SEARCH FOR DARK MATTER PRODUCED IN ASSOCIATION WITH A BOTTOM-QUARK IN PROTON-PROTON COLLISIONS AT $\sqrt{s}$ = 8 TeV WITH THE CMS EXPERIMENT

HOH SIEW YAN

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

SEARCH FOR DARK MATTER PRODUCED IN ASSOCIATION WITH A BOTTOM-QUARK IN PROTON-PROTON COLLISIONS AT $\sqrt{s} = 8$ TeV WITH THE CMS EXPERIMENT

HOH SIEW YAN

DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: HOH SIEW YAN

Registration/Matrix No.: SGR 130120

Name of Degree:

MASTER OF SCIENCE (EXCEPT MATHEMATICS & SCIENCE PHILOSOPHY)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

SEARCH FOR DARK MATTER PRODUCED IN ASSOCIATION WITH A BOTTOM-

QUARK IN PROTON-PROTON COLLISIONS AT $\sqrt{s} = 8$ TeV WITH THE CMS

EXPERIMENT Field of Study: EXPERIMENTAL HIGH ENERGY PHYSICS

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature | Date

Subscribed and solemnly declared before,

Witness's Signature | Date

Name:
Designation:

ABSTRACT

Dark Matter is a hypothetical particle proposed to explain the missing matter discovered from the cosmological observation. The motivation of Dark Matter is overwhelming, however those studies mainly deduced from gravitational interaction which does little to understand the underlying structure of the particle. On the other hand, the WIMP Miracle motivating dark matter production at weak scale, implying the possibility of detecting dark matter in LHC. Assuming Dark Matter is the only new particle accessible in LHC, it is expected to be pair produced in association with a Standard Model particles which can be inferred from the Transverse Missing Energy. Search for dark matter in monojet final state is a dominant channel in placing model-independent constrains on a set of effective operators coupling Dark Matter to Standard Model particle. However in the case of effective operators generated by the exchange of heavy scalar mediator, the inclusive Monojet channel is weakened due to the light quark suppression. An exclusive search on heavy quark final state such as bottom quark can improve the coupling strength of the Dark Matter process and reduce significantly the Standard Model backgrounds. A preliminary study on the search for Dark Matter produced in association with a bottom-quark is carried out as an extension to the Monojet search by requiring a b-tagged jet using the data collected by Compact Muon Solenoid detector in proton-proton collisions at center of mass energy of 8 TeV. The study shows that the data is consistent with Standard Model prediction and a limit is derived on the interaction scale and nucleon-dark matter scattering cross section.

ABSTRAK

Jirim gelap adalah zarah hipotesis yang dicadangkan untuk menjelaskan jirim yang terhilang ditemui dari cerapan kosmologi. Motivasi jirim gelap adalah tidak boleh dipersoalkan, namun kajian kosmologi tersebut terutamanya disimpulkan daripada interaksi graviti dan tidak menberikan gambaran yang jelas tentang struktur asal zarahnya. Sebaliknya, keajaiban WIMP (Zarah berat bertindak secara lemah) telah memotivasikan penghasilan jirim gelap pada skala lemah, mengimplikasikan bahawa kemungkinannya jirim gelap boleh dikesan di LHC (Panlanggar Hadron Besar). Dengan anggapannya jirim gelap adalah zarah baru yang boleh diakses di LHC, ia dijangka akan dihasilkan dalam pasangan bersama dengan zarah Standard Model dan jirim gelap tersebut boleh disimpulkan daripada Tenaga Hilang Lintangan. Pencarian jirim gelap dalam keadaan akhir monojet adalah saluran utama yang digunakan untuk meletakkan kekangan bersifat model bebas pada satu set operator bersifat kesan yang gandingan kepada zarah Standard Model. Walau bagaimanapun dalam kes operator bersifat kesan yang dijanakan oleh penukaran pengantara skalar berat, saluran monojet bersifat inklusif dilemahkan yang disebabkan oleh penindasan quark ringan. Kajian bersifat eksklusif tentang keadaan akhir quark berat seperti quark bawah boleh meningkatkan kekuatan gandingan proses jirim gelap dan dijanka akan mengurangkan latar belakang Standard Model dengen berkesan. Kajian pencarian jirim gelap dihasil bersama dengan penghasilan quark bawah dijalankan sebagai kajian lanjutan kepada pencarian monojet dengan memerlukan jet di-b-tag-kan dikaji kepada data yang dikumpul oleh pengesan Compact Muon Solenoid dalam perlanggaran proton-proton pada pusat tenaga jisim 8 TeV. Kajian tersebut menunjukkan bahawa data adalah konsisten dengan jangkaan Standard Model dan kekangan diletakkan pada skala interaksi dan keratan rentas berselerak.

