC simulation works well.

9.2.3 Results

Since the impact plots and goodness-of-fit tests did not indicate an issue with the fit, the full fit input distributions including data were produced. These are shown in figures 50 and 51, including the ratio of the data both to the pre-fit background MC and the post-fit background MC (i.e. the values of the background MC with the systematic nuisances and rate parameters pulled to their profiled values for the best-fit signal strength).

The distributions generally show good agreement, especially for the post-fit distributions, with the exception that there is an excess of data relative to MC in the final NN bins in many of the signal region distributions, most notably in the SF $t + \text{DM}$ region in 2016 and 2017. This excess is consistent with the signal hypothesis, as can be seen from table 4, which gives the best fit value of the signal strength, r, for each mediator hypothesis considered, as well as the corresponding cross sections and signal significances. The excess generally fits well to all the signal points, except for the very low mass scalar mediators, which have a significantly less hard p_T^{miss} spectrum than the other signal points, and hence would also lead to an excess in the penultimate NN bin, which is not observed. The signal strength increases for the higher mediator masses due to the lower expected cross section, while the corresponding cross sections decrease as the signals are

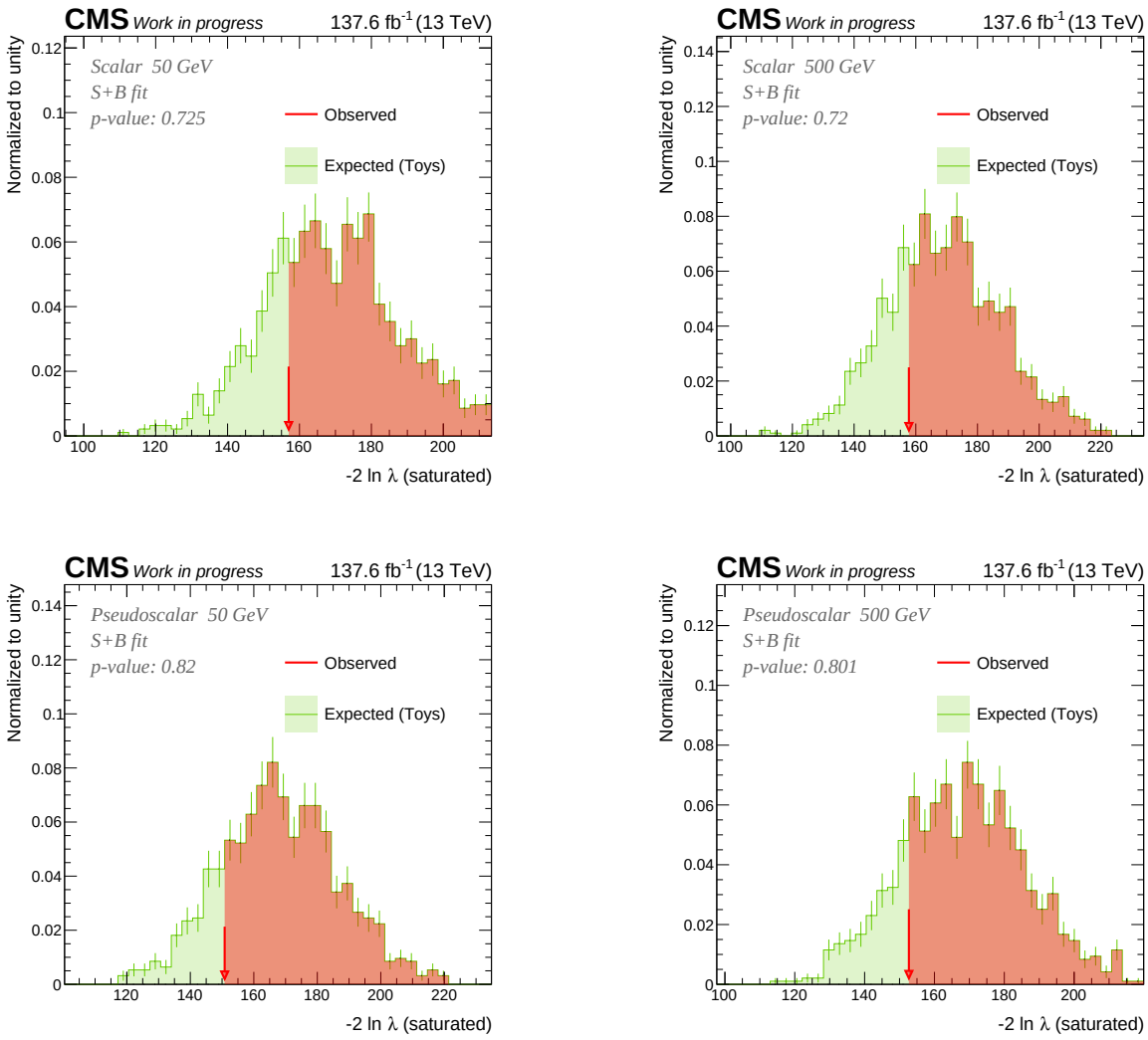


Figure 49: The goodness-of-fit plots for 50 GeV (left) and 500 GeV scalar (top) and pseudoscalar (bottom) meditors. The histogram shows the distribution of $-2\ln\lambda$ for the "toy" datasets, while the red arrow indicates the $-2\ln\lambda$ for data. The red region indicates the range of toy datasets with a worse $-2\ln\lambda$ than the one in data, and the area of this region relative to the total number of "toy" datasets is used to calculate the p-value.

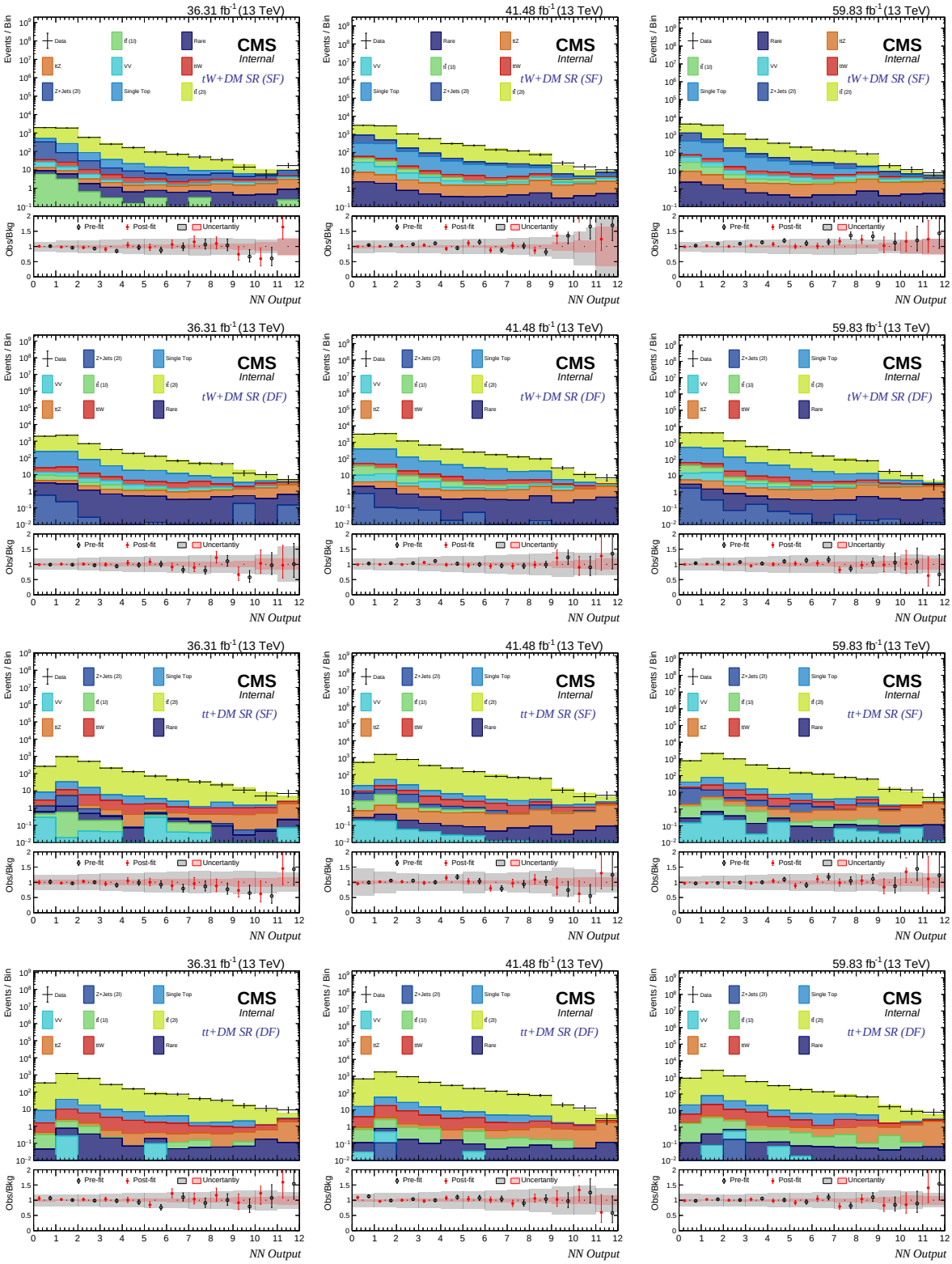


Figure 50: Top to bottom: NN output distributions in the $t + \text{DM}$ SR SF, $t + \text{DM}$ SR OF, $t\bar{t} + \text{DM}$ SR SF and $t\bar{t} + \text{DM}$ SR OF signal regions for 2016 (left), 2017 (middle) and 2018 (right), including the ratio of data to MC background for both the pre- and post-fit distributions of a signal and background fit for the 100 GeV scalar mediator.

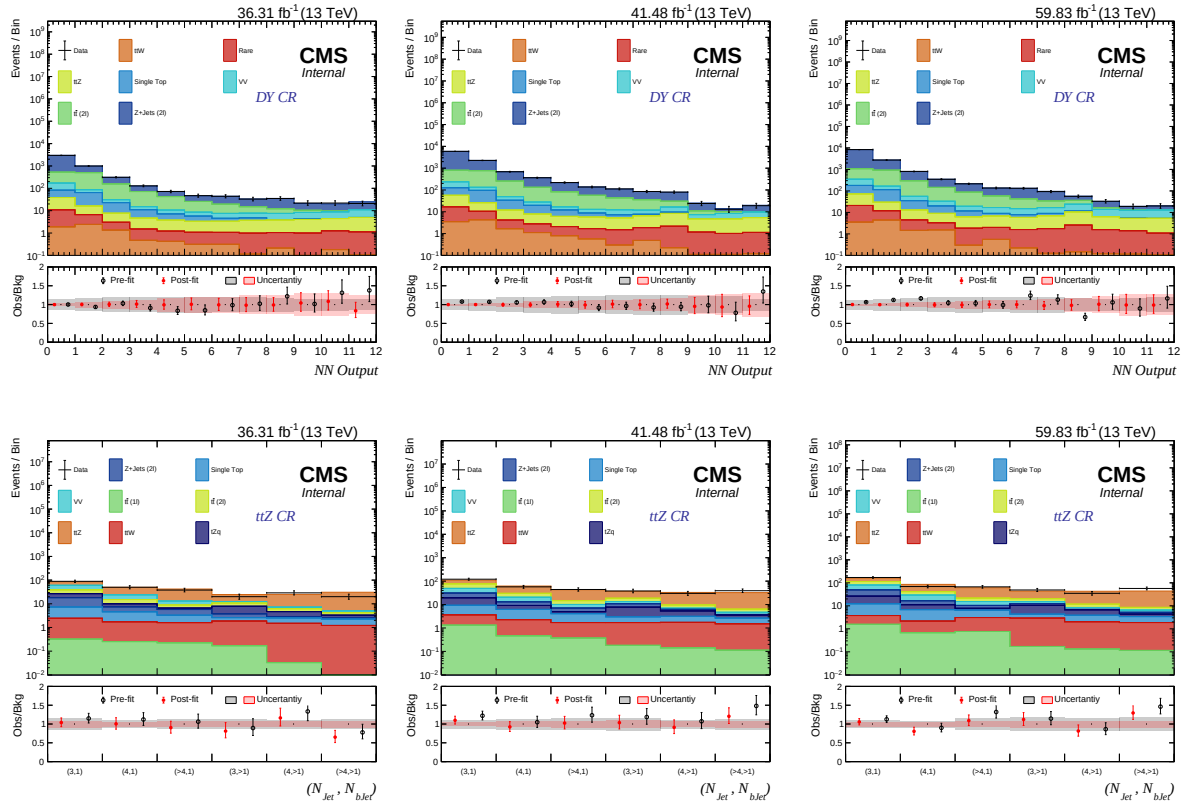


Figure 51: The NN output distributions in the DY CR (top) and $n_{\text{jet}}/n_{\text{b-jet}}$ distributions $t\bar{t}Z$ CR (bottom) for 2016 (left), 2017 (middle) and 2018 (right), including the ratio of data to MC background for both the pre- and post-fit distributions of a signal and background fit for the 100 GeV scalar mediator.

more peaked in the final bin, though start to plateau for the highest mediator masses.

	Scalar				Pseudoscalar			
Mediator mass	r	$\sigma_{t+\text{DM}}$ (pb)	$\sigma_{t\bar{t}+\text{DM}}$ (pb)	significance	r	$\sigma_{t+\text{DM}}$ (pb)	$\sigma_{t\bar{t}+\text{DM}}$ (pb)	significance
50.0	0.14	1.84	47.30	1.08	0.29	0.69	4.78	1.39
100.0	0.25	1.42	20.21	1.31	0.35	0.92	5.45	1.45
150.0	0.41	1.38	10.83	1.48	0.41	1.12	5.73	1.44
200.0	0.57	1.35	6.60	1.46	0.50	1.20	5.57	1.42
250.0	0.78	1.36	4.71	1.46	0.65	1.26	5.28	1.46
300.0	0.77	1.03	2.76	1.15	0.90	0.97	3.71	1.54
350.0	1.02	1.04	2.42	1.21	1.33	0.80	2.80	1.43
400.0	1.66	1.12	2.40	1.24	2.68	0.73	2.35	1.54
450.0	2.49	1.13	2.25	1.25	3.18	0.77	2.32	1.42
500.0	4.23	1.36	2.62	1.47	4.00	0.98	2.80	1.34

Table 4: Best-fit signal strength, r, cross sections, σ, and significances in number of standard deviations for the considered scalar and pseudoscalar mass hypotheses.

While the excess of about 1.5σ for the most compatible mediators could be a first hint of new physics, it is also reasonably compatible with the SM-only hypothesis, and far

below the 3σ threshold to claim evidence for the new process. Limits are therefore set on this process, as shown in figure 52. Due to the excess these are weaker than the median expected limits, however scalar and pseudoscalar mediator masses of up to about 155 GeV are still excluded for the reference cross section.

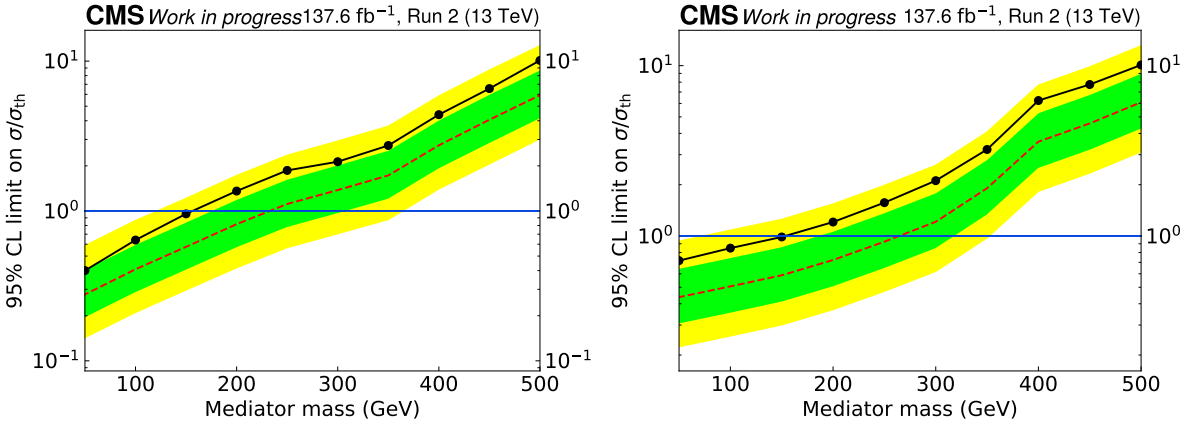


Figure 52: Observed limits for the scalar (left) and pseudoscalar (right) DM mediator.

9.2.4 Comparison with Existing Results

As discussed in section 6.1, other searches have already been performed for this model using the LHC Run 2 dataset. The first is a reinterpretation of a search for stop quarks by the CMS experiment, which was reinterpreted as a search for $t\bar{t}$ + DM [58]. This search did not see an excess for any values of the mediator mass, and set limits up to a mediator mass of about 400 GeV for both the scalar and pseudoscalar mediators, as shown in figure 53. The expected results presented in this thesis are about 40% stronger than the two lepton channel of this reinterpretation for low mediator masses and 60% stronger for high mediator masses, due to also targeting the t + DM final state and the use of MVA techniques to extract the signal. However since this thesis observed a slight excess not seen in the reinterpretation, the observed limits were similar.

The second published Run 2 analysis comes from the ATLAS collaboration, and is also a reinterpretation of a search for stop quarks [59]. This analysis included both $t\bar{t}$ + DM and t+DM signals, however unlike this work the selection was not re-optimised to target the t + DM signal by changing the jet selection criteria. An excess was observed in the all-hadronic $t\bar{t}$ decay mode, which lead to a slight excess for lower mediator masses in the final combination, however the final limits set are close to those of the CMS stop quark search (see figure 54). Despite not being optimised for the t+DM signal nor using MVA techniques, the ATLAS dileptonic limit is about twice as strong as the limit presented in this work. Studies on the origin of this large difference in sensitivity are ongoing, with the most likely candidate being the better p_T^{miss} resolution of the ATLAS detector,

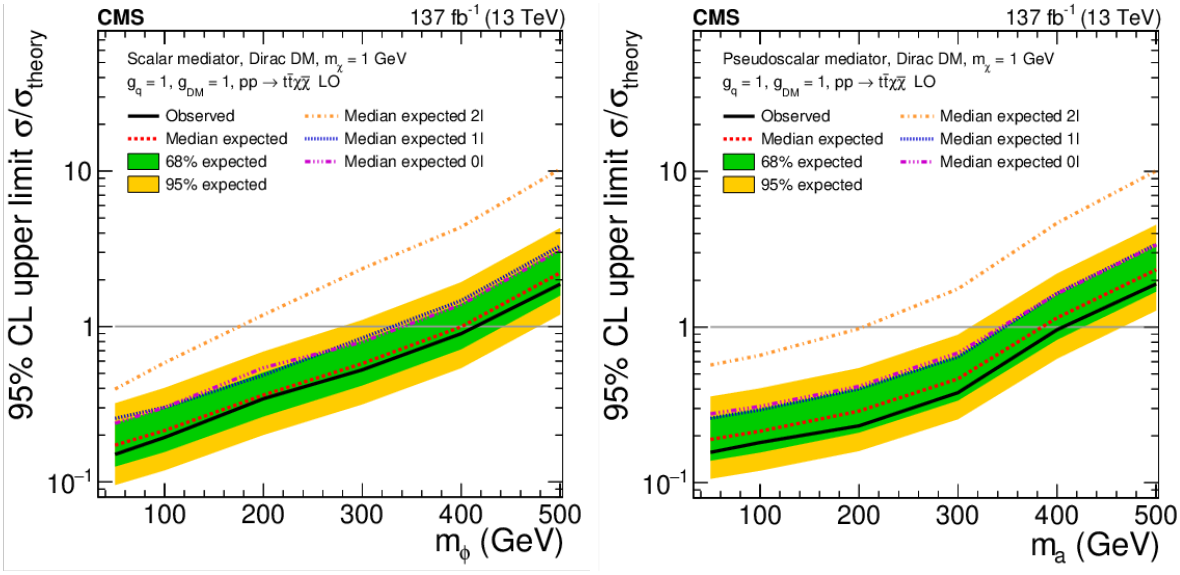


Figure 53: Limits on the $t\bar{t}$ + DM process published in [58] for the scalar (left) and pseudoscalar (right) DM mediator, including the breakdown into expected limits by $t\bar{t}$ decay channel and the combined expected and observed limit.

which also leads to better $M_{\mathrm{T2}}^{\ell\ell}$ resolution, which can dramatically reduce the dominant $t\bar{t}$ background, since this can only enter the signal region due to $M_{\mathrm{T2}}^{\ell\ell}$ mismeasurement.

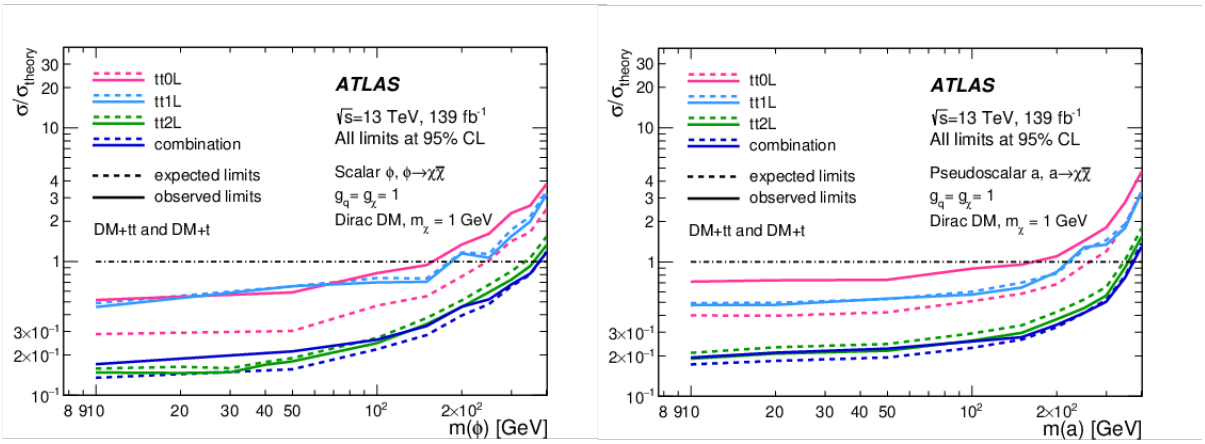


Figure 54: Limits on the top + DM process published in [59] for the scalar (left) and pseudoscalar (right) DM mediator, including the breakdown into expected and observed limits by $t\bar{t}$ decay channel and for the combination of all channels.

The 95% exclusion limits of these two results are close to the best-fit values for the signal strength observed in this work. This suggests the excess observed here is unlikely to be due to the top + DM signal model being searched for. However since the excess is concentrated in the t + DM signal regions, which include a slightly different phase

space to the $t\bar{t}$ + DM based searches, it may be that the signal observed here actually comes from a different process - for instance a similar size of excess for observed by an ATLAS search for dileptonically decaying tW+DM in the context of a 2HDM+a model [105]. Future searches should therefore pay particular attention to this region to determine if these small excesses are the first hint of new physics, or due to fluctuations or mismodelling of background processes.