ACKNOWLEDGEMENTS

The completion of this work is tremendously motivated by someone whom being the closest to me, cherishes every single moment with me, and selflessly cares for me. It is impossible to mention all of them, I wish to mention those whom are significant to me at the level of enlightenment.

To the greatest extent, I would like to express my gratitude to the director of NCPP, Wan Ahmad Tajuddin Wan Abdullah and Zainol Abidin Ibrahim for provision of valuable resources and facilities, most of the work is impossibly done without much of the effort contributed by them. Norraphat Srimanobhas, for whom he is an inspiring figure, a charismatic friend and a visionary scientist whom has offered me a golden opportunity to participate in this international and prestigious collaboration. Also, Jyothsna Rani Komaragiri whom guided me along the way and facilitated me in the analysis group had helped in materializing much of the work in this thesis.

The members of the monojet group during the year 2013 and 2014 deserves more credit for their collective effort invested on the analysis framework used in this thesis. I wish to thank CERN for offering an amazing opportunity for me to work in frontier research and CMS experiments for giving me chance to prove myself as a competent scientist. Lastly, I would like to thank my friends and my love one as a source of solace and inspiration during difficult time, especially Afiq Aizuddin Anuar, Nurfikri Norjoharuddeen, Amjad Kotobi and Chew Sia Ling. Of course my beloved parents, Kin Chin Hoh and Siew Choo Chin for their selflessness and countless sacrifice.

I would like to express my greatest gratitude to all of them, without them I would not be able to work with people from around the world and my academic life would not be as flourishing as what I am currently enjoying. This work is dedicated to all of you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

E_T^{miss}	:	Missing Transverse Energy.
M_*	:	Effective Contact Interaction Scale.
M_T	:	Transverse mass.
M_χ	:	Dark Matter mass.
ΛCDM	:	Cold Dark Matter Model.
η	:	Pseudorapidity.
$\sigma_{\chi-nucleon}$	:	Nucleon-DM scattering cross section.
p_T	:	Transverse Momentum.
ALICE	:	A Large Ion Collider.
ATLAS	:	A Toroidal LHC Apparatus.
CERN	:	European Organization for Nuclear Research.
CKM	:	Cabibbo-Kobayashi-Maskawa.
CL	:	Confident Level.
CMS	:	Compact Muon Solenoid.
CSV	:	Combine Secondary Vertex.
DAQ	:	CMS Data Acquisition System.
DM	:	Dark Matter.
EFT	:	Effective Field Theory.
FSR	:	Final State Radiation.
HLT	:	High-Level Trigger.
ISR	:	Initial State Radiation.
L1	:	Level-1 Trigger.
LHC	:	Large Hadron Collider.
LHCb	:	Large Hadron Collider beauty.
LHE	:	Les Houches Event Files.
LO	:	Leading Order.
MACHO	:	Massive Compact Halo Object.
MC	:	Monte Carlo.
ME	:	Matrix Element.
MFV	:	Minimum Flavor Violation.
MSSM	:	Minimal Supersymmetric Standard Model.
PDF	:	Parton Density Function.
PF	:	Particle Flow algorithm.
POG	:	Physics Object Group.
QCD	:	Quantum Chromo-Dynamics.
SM	:	Standard Model.
SUSY	:	Supersymmetry.
WIMP	:	Weakly Interacting Massive Particle.