9.3 Outlook

This result will form part of a combined search across all decay channels targeting both the $t\bar{t}$ + DM and t + DM signals, which aims to provide the most sensitive search for this model with the LHC Run 2 dataset. However a significant range of parameter space remains to be explored - some thermal freeze-out models have mediator masses up to about 1 TeV, and the cross section to achieve the observed relic density may be lower than the one for $g_q = g_\chi = 1$, depending on the mass of the dark matter, m_χ. Fortunately this search in the dileptonic channel is mainly limited by statistics, and hence adding more data from future LHC runs could significantly expand the reach of this analysis; figure 55 shows that one can get a significant improvement in the expected limit just by scaling up the luminosity to include the data being taken in the ongoing Run 3 of the LHC, which is expected to provide a further 160 fb^{-1} of luminosity, and then the planned high luminosity-LHC (HL-LHC) upgrade, which is expected to provide 3000 fb^{-1}, will extend the reach even further.

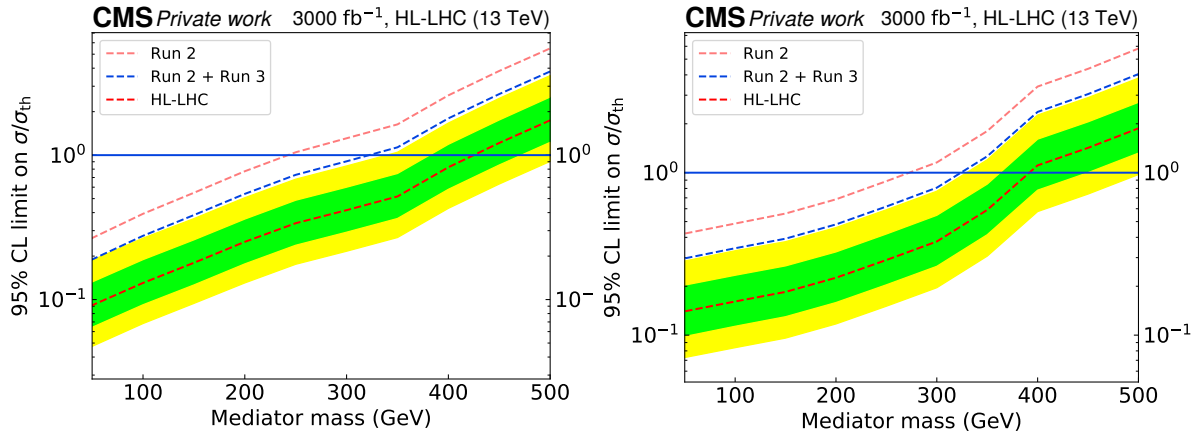


Figure 55: Projected expected limits for this analysis, with the luminosity scaled to the expected values for Run-2 and Run-3 combined (300 fb^{-1}), and the full high luminosity run of the LHC (3000 fb^{-1}), for the scalar (left) and pseudoscalar (right) DM mediator.

In fact, the increase in collected data is not the only improvement which can be expected for this search in coming years. The collision energy has been increased to 13.6 TeV for Run 3, and the LHC aims to reach $\sqrt{s}$ = 14 TeV for its high luminosity upgrade.

These relatively small increases in collision energy have the potential to significantly increase the cross section for the highest mediator masses, as these massive final states are produced by partons in the high energy tails of the PDF, which will increase significantly as the total energy increases. Further improvements could be expected as the higher luminosity would allow using a finer binning in the tails of the NN distribution used to identify the signal, which would particularly help identifying heavy mediators which can be differentiated very well from the SM backgrounds. In addition, there may still be room for improvements in the analysis strategy; it has recently been shown that neural networks can achieve better reconstruction of $t\bar{t}$ pairs than traditional algorithms [106], it would be interesting to explore if these methods can also assign some fraction of the p_T^{miss} to DM in the event, which could dramatically improve the performance of this search. Alternatively, it may be possible to train a neural network on the full kinematics of the events (i.e., the 4-momenta of the different particles) to directly differentiate the signal events from background, which with sufficient training data may achieve higher sensitivity than training only on specially constructed variables which are expected to be sensitive to the model. This analysis will therefore continue to explore new parameter space for DM production at the LHC for many years to come.

10 Dark Showers with the Herwig Generator

As discussed in section 3.5, it is possible for dark matter to be part of a "dark sector" with strong interactions between the particles, resulting in "showers" of dark particles which then hadronise, similarly to SM QCD. Some of these dark hadrons can then decay to SM particles, which can lead to a range of signatures which might elude traditional DM searches, such as semi-visible jets with a different substructure to standard QCD jets, or displaced jets, if the dark hadrons have a long lifetime and hence decay some distance from the primary interaction vertex. Until now simulation of such models has only been available in Pythia [107, 108]. However since many of the variables which can help to differentiate these signatures from the SM are also sensitive to the details of the parton shower and hadronisation, it is useful to have multiple different models to simulate this scenario, to ensure the conclusions drawn are not model-dependent, and gain further insight into the implications of the physics scenario. This section therefore discusses the development of dark shower models in the Herwig generator [43, 44], and some first studies of the sensitivity of potential future searches to these models.

10.1 Dark Shower Models

Similarly to the simplified DM models, dark shower models usually involve a high mass mediator which couples the SM to the new dark sector, typically either spin-0 or spin-1 boson, with different couplings motivating searches in different initial states. However, instead of just considering a single DM fermion, the dark sector is considered to have n_f "dark quarks", which are charged under the new strong interaction. This interaction can correspond to various new gauge symmetries, such as the special unitary groups, $SU(N_c)$, the special orthogonal groups, $SO(N_c)$, or the symplectic groups, $Sp(N_c)$, where N_c is the number of colour charges under the new group. Regardless of the exact gauge group, the new force will have a mediator which transforms in the adjoint representation of the group, referred to as a "dark gluon", and a running coupling constant $\alpha_D(s)$, which will increase as the scale decreases, reaching an infinite value at its Landau pole Λ_D. As with QCD in the SM, this then leads to the dark quarks being confined to bound states at low energies. The exact structure of these bound states varies with the gauge group; most groups lead to the formation of dark mesons of one dark quark and one dark anti-quark, however the formation of dark baryons varies with N_c - for $N_c = 2$ these are degenerate with the mesons, while for $N_c > 3$ the production of baryons is suppressed since it becomes progressively harder to form bound states of N_c dark quarks. Therefore this discussion will mostly focus on dark mesons, for which the mass spectrum has been predicted for a wide range of N_c and n_f [109]. This is done using lattice QFT, an alternative formalism to perturbative QFT in which the Lagrangian is solved numerically on a lattice of points, which is able to predict the mass and properties of bound states even in non-perturbative regimes.

10.2 Herwig Implementation

10.2.1 Hard Process

As discussed, the production of the initial dark quarks proceeds via a high energy mediator, and hence one can use the same models as have been developed for simplified DM scenarios to simulate the hard process. Production of the dark fermions via a simple LO s-channel diagram can be generated using the internal ResonantProcessConstructor in Herwig once a suitable BSM model has been specified, while more complicated processes, such as associated production with SM particles or NLO diagrams can be simulated using an external ME provider such as MadGraph5_aMC@NLO using the Herwig Matchbox module [44, 110].

10.2.2 Parton Shower

The showering of the dark quarks is implemented by extending the Herwig parton shower (as described in section 5.3.1) to include the new interaction, which can be turned on using the new "ShowerHandler:DarkInteraction" switch. There are then $3n_f + 1$ possible parton emission diagrams, as shown in figure 56; only final state radiation need be considered, since there are no dark quarks in the initial state.

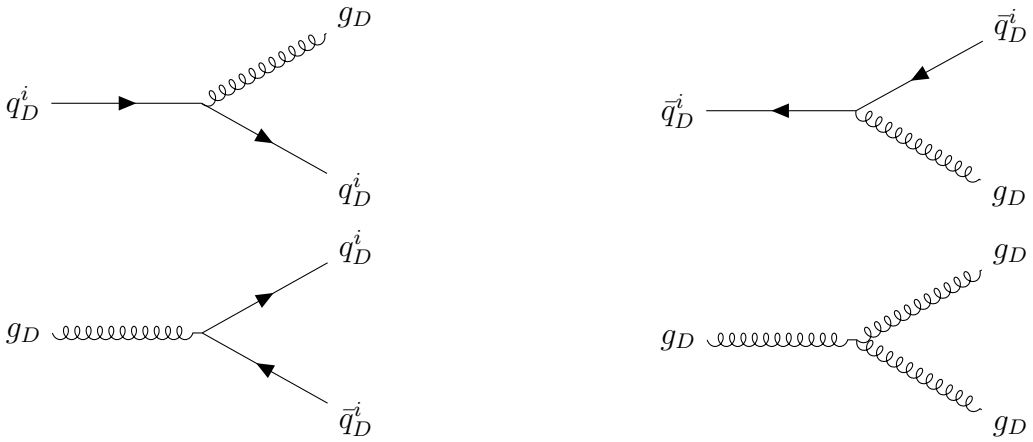


Figure 56: The possible splitting functions in the dark shower: top row - emission of a dark gluon from a dark quark (left) or dark anti-quark (right); bottom row - splitting of a dark gluon into a dark quark - anti-quark pair (left) or two dark gluons (right). The index i runs through the n_f dark gluons.

The rate of these emissions is controlled by the coupling constant $\alpha_D(s)$, which is allowed to evolve with energy following the renormalisation group equations at two loop level, and can be defined either explicitly as $\alpha_D(M_Z)$ or implicitly via Λ_D. The dark shower emissions must be cut-off at some energy scale for which $\alpha_D(s) << 1$ to ensure the parton shower is only used in a region which can reasonably be described perturbatively.

10.2.3 Hadronisation

In order to include hadronisation of dark quarks in the Herwig cluster hadronisation model, described in section 5.3.2, the Herwig code (documented in reference [43]) was refactored to enable the addition of new confining interactions. The main class which performs the hadronisation in Herwig is the ClusterHadronisationHandler, which has a series of sub-handlers which perform the various steps: the GluonMassGenerator sets the gluons to have a sufficiently high mass that they can be split into a pair of quarks, the PartonSplitter performs this splitting and the ClusterFinder forms clusters from the colour connected partons, while the ColourReconnector can reshuffle these clusters according to different models, which can potentially lead to the formation of baryonic clusters (i.e. clusters of either three quarks or three anti-quarks, rather than a quark and an anti-quark). Very heavy clusters are then split into smaller clusters by the ClusterFissioner, then the lightest clusters, which cannot decay to two hadrons, are decayed to a single hadron by the LightClusterDecayer, before the ClusterDecayer decays all remaining clusters to two hadrons. The last two handlers use employ a HadronSelector class which selects which hadrons are produced in the decay from those kinematically available, including baryon - anti-baryon pairs if the cluster has sufficient energy to produce two quarks and two anti-quarks, which are referred to as diquarks.

This set-up was quite specific to the Standard Model, with classes such as the ClusterFissioner having different options for clusters including charm and bottom quarks, while the HadronSelector had parameters relating how often strange quark pairs should be produced in comparison to up and down pairs, as well as specific treatment of the mixing of various SM hadrons. All model specific information was therefore moved into a new class, called a "HadronSpectrum", which expanded the functionality of the HadronSelector, while also allowing separate parameters to be set for an arbitrary number of heavy quarks in the Herwig input settings. A subclass called the "StandardModelHadronSpectrum" implemented the remaining SM QCD specific settings, while another subclass, called the "DarkHadronSpectrum", allows the user to add up to nine hadronising dark quarks and a corresponding spectrum of hadrons that they could form. By default the ClusterHadronisationHandler and its sub-handlers are configured to just use the StandardModelHadronSpectrum, however one can add further HadronSpectra such as the DarkHadronSpectrum to handle hadronisation of other sectors.

Being semi-empirical, the hadronisation model has a number of parameters, which for SM QCD can be tuned to obtain the best possible agreement with data. For the dark sector this is not possible, so these parameters must be set to estimates of reasonable values, based on their physical meaning and the values found to give a good fit for QCD. The most relevant of these parameters will be discussed here, as well as some recommendations for reasonable values.