CHAPTER 1: GENERAL INTRODUCTION

Mankind always finds themselves constantly struggling to push beyond the limit imposed by the nature. The basic question such as what is the universe made of unfortunately requires multidisciplinary studies to answer. Currently we can confidently state that we do not know the complete answer to the question. This state of affairs is probably unique in the history of mankind. But with the persistence and the gifted sentient capability bestowed on mankind, a surge of hypothesis and speculation which evolved into a working model has been seen on the face of the earth since the dawn of mankind. One of the fundamental questions is the existence of Dark Matter (DM) consistently hypothesized by Astrophysical studies. Many models are being proposed to explain the DM phenomena, particularly the popular model such as Weakly Interacting Massive Particle (WIMP) model which postulates the existence of non-relativistic weak-scale mass particle interacts weakly with Standard Model (SM) particles. If DM is a new kind of particle produced thermally in the early universe, the Large Hadron Collider (LHC) which serves as an instrument of discovery can potentially produce and eventually detected by the detector if and only if the energy of the LHC is high enough for its production.

So far DM has not been observed in particle physics experiments, nor it has been observed in non-gravitational interaction in astrophysical study. In the collider experiment, the only way to study about DM is when they are produced in association with SM particles. Such interaction might be observed in LHC as a jet or particle recoiled with an invisible state, which is interpreted as Missing Transverse Energy (E_T^{miss}). Most of the studies assume the interaction exists between SM particles and the DM, if this is not the case, then proton collisions will not directly produce DM particles, and will not scatter off nuclei in direct detection experiments. To further expand the assumption, often DM is taken to be a Dirac fermion WIMP and it is stable on collider timescales and non-interacting with the detector. In high energy physics, the subsequent process after the collision usually involves only one mediator and one search channel playing the dominant role in the discovery of new physics in LHC.

1.1 Project Statement

The study is carried out on the data collected by Compact Muon Solenoid (CMS) experiment with an integrated luminosity of 19.7 fb^{-1} in center of mass energy of 8 TeV in the proton-proton collision scenario. The search for DM has a characteristic topology of X + E_T^{miss} channel which X is any SM particles, one such channel is a single hadronic jet (monojet) with large E_T^{miss}. The sensitivity of the monojet channel is challenged by the modeling of dominant background such as the Z+jets process which decays into a pair of neutrino and the Initial State Radiation (ISR) or Final State Radiation (FSR) in the monojet final state. In order to minimize the impact of the incurred uncertainty, several studies are proposed to improve the collider limit, in the aspect of background modeling and sensitivity study by b-tag the monojet channel.

1.1.1 Data-Driven Background Estimation on $Z(\nu\bar{\nu})$ Backgrounds

Since the channel is dominated by irreducible background process such as $Z(\nu\bar{\nu})$ events, a control sample is required to model from background from data. The control sample is necessary to have similar topology and kinematics with respect to the background process. One of the shortcoming of the monojet analysis is the standing issue of limited size of control sample such as $Z(\mu\bar{\mu})$, which is indicated in higher E_T^{miss} region (Khachatryan et al., 2015). Therefore a sufficient number of events from the control sample is essential to reduce statistical uncertainty in high E_T^{miss} region. The study on using $W(\mu\nu)$ events as a control sample to predict the number of $Z(\nu\bar{\nu})$ events is conducted, later the predicted background is propagating into the sensitivity study on mono b-jet channel.

1.1.2 Sensitivity Study on Mono b-jet Channel

Besides improving the modeling of backgrounds, choices of search channel constituting a significant role in the analysis attributing to the different physics properties of the process. Based on the phenomenological study (Lin, Kolb, & Wang, 2013) on mono b-jet + E_T^{miss} final state, the corresponding collider limit has improved with a factor of $\sim$10 compared to inclusive monojet channel. By interpreting mono b-jet channel in Effective Field Theory (EFT), the strength of a scalar interaction between DM particles coupling to SM particles is proportional to the mass of the quarks. Therefore focusing on heavy flavour quark such as b quark, the coupling strength is expected to increase significantly compares to light flavour quarks, which may provide a potential signal in LHC. As a consequence of EFT, the interaction between DM and SM particles is model-independent. Therefore the collider-derived limit on DM phase space can be translated into limit derived from DM direct detection experiment.

1.2 Objectives

The thesis intended to achieve the objectives listed below:

- To study and to access the feasibility of employing alternative control samples in estimating $Z(\nu\bar{\nu})$ background process.
- To study the Mono b signal sensitivity by benefiting from the existing Monojet analysis in term of limit on the interaction scale and DM-nucleon cross section.
- To access the feasibility of a dedicated analysis based on mono b-jet channel.

1.3 Thesis Outline

Throughout the thesis, high energy physics quantity is expressed in natural units by assigning c=1. Therefore mass, energy, and momentum have the same dimension which is GeV (Giga-electrovolt) (Nakamura & Group, 2010) for ease of interpretation.

The first chapter served as a general introduction to the study, highlighting the motivation and objectives of each chapter. The second chapter describes the novelty relevant to the aspect of the CMS detector and how data is collected. Besides, DM searches at LHC and the other DM detection technique are briefly discussed to provide a picture of generic DM searches. Chapter 3 touches on the data-driven background estimation study by comparing two different control samples. Chapter 4 concerns the sensitivity study on mono-b + E_T^{miss} channel on the CMS collected data interpreted in DM search. Chapter 5 summarizing all the work from the studies.

CHAPTER 2: LITERATURE REVIEW

2.1 The Standard Model of Particle Physics

The SM of particle physics has, for many years, accounted for all observed particles and interactions between them (Nakamura & Group, 2010). Despite of this success, it is clear that a more fundamental theory must exist, whose low energy realization should coincide with the SM. In the SM framework, the fundamental constituents of matter are fermions: quarks and leptons. Their interaction is mediated by integer spin particles called gauge bosons. Strong interactions are mediated by gluons G_a, electroweak interaction by $W^{\pm}$, Z_0, γ and the Higgs boson H^0. The left-handed leptons and quarks are arranged into three generations $SU(2)_L$ doublets with the corresponding right-handed fields transferring as singlets under $SU(2)_L$.

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \tag{2.1}$$

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L \tag{2.2}$$

Each generation contains two flavours of quarks with baryon number B = 1/3 and lepton number L = 0 and two leptons with B = 0 and L = 1. Each particle also has a corresponding antiparticle with the same mass and opposite quantum numbers. The quarks which are primed are weak eigenstates related to mass eigenstates by the Cabibbo-Kobayashi-Maskawa (CKM) Matrix (Cabibbo, 1963; Kobayashi & Maskawa, 1973).

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \hat{V}_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.3}$$

Gauge symmetries play a fundamental role in particle physics. It is in fact in terms of symmetries and using the formalism of gauge theories that we describe electroweak and strong interaction.

The SM is based on the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge theory, which undergoes the spontaneous breakdown:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_C \otimes U(1)_Q \tag{2.4}$$

where Y and Q denote the weak hypercharge and the electric charge generators, respectively, and $SU(3)_C$ describes the strong (color) interaction, known as Quantum Chromo-Dynamics (QCD). This spontaneous symmetry breaking results in the generation of the massive $W^{\pm}$ and Z gauge bosons as well as a massive scalar Higgs field.

2.2 The Large Hadron Collider

The LHC is the world largest and most powerful particle accelerator. It first started up on 10 September 2008, and remains the latest addition to European Organization for Nuclear Research (CERN) accelerator complex. The LHC consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way (CERN, 2014).

Inside the accelerator, two high energy particle beams travel at close to the speed of light before they are made to collide. The beams travel in opposite directions in separate beam pipes, two tubes kept at ultrahigh vacuum. They are guided around the accelerator ring by a strong magnetic field maintained by superconducting electromagnets. The electromagnets are built from coils of special electric cable that operates in a superconducting state, efficiently conducting electricity without resistance or loss of energy. This requires chilling the magnets to $-271.3\,^{\circ}$C, a temperature colder than outer space. For this reason, much of the accelerator is connected to a distribution system of liquid helium, which cools the magnets, as well as to other supply services.