The first set of relevant parameters are the constituent masses of the dark quarks and gluon, which control the mass these particles can be considered to have in the hadronisation (which can differ significantly from the mass of these particles at higher scales).

For dark quarks this should be at least the mass of the lightest hadron which they can form to ensure their clusters have sufficient energy to decay to into hadronic final states, while for the dark gluons it must be at least twice the constituent mass of the lightest dark quark to allow these to be split by the PartonSplitter.

The next parameters control the splitting of heavy clusters; a cluster will be split if it has a mass of over:

$$\left(Cl_{\rm max}^{Cl_{\rm pow}} + (m_1 + m_2)^{Cl_{\rm pow}}\right)^{1/Cl_{\rm pow}}, \tag{28}$$

where m_1 and m_2 are the constituent masses of the quarks in the cluster, and $Cl_{\rm max}$ and $Cl_{\rm pow}$ are parameters of the model, which are the same for clusters containing only the lightest dark quarks, but can be set separately for clusters containing heavier dark quarks. $Cl_{\rm max}$ represents the highest mass for which a cluster can reasonably be considered a pseudo-hadronic bound state; while the exact value this would take for the spectrum of dark hadrons is unclear, in the SM this is about $17\Lambda_{QCD}$, so a similar multiple of Λ_D would make sense for the dark sector. The best value for $Cl_{\rm pow}$ is unclear; for the SM a value of 2.78 is found to work well for light clusters, but the best fit value for bottom clusters is 0.547. However provided $Cl_{\rm max}$ is significantly larger than the dark quark constituent masses this parameter should not have a major impact.

A dark quark - anti-quark pair is then sprung from the vacuum, with a set of parameters Pwt_i giving the probabilities of selecting a dark quark with flavour i. Pwt_i should be set to 1 for the lightest dark quarks, and 0 for significantly heavier dark quarks that are very unlikely to be pair-produced; it can also take a value in-between for a dark quark with a relatively small mass splitting from the lightest dark quarks, similarly to the strange quark in the SM, though this scenario has not yet been explored. The masses of the two outgoing clusters, M_1 and M_2 are then given by power-like distributions:

$$\begin{aligned} M_1 &= m_1 + \left(r_1 m^P + (1-r_1)(M - m_1 - m_2 - m)^P\right)^{1/P} \\ M_2 &= m_2 + \left(r_2 m^P + (1-r_2)(M - m_1 - m_2 - m)^P\right)^{1/P}, \end{aligned} \tag{29}$$

where M is the mass of the parent cluster, m is the mass of the quark sprung from the vacuum, r_1 and r_2 are independent random variables sampled in the range $[0, 1]$, and P is a free parameter; if $M_1 + M_2 > M$ new values are generated. The best value of P, which will affect how much of the parent cluster masses is converted to mass of the child clusters, compared to their momentum, is unclear, and varies in the SM between 0.625 (for bottom clusters) and 0.994 (for charm clusters), though 0.899, which is the SM value for light clusters, and will lead to most of the energy being converted to mass of the child clusters, seems a good starting point.

The remaining important set of parameters which need to be set control the decay of clusters to two hadrons. Similarly to cluster fissioning, this involves producing a quark - anti-quark pair with flavour determined by parameters Pwt_i. These can be set independently from the values used in the ClusterFissioner, since the energy scale for decay is lower than for fissioning of heavy cluster, so dark quarks heavier than the lightest ones might be slightly more suppressed, however the same general principles apply when selecting the values.

Generally, the hadrons produced in cluster decays are emitted isotropically, however it is found for the SM a better description of the kinematics is achieved if the hadrons which contain a quark from the perturbative parts of the calculation (i.e. not from gluon or heavy cluster splittings) are emitted in broadly the same direction as this quark. There is therefore an option to emit these hadrons in this direction, smeared by an angle θ, where:

$$\cos\theta = 1 + Cl_{\text{Smr}} \log r. \tag{30}$$

r is a random number uniformly distributed between 0 and 1, while Cl_{Smr} is a parameter which controls the degree to which the hadron direction is smeared compared to the parent quark. This option is used in the SM, with Cl_{Smr} varying between 0.78 for light quarks and 0.078 for bottom quarks; in the dark sector it would similarly make sense to use a reasonably large value for dark quarks with masses significantly below Λ_D, where the direction is likely to be smeared by hadronisation effects, and small values for heavier dark quarks which are more likely to retain their direction.

10.2.4 Decays

Decays of unstable dark hadrons can be reasonably easily included in Herwig, which already has a large number of possible hadron decay matrix elements implemented, which can be initialised by the user to also perform decays of dark hadrons. There are two possible cases for these decays: either dark hadrons decay to other dark hadrons, which can use the same matrix elements as the corresponding decays in the SM, or they decay back to SM quarks by mixing with the (off-shell) DM mediator. Since most dark shower models have dark hadrons significantly higher than the QCD confinement scale, these will not decay directly to SM hadrons, but instead to SM quarks, which will undergo further parton showering before hadronising again. This can be achieved using the QuarkoniumDecayer, which decays hadrons to either quarks or gluons, before calling the relevant classes to perform the parton shower and hadronisation. In addition some dark hadrons can decay to a lighter dark hadron and an SM quark anti-quark pair - a new DarkQuarkoniumDecayer class was implemented to handle these decays.

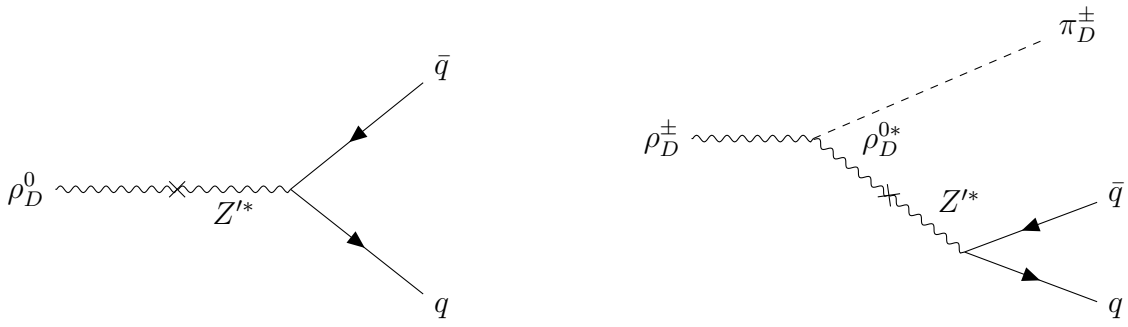


Figure 57: Possible decays of dark ρ mesons: left - decay of a neutral ρ_D^0 meson to a SM $q\bar{q}$ pair, right - decay of a $\rho_D^\pm$ meson, which is charged under a $U(1)$ symmetry associated with the dark mediator, to a lighter dark pion and a SM $q\bar{q}$ pair.

10.3 Validation

In order to validate the implementation in the Herwig code and test reasonable values for the shower and hadronisation parameters, a specific benchmark model was implemented. The plots shown here are custom routines implemented using the Rivet framework [111].

10.3.1 Benchmark Model

Section 4 of [112] outlines a family of benchmarks - these all involve a massive spin-1 "dark photon" mediator, which couples to the SM via mixing with the photon. There are then n_f dark quarks, which are mass degenerate and charged under the $U(1)$ group associated with the dark photon, and under a $SU(N_c)$ group associated with the dark gluon. This leads to a relatively simple hadron spectrum, with $n_f^2 - 1$ mass-degenerate pseudo-scalar dark pions, a significantly heavier flavour singlet pseudo-scalar η'_D, $n_f^2 - 1$ mass-degenerate vector mesons, ρ_D, and the flavour singlet vector ω_D. The mass spectrum of all of these hadrons has been computed as a function of Λ_D and the mass of the dark quarks, m_{q_D}, though from a practical point of view it is easier to parameterise everything in terms of the dark pion mass, m_{π_D}. Fits to lattice calculations of dark hadron masses [113] give:

$$\begin{aligned} m_{q_D} &= 0.033\frac{m_{\pi_D}^2}{\Lambda_D} \\ m_{\eta'_D} &= \sqrt{m_{\pi_D}^2 + 9\frac{n_f}{N_c}\Lambda_D^2} \\ m_{\omega_D} = m_{\rho_D} &= \sqrt{5.76\Lambda_D^2 + 1.5m_{\pi_D}^2}. \end{aligned} \tag{31}$$

If $m_{\rho_D} > 2m_{\pi_D}$, the ρ_D mesons can decay to a pair of dark pions, which in general will be the dominant decay mode since decays to SM particles are suppressed due to the

DM mediator being significantly off-shell. Some of the dark pions can then be unstable and decay back to SM particles, however these decays are further suppressed compared with the decay of ρ_D to visible particles due to helicity considerations, and hence these dark pions are often long-lived and may not decay inside the detector. It was therefore decided to focus on the benchmark with $m_{\pi_D}/\Lambda_D = 1.7$, which has $m_{\rho_D} < 2m_{\pi_D}$, and so the ρ_D mesons decay to SM particles, while the dark pions are assumed to be stable. This benchmark also has $N_c = 3$ and $n_f = 4$, with the different dark quarks having charges of -1, 2, 3 and -4 under the $U(1)$ group, ensuring none of the ρ_D mesons are stable. There are two possible types of decays in this model, as shown in figure 57: the flavour diagonal ρ_D mesons (and the ω_D meson), which are not charged under $U(1)$ can decay directly to a SM $q\bar{q}$ pair, while the other ρ_D mesons, which are charged, decay to a dark pion in addition to the $q\bar{q}$ pair. The benchmark in [112] recommends setting $\Lambda_D = 10$ GeV, however to investigate the effect of varying this scale a second benchmark was considered with $\Lambda_D = 1$ GeV and m_{π_D}/Λ_D kept at the same value.

10.3.2 Parton Shower Validation

The first check was to investigate the effect of the dark parton shower, before considering the hadronisation and decays to the visible sector. To simplify these checks, it was decided to initially focus on production of the dark showers in e^+e^- collisions, since these do not have QCD ISR, and hence the only jets in the event will come from the decays of dark hadrons. The center of mass energy was taken to be $\sqrt{s} = 1$TeV, which was also used as the mass of Z' mediator. Only the production of one flavour of dark quarks was simulated, which aids differentiating the quarks which come from the hard process from the ones produced in the parton shower. The parton shower cut-off was taken to be $3\Lambda_D$; the value of the dark strong coupling at this cut-off is $\alpha_D = 0.59$, which is already quite high for the purely perturbative PS to describe, however it was decided to take this reasonably low scale to allow as many emissions as possible to be described by the perturbative PS rather than the semi-empirical hadronisation model.

The first observable investigated was the multiplicities of dark particles before the parton shower, which are shown in table 5. From this it is clear there are many more dark gluons produced in the PS than dark quarks, which is expected since the initial $q_{D1}\bar{q}_{D1}$ pair can

Particle	Multiplicity ($\Lambda_D = 1$ GeV)	Multiplicity ($\Lambda_D = 10$ GeV)
g_D	4.68	1.30
q_D^1	2.051	2.0054
q_D^2	0.054	0.0056
q_D^3	0.046	0.0058
q_D^4	0.056	0.0058

Table 5: Average multiplicities of dark particles after the parton shower. Multiplicities for quarks also include the corresponding anti-quarks.

only radiate dark gluons, and any additional dark quarks would have to be produced by splitting of these dark gluons, which generally will have lower energy. Similarly, the multiplicities of all particles are higher for $\Lambda_D = 1$ GeV than $\Lambda_D = 10$ GeV, since the lower shower cut-off allows more phase space for emissions, however this is particularly pronounced for the dark quarks, which can only be produced by radiation from other emitted gluons, and hence are quite close to the kinematic boundary.

One can also analytically calculate the expected multiplicity of particles at a particular energy scale, s, [114, 115], which gives:

$$\langle n(s) \rangle \propto \exp\left(\frac{1}{b}\sqrt{\frac{6}{\pi\alpha_D(s)}} + \frac{13\pi b + 5n_f}{54\pi b}\ln\alpha_D(s)\right), \tag{32}$$

where $b = (11N_c - 2n_f)/12\pi$ is the first coefficient of the β function, $\beta(\alpha_D) = Q^2\frac{\partial\alpha_D}{\partial Q^2} = -b\alpha_D^2 + O(\alpha_D^3)$. Inserting the values of $\alpha_D(s)$ calculated at the center of mass energy $s = 1$TeV for the different values of Λ_D suggests the ratio between these multiplicities should be 2.7, which is reasonably close to the ratio for gluons, 3.6, but a long way off the ratio for quarks, which is close to 10. This can be understood as equation 32 only holds if the center-of-mass energy is a long way above the cut-off scale, particularly for the dark quarks which are emitted from dark gluons produced earlier in the shower. Comparing $\Lambda_D = 1$ GeV with $\Lambda_D = 0.1$ GeV, equation 32 predicts a ratio of 2.32, which is very close to the 2.35 observed for dark gluons, and the dark quark ratio, at around 3.2, is also significantly closer.

The distribution of generated momenta of particles after the PS, shown in figure 58, is generally as one would expect; the quarks of the flavour produced in the hard process have a p_T spectrum up to half the mediator mass - for the $\Lambda_D = 10$ GeV benchmark, which has relatively little radiation, the spectrum is peaked towards this value, while

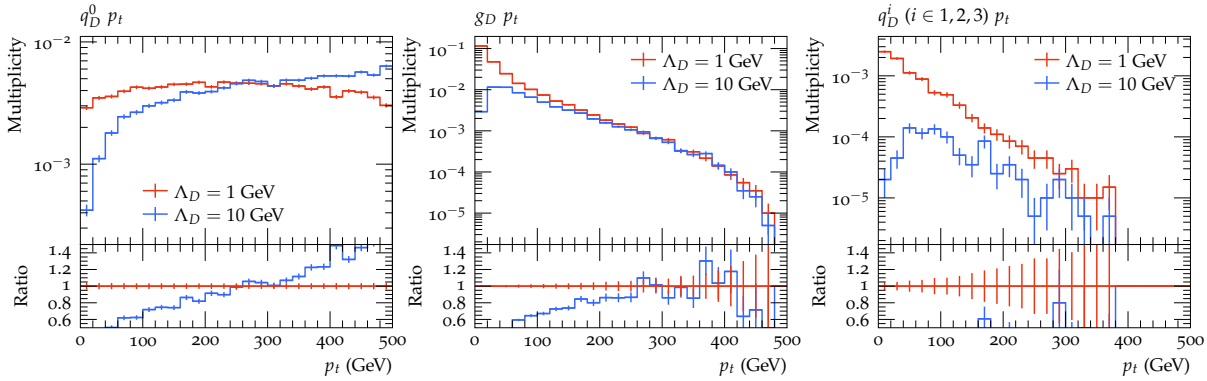


Figure 58: Generated p_T of the dark quarks from the hard process (left), dark gluons (center) and dark quarks produced in the parton shower (right) for the benchmarks with $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV.

for $\Lambda_D = 1$ GeV the spectrum is more flat as more energy is lost to PS emissions. The momentum spectra of the other particles is peaked towards low values due to soft singularities, though the $\Lambda_D = 10$ GeV benchmark has less radiation at low p_T due to the shower cut-off.

10.3.3 Hadronisation Parameters

As discussed in section 10.2.3, the cluster hadronisation model has a number of parameters which need to be set to physically sensible values. For these preliminary checks there was not time to explore the impact of varying all of these scales, so most parameters were set to the same values as for light quarks in QCD, with the exception of Cl_{max}, which is set to $17\Lambda_D$, and the constituent masses, which are set to be equal to the dark pion mass for dark quarks, and twice this for dark gluons. To check the cluster fissioning works well with these parameters, the masses of these clusters were plotted before and after cluster fissioning (figure 59). One can see that before that before cluster fissioning the cluster masses are spread over a large range from $3.8\Lambda_D$, twice the dark quark constituent mass, to 1000 GeV, which corresponds to the case there are no parton shower emissions, so the two dark quarks from the hard process form a single cluster. The cluster masses are peaked around $6\Lambda_D$ due to the shower cut-off, then decrease, however there are still a large number of clusters with masses too high to be described in the cluster hadronisation model. After fissioning the distributions look well-behaved, with a similar shape to before fissioning, but the heaviest clusters split into more light clusters in the bulk.

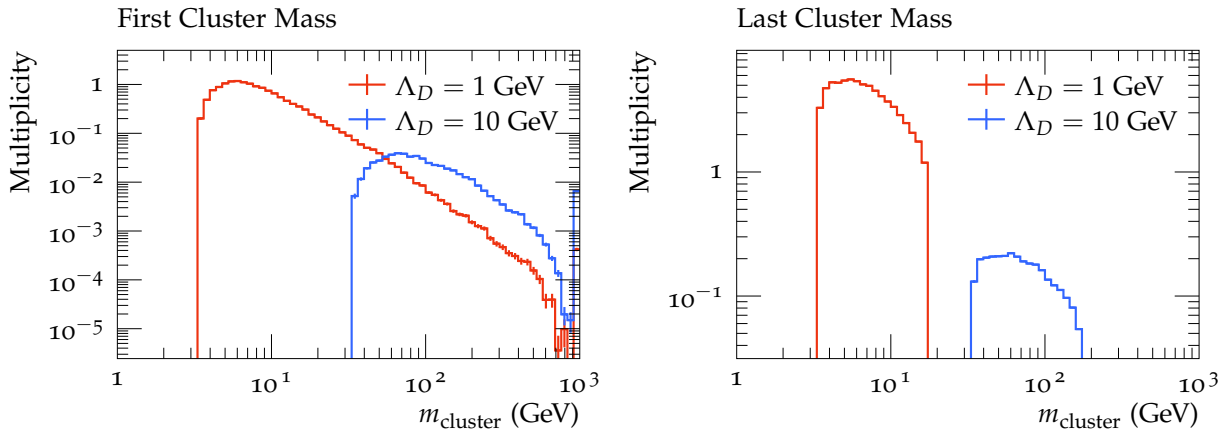


Figure 59: Cluster masses in the cluster hadronisation model before cluster fissioning (left) and after (right), for the $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV benchmarks.

These clusters were then decayed to dark hadrons- the average multiplicities of the different types of mesons produced are shown in table 6. One of the advantages of the cluster hadronisation model is that it is predictive of the flavours of produced mesons (albeit with dependency on the cluster hadronisation parameters), in contrast to the

string hadronisation model used in Pythia, where the probability of producing a meson must be supplied as a model parameter. From the average multiplicities the probability to produce a vector meson (ρ_D or ω_D) is 0.546 for the $\Lambda_D = 1$ GeV benchmark and 0.555 for $\Lambda_D = 10$ GeV, which is slightly lower than the value of 0.58 proposed in [112], though more studies of the impact of cluster parameters on these values are needed before this could be considered a recommendation for future model building. η_D mesons, which are completely neglected in [112], form about 0.9% of all mesons produced my cluster hadronisation for both benchmarks, though for this scenario these are not hugely relevant, since $m_{\eta_D} > 2m_{\pi_D}$, so these will decay preferentially to stable dark pions rather than SM particles.

Particle	Multiplicity ($\Lambda_D = 1$ GeV)	Multiplicity ($\Lambda_D = 10$ GeV)
π_D	8.81	3.59
η_D	0.18	0.07
$\rho_D^0 + \omega_D$	2.72	1.15
$\rho_D^\pm$	8.12	3.43

Table 6: Average multiplicities of dark hadrons before decays. The vector mesons are split into those that decay to a SM $q\bar{q}$ pair ($\rho_D^0 + \omega_D$), and those that decay to an SM $q\bar{q}$ plus a dark pion ($\rho_D^\pm$), as discussed in section 10.2.4

The transverse momenta of dark pions and ρ_D mesons before hadron decays are shown in figure 60. These are as one would expect: for the $\Lambda = 1$ GeV benchmark the p_T spectrum is peaked at very low values, though decreases relatively slowly, with some mesons having $p_T > 200$ GeV; the $\Lambda = 10$ GeV benchmark has slightly more mesons

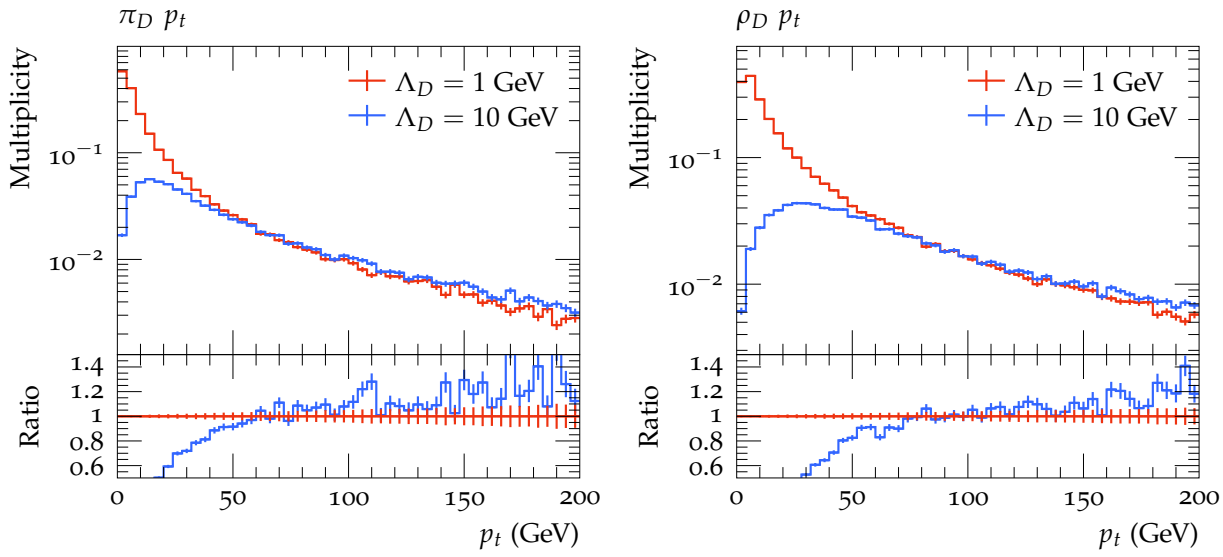


Figure 60: p_T of the dark pions (left) and ρ_D mesons (right), for the $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV benchmarks.

at high p_T, but the most notable difference is that are significantly fewer mesons at low p_T due to the higher shower cut-off and cluster hadronisation scales suppressing soft emissions. In both benchmarks, the ρ_D mesons have a slightly harder p_T spectrum than the dark pions, which can be understood as heavier clusters will decay preferentially to ρ_D mesons due to the larger number of available spin states, whereas lighter clusters which do not have the energy to produce ρ_D mesons will decay to dark pions.

10.4 Sensitive Variable Studies

After the ρ_D mesons are allowed to decay to the SM quarks, one can compare observables produced from the visible SM final state particles to SM backgrounds to identify which variables could be useful for future searches for this model. As has been discussed, traditional DM search variables such as p_T^{miss} are less sensitive to dark showers since some of the dark hadrons decay back to the Standard Model rather than contributing to p_T^{miss}. However variables which are sensitive to the substructure of the jets may aid discrimination of dark showers, since the jets produced in dark showers will tend to contain a number of separated energy deposits coming from the decays of the different dark hadrons.

10.4.1 Missing Transverse Momentum

The first variable investigated was the p_T^{miss}, since this can still provide some discriminating power for dark shower models. The p_T^{miss} distribution is shown in figure 61; both

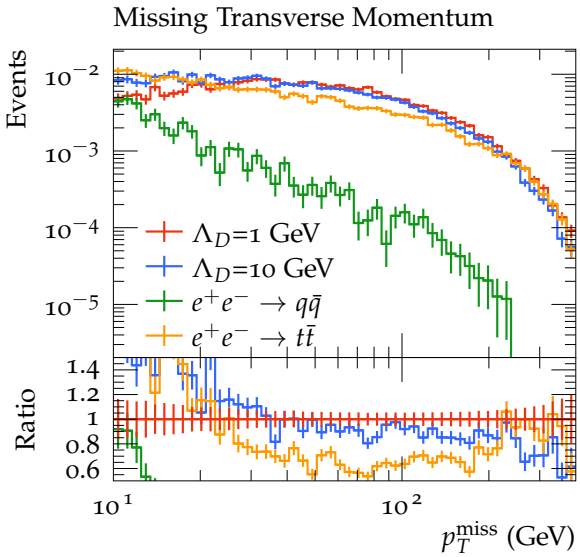


Figure 61: $p_T^{\text{miss}}$of the $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV benchmarks dark shower benchmarks, as well as SM $e^+e^- \to q\bar{q}$ and $e^+e^- \to t\bar{t}$ scatterings. All processes are normalised to unity.

benchmarks have significantly more p_T^{miss} than $e^+e^- \to q\bar{q}$ ($q \in [u, d, s, c, b]$) scatterings, which do not contain any particles which contribute to p_T^{miss} other than neutrinos in hadron decays, however the p_T^{miss} spectrum is quite similar to $e^+e^- \to t\bar{t}$, which has a moderately hard p_T^{miss} spectrum due to the neutrinos in W boson decays. At the LHC there would be other significant backgrounds such a $Z \to \nu\bar{\nu}$ + jets production, and other processes might have spurious p_T^{miss} due to mismeasurement, which is not considered here. p_T^{miss} may therefore be a useful variable for suppressing some backgrounds, however it is not sufficient as the main search variable.

10.4.2 Angularities

One commonly used set of variables sensitive to the substructure of jets are angularities [116, 117]. There are number of slightly different definitions for these variables, but the ones used in this work are:

$$\tau_{\alpha,\beta} = \sum_{\text{jet}\in\text{jets}} \sum_{i\in\text{jet}} \left(2\sqrt{1 - cos\theta_i}\right)^{\alpha} \left(\frac{E_i}{E_{Tot}}\right)^{\beta} \approx \sum_{\text{jet}\in\text{jets}} \sum_{i\in\text{jet}} \theta_i^{\alpha} \left(\frac{E_i}{E_{Tot}}\right)^{\beta}, \qquad (33)$$

where θ_i is the angle of the i^{th} particle in a jet to the jet axis, E_i is the energy of this particle, E_{Tot} is the total energy in the event, and the final approximation holds for $\theta_i << 1$. These variables probe the structure of the jet at different angular and energy scales depending on the values of α and β - small values of α probe particles close to the jet axis and larger values probe wide-angle emissions, and similarly small values of β probe soft emissions, while higher values probe the harder particles in the jet. Combinations of α and β in the range $[0.1, 2]$ were tested, and two sensitive combinations are shown in figure 62. The angularity with $\alpha = 0.3, \beta = 2$, which is sensitive to high energy particles relatively close to the center of the jet, seems particularly sensitive to the $\Lambda_D = 1$ GeV benchmark, likely because this angular scale corresponds to either the distance between the $q\bar{q}$ pair from a ρ_D decay, or between a pair of ρ_D mesons in the decay of a dark cluster, either of which could give rise to two high energy clusters within the jet with reasonably small angular separation. For the $\Lambda_D = 10$ GeV benchmark, by contrast, the most sensitive angularity comes at $\alpha = 2, \beta = 1$, i.e. wide angle emissions which are neither particularly hard nor particularly soft. This likely reflects the higher mass of the ρ_D mesons and dark clusters, which will lead to much more separated emissions than for the $\Lambda_D = 1$ GeV benchmark; the fact that the emissions are neither particularly hard nor particularly soft probably reflects the fact that the SM quarks produced from the decays of these heavy ρ_D mesons will have sufficient energy to undergo a reasonably long parton shower, and hence will have particles covering a p_T spectrum similar to QCD, whereas in the $\Lambda_D = 1$ GeV benchmark the SM quarks are produced quite close to the QCD confinement scale, so will probably not produce a full shower of softer particles.

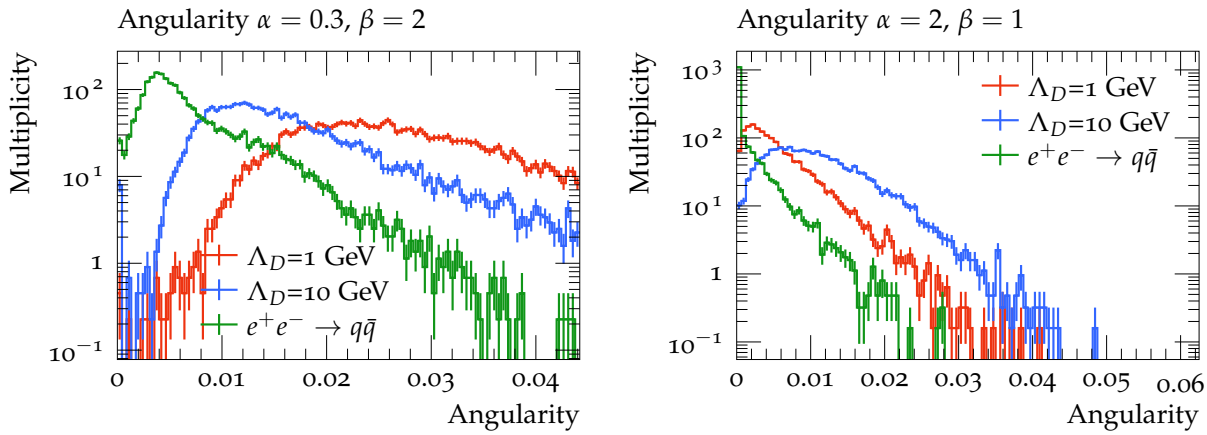


Figure 62: Angularities of visible jets with $\alpha = 0.3, \beta = 2$ (left) and $\alpha = 2, \beta = 1$ (right) for the $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV dark shower benchmarks and $e^+e^- \to q\bar{q}$ scattering. All processes are normalised to unity.

10.4.3 Angular Correlation Functions

Angular correlation functions are a different type of jet substructure observables. In contrast to the angularities, which sum over all particles within a jet, these correlation function are formed by considering the angles between all visible particles within a jet, weighted by the energy of the particles to a particular power. These are shown for the energy powers of 0, 1 and 2 in figure 63. The correlation function with energy raised to the 0^{th} power (i.e., no energy weighting) shows good discrimination for the $\Lambda_D = 10$ GeV benchmark, which has much more particles at large angles from one another, which is likely related to the observation that this benchmark has a significant excess for the $\alpha = 2, \beta = 1$ angularity. Interestingly this correlation function offers almost no discrimination for the $\Lambda_D = 1$ GeV benchmark, which instead is best probed by the correlation function weighted to the second power of energy, for which this benchmark has many more high energy particles at small angles, as was observed by the $\alpha = 0.3, \beta = 2$ angularity. For

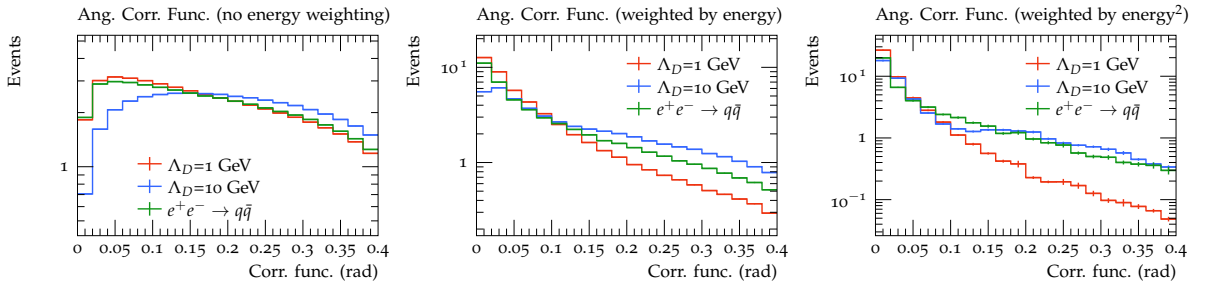


Figure 63: Angular correlation functions within all jets, without energy weighting (left) and weighted by the energy (center) and energy squared (right) for the $\Lambda_D = 1$ GeV and $\Lambda_D = 10$ GeV dark shower benchmarks and $e^+e^- \to q\bar{q}$ scattering. All processes are normalised to unity.

the correlation function weighted to one power of the energy, both behaviours are evident - the $\Lambda_D = 1$ GeV benchmark tends towards smaller angles than the QCD background, $\Lambda_D = 10$ GeV tends to large angles.

10.5 Outlook

The jet substructure variables outlined here provide a useful probe of dark showers - a notable observation from these studies is that the most sensitive variables can change significantly depending on the benchmark scenario being investigated, so more studies are required on further benchmarks to understand the full range of signatures which can be produced by these models. More studies would also be required in order to use these variables at the LHC; firstly, the variables could not be constructed with the particles inside the jet, which are not identified, but must be instead reconstructed from the tracks of charged particles in the tracker and calorimeter deposits, which would lead to worse resolution. Secondly, there are more background processes which must be accounted for at the LHC; there is a significantly higher rate of jet production at a hadron collider compared to the e^+e^- collider backgrounds considered here, which leads to both more background events and means that signal events can contain a mixture of dark sector and QCD jets due to ISR.

A first search for dark showers producing these semi-visible jet signatures at the LHC has already been performed by the CMS experiment [118], which targets signatures with p_T^{miss} aligned to a jet, as well as using a BDT trained on some basic substructure variables to identify semi-visible jets. However a large region of possible models remains to be explored; it is hoped that providing the capability to simulate these models in Herwig will bolster these efforts, as it offers better predictivity for some model parameters such as the fraction of vector mesons produced for a given benchmark, and can help quantify the modelling uncertainties through comparisons with Pythia, or variations of the hadronsiation parameters. In addition usage of multiple generators for the background predictions may also be relevant to these searches, as jet substructure variables can also be sensitive to differences between the parton shower and hadronisation models, so a dark shower search programme could also promote improvements in modeling of SM QCD in these generators.

11 Conclusion

Searching for dark matter produced in association with Standard Model particles is one of the main objectives of the current LHC physics programme. As part of this effort, this work presents a search for dark matter produced in association with top quarks decaying into two lepton final states with the CMS experiment using the LHC Run 2 dataset. This search contains numerous improvements to enhance sensitivity; firstly the selection was expanded to target the t + DM signal in addition to the $t\bar{t}$ + DM signal for the first time in the dileptonic channel, which adds sensitivity particular for higher mediator mass hypotheses. A new kinematic reconstruction algorithm was explored to improve the reconstruction of the $t\bar{t}$ pair for the $t\bar{t}$ + DM signal, which is challenging since most algorithms assume all the p_T^{miss} is due to neutrinos, which is not true for this signal. Neural networks were trained on a number of sensitive variables to differentiate the signal from the backgrounds, and new control regions were added to estimate the rate of the $t\bar{t}Z$ and Drell-Yan backgrounds. During checks of the systematic uncertainties, it was found some had spurious shapes due to limited MC simulation statistics. A new "smoothing" procedure was therefore used on the affected systematic templates to obtain a better estimate for the systematic uncertainties considered.

As a result of these improvements, the expected sensitivity for the 2016 dataset improves by approximately 40% for low mediator masses and 65% for high mediator masses with respect to the results published in [55]. When also including the 2017 and 2018 data the improvement becomes 65% and 80% respectively.

A slight excess was observed, particularly in the channels targeting the t + DM signal. This excess is consistent with all of the signal models considered, but most consistent with a pseudoscalar mediator with a mass of 300-400 GeV, for which a significance of 1.54σ is observed. Limits are also set, excluding scalar mediator masses up to 150 GeV (230 GeV expected) and pseudoscalar mediator masses up to 150 GeV (260 GeV expected) for the benchmark couplings $g_q = g_\chi = 1$. Comparison with other results suggests the observed excess is unlikely to be due to the top + DM model, though further investigation is required to determine if the excess is genuine and due to a different type of signal process, or a statistical fluctuation.

Finally, a "dark shower" model, where dark matter is part of a "dark sector" with strong self-interactions, was implemented into the Herwig MC generator. Jet substructure variables were then investigated, which showed potential discrimination power for future LHC searches for this model.

While no clear evidence has yet been observed for the production of dark matter at the LHC, significant discovery potential remains, both in terms of the potential expanded reach of traditional searches for p_T^{miss} in association with SM particles from the ongoing Run 3 data-taking and HL-LHC datasets, and from new analyses targeting alternative models such as the dark showers presented here. The coming years therefore hold the potential for exciting new discoveries in this key area of modern physics.

References

1. Green, A. M. Dark matter in astrophysics/cosmology. *SciPost Phys. Lect. Notes,* 37. https://scipost.org/10.21468/SciPostPhysLectNotes.37 (2022) (cit. on pp. 10, 20).
2. Planck 2018 results - VI. Cosmological parameters. *A&A* **641,** A6. https://doi.org/10.1051/0004-6361/201833910 (2020) (cit. on pp. 10, 24).
3. Feng, J. L. The WIMP paradigm: Theme and variations. *SciPost Phys. Lect. Notes,* 71. https://scipost.org/10.21468/SciPostPhysLectNotes.71 (2023) (cit. on pp. 10, 25).
4. https://en.wikipedia.org/wiki/Standard_Model. Accessed: 12.06.2023 (cit. on p. 12).
5. Woithe, J., Wiener, G. J. & der Veken, F. F. V. Let's have a coffee with the Standard Model of particle physics! *Physics Education* **52,** 034001. https://dx.doi.org/10.1088/1361-6552/aa5b25 (Mar. 2017) (cit. on p. 13).
6. The Particle Data Group collaboration. Review of Particle Physics. *PTEP* **2022,** 083C01 (2022) (cit. on p. 15).
7. The CMS collaboration. Measurement of the top quark mass using proton-proton data at $\sqrt{(s)} = 7$ and 8 TeV. *Phys. Rev. D* **93,** 072004. https://link.aps.org/doi/10.1103/PhysRevD.93.072004 (7 Apr. 2016) (cit. on p. 18).
8. Bernreuther, W., Brandenburg, A., Si, Z. G. & Uwer, P. Spin properties of top quark pairs produced at hadron colliders. *Acta Phys. Polon. B* **34** (eds Fialkowski, K., Jezabek, M. & Rozanska, M.) 4477–4490. arXiv: hep-ph/0304244 (2003) (cit. on p. 18).
9. Rubin, V. C. & Ford W. Kent, J. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal* **159,** 379 (Feb. 1970) (cit. on p. 20).
10. Freeman, K. C. On the Disks of Spiral and S0 Galaxies. *The Astrophysical Journal* **160,** 811 (June 1970) (cit. on p. 20).
11. Karukes, E. V., Salucci, P. & Gentile, G. The dark matter distribution in the spiral NGC 3198 out to 0.22 Rvir. *A&A* **578,** A13. https://doi.org/10.1051/0004-6361/201425339 (2015) (cit. on p. 21).
12. Zwicky, F. On the Masses of Nebulae and of Clusters of Nebulae. *The Astrophysical Journal* **86,** 217 (Oct. 1937) (cit. on p. 21).
13. Ryden, B. *Introduction to Cosmology* 2nd ed. (Cambridge University Press, 2016) (cit. on p. 22).
14. ESA/Hubble & NASA. *Astronomy Picture of the Day, 2011 December 21* https://apod.nasa.gov/apod/ap111221.html. Accessed: 12.06.2023 (cit. on p. 22).

15. Clowe, D. *et al.* A Direct Empirical Proof of the Existence of Dark Matter*. *The Astrophysical Journal* **648,** L109. https://dx.doi.org/10.1086/508162 (Aug. 2006) (cit. on p. 23).

16. Chatterjee, C., Majumdar, D. & Chatterjee, S. *Dark Matter Self Interactions and its Impact on Large Scale Structures* Jan. 2019 (cit. on p. 26).

17. Et al., A. B. Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter. *Physics of the Dark Universe* **27,** 100365. https://www.sciencedirect.com/science/article/pii/S2212686419301633 (2020) (cit. on p. 27).

18. Branco, G. *et al.* Theory and phenomenology of two-Higgs-doublet models. *Physics Reports* **516.** Theory and phenomenology of two-Higgs-doublet models, 1–102. https://www.sciencedirect.com/science/article/pii/S0370157312000695 (2012) (cit. on p. 27).

19. Brüning, O. S. *et al. LHC Design Report* http://cds.cern.ch/record/782076 (CERN, Geneva, 2004) (cit. on p. 31).

20. Lopienska, E. The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022. General Photo. https://cds.cern.ch/record/2800984 (2022) (cit. on p. 32).

21. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B* **716,** 30–61. https://www.sciencedirect.com/science/article/pii/S0370269312008581 (2012) (cit. on p. 31).

22. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B* **716,** 1–29. https://www.sciencedirect.com/science/article/pii/S037026931200857X (2012) (cit. on p. 31).

23. *CMS Collaboration* https://cms.cern/collaboration (cit. on p. 33).

24. Mousa, J. *et al.* A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider. *Science* **338,** 1569–1575 (Dec. 2012) (cit. on p. 34).

25. The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation* **9,** P10009. https://dx.doi.org/10.1088/1748-0221/9/10/P10009 (Oct. 2014) (cit. on p. 35).

26. The CMS collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software (2006) (cit. on pp. 36–38).

27. The CMS collaboration. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST* **10,** P06005. arXiv: 1502.02701 [physics.ins-det] (2015) (cit. on p. 40).

28. The CMS collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST* **13,** P06015. arXiv: 1804.04528 [physics.ins-det] (2018) (cit. on pp. 40, 57).

29. Cacciari, M., Salam, G. P. & Soyez, G. The anti-kt jet clustering algorithm. *Journal of High Energy Physics* **2008,** 063. https://dx.doi.org/10.1088/1126-6708/2008/04/063 (Apr. 2008) (cit. on p. 40).

30. The CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *Journal of Instrumentation* **12,** P10003. https://dx.doi.org/10.1088/1748-0221/12/10/P10003 (Oct. 2017) (cit. on p. 40).

31. The CMS collaboration. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST* **12,** P02014. arXiv: 1607.03663 [hep-ex] (2017) (cit. on p. 40).

32. The CMS collaboration. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector. *JINST* **14,** P07004. arXiv: 1903.06078 [hep-ex] (2019) (cit. on p. 40).

33. The NNPDF collaboration. Parton distributions from high-precision collider data. *Eur. Phys. J. C* **77,** 663. arXiv: 1706.00428 [hep-ph] (2017) (cit. on pp. 42, 52).

34. The NNPDF collaboration. Parton distributions for the LHC Run II. *JHEP* **04,** 040. arXiv: 1410.8849 [hep-ph] (2015) (cit. on pp. 42, 52).

35. Alwall, J. *et al.* The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP* **07,** 079. arXiv: 1405.0301 [hep-ph] (2014) (cit. on pp. 43, 52).

36. Frixione, S. & Webber, B. R. Matching NLO QCD computations and parton shower simulations. *JHEP* **06,** 029. arXiv: hep-ph/0204244 (2002) (cit. on p. 43).

37. Nason, P. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP* **11,** 040. arXiv: hep-ph/0409146 (2004) (cit. on p. 43).

38. Frixione, S., Nason, P. & Oleari, C. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP* **11,** 070. arXiv: 0709.2092 [hep-ph] (2007) (cit. on p. 43).

39. Alioli, S., Nason, P., Oleari, C. & Re, E. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP* **06,** 043. arXiv: 1002.2581 [hep-ph] (2010) (cit. on p. 43).

40. Ježo, T. & Nason, P. On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower. *JHEP* **12,** 065. arXiv: 1509.09071 [hep-ph] (2015) (cit. on p. 43).

41. Alwall, J. *et al.* Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C* **53,** 473–500. arXiv: 0706.2569 [hep-ph] (2008) (cit. on p. 44).

42. Bierlich, C. *et al.* A comprehensive guide to the physics and usage of PYTHIA 8.3. arXiv: 2203.11601 [hep-ph] (Mar. 2022) (cit. on pp. 44, 52).

43. Bahr, M. *et al.* Herwig++ Physics and Manual. *Eur. Phys. J. C* **58,** 639–707. arXiv: 0803.0883 [hep-ph] (2008) (cit. on pp. 44, 106, 108).

44. Bellm, J. *et al.* Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C* **76,** 196. arXiv: 1512.01178 [hep-ph] (2016) (cit. on pp. 44, 106, 107).

45. Altarelli, G. & Parisi, G. Asymptotic freedom in parton language. *Nuclear Physics B* **126,** 298–318. https://www.sciencedirect.com/science/article/pii/0550321377903844 (1977) (cit. on p. 45).

46. Dokshitzer, Y. L. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP* **46,** 641–653 (1977) (cit. on p. 45).

47. Gribov, V. N. & Lipatov, L. N. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.* **15,** 438–450 (1972) (cit. on p. 45).

48. Andersson, B., Gustafson, G. & Soderberg, B. A General Model for Jet Fragmentation. *Z. Phys. C* **20,** 317 (1983) (cit. on p. 46).

49. Sjöstrand, T. Jet fragmentation of multiparton configurations in a string framework. *Nuclear Physics B* **248,** 469–502. https://www.sciencedirect.com/science/article/pii/0550321384906072 (1984) (cit. on p. 46).

50. Webber, B. R. A QCD Model for Jet Fragmentation Including Soft Gluon Interference. *Nucl. Phys. B* **238,** 492–528 (1984) (cit. on p. 46).

51. *CMS PU twiki* https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData. CMS internal, accessed: 09.08.2023 (cit. on p. 48).

52. Agostinelli, S. *et al.* Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506,** 250–303. https://www.sciencedirect.com/science/article/pii/S0168900203013688 (2003) (cit. on pp. 48, 52).

53. Allison, J. *et al.* Geant4 developments and applications. *IEEE Transactions on Nuclear Science* **53,** 270–278 (2006) (cit. on pp. 48, 52).

54. Allison, J. *et al.* Recent developments in Geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835,** 186–225. https://www.sciencedirect.com/science/article/pii/S0168900216306957 (2016) (cit. on pp. 48, 52).

55. The CMS collaboration. Search for Dark Matter Particles Produced in Association with a Top Quark Pair at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.* **122,** 011803. https://link.aps.org/doi/10.1103/PhysRevLett.122.011803 (1 Jan. 2019) (cit. on pp. 49, 92, 93, 120).

56. The CMS collaboration. Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP* **03,** 141. arXiv: 1901.01553. https://cds.cern.ch/record/2652856 (2019) (cit. on p. 49).

57. The CMS collaboration. *Search for dark matter in association with a* $\mathrm{t\bar{t}}$ *pair at* $\sqrt{s} = 13$ TeV *in the dilepton channel with 2016 data* tech. rep. (CERN, Geneva, 2018). http://cds.cern.ch/record/2308269 (cit. on pp. 49, 66).

58. The CMS collaboration. Combined searches for the production of supersymmetric top quark partners in proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$. *Eur. Phys. J. C* **81,** 970. arXiv: 2107.10892 [hep-ex] (2021) (cit. on pp. 49, 102, 103).

59. The ATLAS collaboration. Constraints on spin-0 dark matter mediators and invisible Higgs decays using ATLAS 13 TeV *pp* collision data with two top quarks and missing transverse momentum in the final state. *Eur. Phys. J. C* **83,** 503. arXiv: 2211.05426 [hep-ex] (2023) (cit. on pp. 49, 102, 103).

60. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362. https://doi.org/10.1038/s41586-020-2649-2 (Sept. 2020) (cit. on p. 50).

61. Pivarski, J. *et al. Awkward Array* version v2.2.0. May 2023. https://doi.org/10.5281/zenodo.7919197 (cit. on p. 51).

62. Gray, L. *et al. CoffeaTeam/coffea: Release v2023.5.0.rc0* version v2023.5.0.rc0. May 2023. https://doi.org/10.5281/zenodo.7926179 (cit. on p. 51).

63. Rübenach, J. & Stafford, D. *Pepper: Particle Physics Processor* May 2023. https://gitlab.cern.ch/pepper/pepper (cit. on p. 51).

64. The CMS collaboration. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C* **81,** 800. arXiv: 2104.01927 [hep-ex] (2021) (cit. on pp. 51, 86).

65. The CMS collaboration. *CMS luminosity measurement for the 2017 data-taking period at* $\sqrt{s} = 13$TeV CMS Physics Analysis Summary CMS-PAS-LUM-17-004 (2018). https://cds.cern.ch/record/2621960 (cit. on pp. 51, 86).

66. The CMS collaboration. *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s} = 13$TeV CMS Physics Analysis Summary CMS-PAS-LUM-18-002 (2019). https://cds.cern.ch/record/2676164 (cit. on pp. 51, 86).

67. Frixione, S., Nason, P. & Ridolfi, G. A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP* **09,** 126. arXiv: 0707.3088 [hep-ph] (2007) (cit. on p. 51).

68. Melia, T., Nason, P., Rontsch, R. & Zanderighi, G. W+W-, WZ and ZZ production in the POWHEG BOX. *JHEP* **11,** 078. arXiv: 1107.5051 [hep-ph] (2011) (cit. on p. 52).

69. Nason, P. & Zanderighi, G. W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2. *Eur. Phys. J. C* **74,** 2702. arXiv: 1311.1365 [hep-ph] (2014) (cit. on p. 52).

70. Hartanto, H. B., Jager, B., Reina, L. & Wackeroth, D. Higgs boson production in association with top quarks in the POWHEG BOX. *Phys. Rev. D* **91,** 094003. arXiv: 1501.04498 [hep-ph] (2015) (cit. on p. 52).

71. The CMS collaboration. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C* **76,** 155. arXiv: 1512.00815 [hep-ex] (2016) (cit. on p. 52).

72. The CMS collaboration. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. *Eur. Phys. J. C* **80,** 4. arXiv: 1903.12179 [hep-ex] (2020) (cit. on p. 52).

73. *CMS EGamma HLT summary* https://twiki.cern.ch/twiki/bin/view/CMS/EgHLTRunIISummary. CMS internal, accessed: 22.08.2023 (cit. on pp. 52, 54).

74. The CMS Collaboration. Measurement of the trigger efficiencies for a dilepton selection for a $t\bar{t}$ analysis with the full run 2 dataset. *CMS Note AN-2019/140* **2019/140** (2019) (cit. on p. 54).

75. *CMS EGamma POG Run 2 recommendations* https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations. CMS internal, accessed: 13.08.2023 (cit. on p. 57).

76. *CMS Muon POG recommendations* https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG#User_Recommendations. CMS internal, accessed: 13.08.2023 (cit. on p. 57).

77. *CMS Jet Energy Correction Introduction* https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC. CMS internal, accessed: 13.08.2023 (cit. on p. 57).

78. *CMS Jet ID* https://twiki.cern.ch/twiki/bin/view/CMS/JetID. CMS internal, accessed: 13.08.2023 (cit. on p. 58).

79. *CMS Jet ID* https://twiki.cern.ch/twiki/bin/view/CMS/PileupJetID. CMS internal, accessed: 13.08.2023 (cit. on p. 58).

80. The CMS Collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation* **13,** P05011. https://dx.doi.org/10.1088/1748-0221/13/05/P05011 (May 2018) (cit. on p. 58).

81. *CMS b tag SF Methods* https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods#1a_Event_reweighting_using_scale. CMS internal, accessed: 13.08.2023 (cit. on p. 59).

82. *CMS 2017 MET recipe v2* https://hypernews.cern.ch/HyperNews/CMS/get/JetMET/1865.html. CMS internal, accessed: 13.08.2023 (cit. on p. 59).

83. Lester, C. G. & Summers, D. J. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders. *Phys. Lett. B* **463,** 99–103. arXiv: hep-ph/9906349 (1999) (cit. on p. 60).

84. The Particle Data Group collaboration. Review of Particle Physics. *Phys. Rev. D* **98,** 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001 (3 Aug. 2018) (cit. on p. 61).

85. The CMS collaboration. Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV. *Phys. Rev. D* **95,** 092001. arXiv: 1610.04191 [hep-ex] (2017) (cit. on p. 61).

86. Czakon, M. *et al.* Top-pair production at the LHC through NNLO QCD and NLO EW. *JHEP* **10,** 186. arXiv: 1705.04105 [hep-ph] (2017) (cit. on p. 61).

87. Betchart, B. A., Demina, R. & Harel, A. Analytic solutions for neutrino momenta in decay of top quarks. *Nucl. Instrum. Meth. A* **736,** 169–178. arXiv: 1305.1878 [hep-ph] (2014) (cit. on p. 64).

88. Sonnenschein, L. Analytical solution of ttbar dilepton equations. *Phys. Rev. D* **73.** [Erratum: Phys.Rev.D 78, 079902 (2008)], 054015. arXiv: hep-ph/0603011 (2006) (cit. on p. 64).

89. Anuar, A. A. *Top Quark Spin and Polarization Properties in Searches for New Phenomena with the CMS Detector at the LHC* Dissertation (Universität Hamburg, Hamburg, 2019). https://bib-pubdb1.desy.de/record/434642 (cit. on p. 67).

90. Chollet, F. *et al. Keras* https://keras.io. 2015 (cit. on p. 69).

91. Kingma, D. P. & Ba, J. *Adam: A Method for Stochastic Optimization* 2017. arXiv: 1412.6980 [cs.LG] (cit. on p. 69).

92. The CMS Collaboration. Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s}$= 13 TeV. *Journal of High Energy Physics* **56.** arXiv: 1907.11270 [hep-ph] (2020) (cit. on pp. 75, 98).

93. Butterworth, J. *et al.* PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics* **43,** 023001. https://dx.doi.org/10.1088/0954-3899/43/2/023001 (Jan. 2016) (cit. on p. 85).

94. Carrazza, S., Forte, S., Kassabov, Z., Latorre, J. I. & Rojo, J. An Unbiased Hessian Representation for Monte Carlo PDFs. *Eur. Phys. J. C* **75,** 369. arXiv: 1505.06736 [hep-ph] (2015) (cit. on p. 85).

95. Barlow, R. & Beeston, C. Fitting using finite Monte Carlo samples. *Computer Physics Communications* **77,** 219–228. https://www.sciencedirect.com/science/article/pii/001046559390005W (1993) (cit. on pp. 86, 90).

96. The CMS collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST* **13,** P05011. arXiv: 1712.07158 [physics.ins-det] (2018) (cit. on p. 87).

97. *CMS Jet Energy Correction Uncertainty Sources* https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources. CMS internal, accessed: 13.08.2023 (cit. on p. 87).

98. Popov, A. *et al.* Artificial constraints from statistical fluctuations in systematic variations. *CMS Note AN-2018/417* **2018/077** (2018) (cit. on p. 88).

99. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C* **71.** [Erratum: doi:10.1140/epjc/s10052-013-2501-z], 1554. arXiv: 1007.1727 [physics.data-an] (2011) (cit. on p. 91).

100. Read, A. L. Presentation of search results: The CL_s technique. *J. Phys. G* **28,** 2693 (2002) (cit. on p. 91).

101. Junk, T. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A* **434,** 435. arXiv: hep-ex/9902006 [hep-ex] (1999) (cit. on p. 91).

102. *Combine: a tool for statistical fits* 2023. http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/ (cit. on p. 92).

103. Moneta, L. *et al. The RooStats Project* in *13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010)* PoS(ACAT2010)057 (SISSA, 2010). arXiv: 1009.1003 [physics.data-an]. http://pos.sissa.it/archive/conferences/093/057/ACAT2010_057.pdf (cit. on p. 92).

104. Cousins, R. D. *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms* in (2013). https://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf (cit. on p. 98).

105. The ATLAS collaboration. Search for dark matter produced in association with a single top quark in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *Eur. Phys. J. C* **81,** 860. arXiv: 2011.09308 [hep-ex] (2021) (cit. on p. 104).

106. Rübenach, J. *Search for heavy Higgs bosons in conjunction with neural-network-driven reconstruction and upgrade of the Fast Beam Condition Monitor at the CMS experiment* Dissertation (Universität Hamburg, Hamburg, 2023). https://bib-pubdb1.desy.de/record/584415 (cit. on p. 105).

107. Carloni, L. & Sjostrand, T. Visible Effects of Invisible Hidden Valley Radiation. *JHEP* **09,** 105. arXiv: 1006.2911 [hep-ph] (2010) (cit. on p. 106).

108. Carloni, L., Rathsman, J. & Sjostrand, T. Discerning Secluded Sector gauge structures. *JHEP* **04,** 091. arXiv: 1102.3795 [hep-ph] (2011) (cit. on p. 106).

109. Bali, G. S. *et al.* Mesons in large-N QCD. *JHEP* **06,** 071. arXiv: 1304.4437 [hep-lat] (2013) (cit. on p. 106).

110. Platzer, S. & Gieseke, S. Dipole Showers and Automated NLO Matching in Herwig++. *Eur. Phys. J. C* **72,** 2187. arXiv: 1109.6256 [hep-ph] (2012) (cit. on p. 107).

111. Buckley, A. *et al.* Rivet user manual. *Comput. Phys. Commun.* **184,** 2803–2819. arXiv: 1003.0694 [hep-ph] (2013) (cit. on p. 111).

112. Albouy, G. *et al.* Theory, phenomenology, and experimental avenues for dark showers: a Snowmass 2021 report. *Eur. Phys. J. C* **82,** 1132. arXiv: 2203.09503 [hep-ph] (2022) (cit. on pp. 111, 112, 115).

113. Fischer, C. S. Infrared properties of QCD from Dyson-Schwinger equations. *J. Phys. G* **32,** R253–R291. arXiv: hep-ph/0605173 (2006) (cit. on p. 111).

114. Ellis, R. K., Stirling, W. J. & Webber, B. R. *QCD and Collider Physics* (Cambridge University Press, 1996) (cit. on p. 113).

115. Gaffney, J. & Mueller, A. $\alpha(Q^2)$ corrections to particle multiplicity ratios in gluon and quark jets. *Nuclear Physics B* **250,** 109–142. https://www.sciencedirect.com/science/article/pii/0550321385904766 (1985) (cit. on p. 113).

116. Almeida, L. G. *et al.* Substructure of high-p_T Jets at the LHC. *Phys. Rev. D* **79,** 074017. arXiv: 0807.0234 [hep-ph] (2009) (cit. on p. 117).

117. Berger, C. F., Kucs, T. & Sterman, G. F. Event shape / energy flow correlations. *Phys. Rev. D* **68,** 014012. arXiv: hep-ph/0303051 (2003) (cit. on p. 117).

118. The CMS collaboration. Search for resonant production of strongly coupled dark matter in proton-proton collisions at 13 TeV. *JHEP* **06,** 156. arXiv: 2112.11125 [hep-ex] (2022) (cit. on p. 119).