Thousands of magnets of different varieties and sizes are used to direct the beams around the accelerator. These include 1232 dipole magnets 15 meters in length which bend the beams, and 392 quadrupole magnets, each 5$\sim$7 meters long, which focus the beams. Just prior to collision, another type of magnet is used to "squeeze" the particles closer together to increase the chances of collisions. The particles are so tiny that the task of making them collide is akin to firing two needles 10 kilometers apart with such precision that they meet halfway.

All the controls for the accelerator, its services and technical infrastructure are housed under one roof at the CERN Control Center. From here, the beams inside the LHC are made to collide at four locations around the accelerator ring, corresponding to the positions of four large particle detectors, there are experiment A Toroidal LHC Apparatus (ATLAS), CMS, A Large Ion Collider (ALICE) and Large Hadron Collider beauty (LHCb). More information on each different particle detectors can be found in `http://home.web.cern.ch/`.

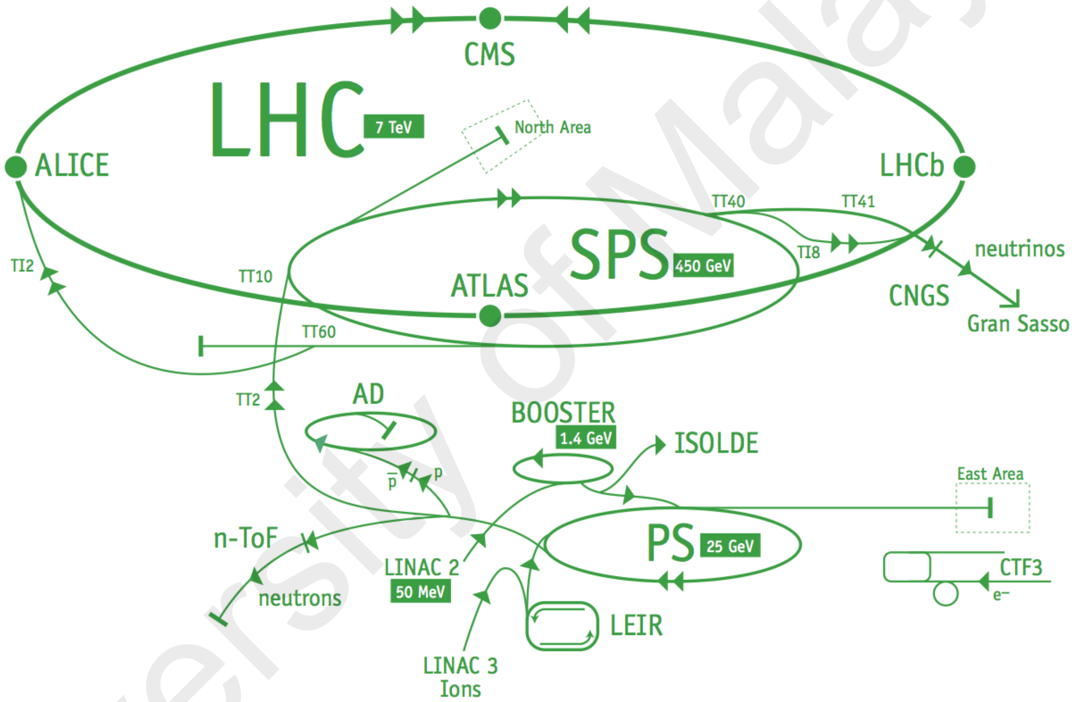

Figure 2.1: A schematic view of the LHC housed by the France and Switzerland border. (TE-EPC-LPC, 2012)

2.3 The CMS Detector

The CMS detector is a multi-purpose apparatus operating at the LHC. The CMS detector features a superconducting solenoid, 12.5 m long with an internal diameter of 6 m, providing a uniform magnetic field of 3.8 T (CMS Collaboration, 2008). The CMS coordinate system is oriented such that the x-axis points south to the center of the LHC ring, the y-axis points vertically upward and the z-axis is in the direction of the beam to the west. The azimuthal angle ϕ is measured from the x-axis in the x-y plane and the radial coordinate in this plane is denoted by r . The polar angle θ is defined in the r-z plane and the pseudorapidity is $\eta = -\ln \tan(\theta/2)$. The momentum component transverse to the beam direction, denoted by p_T , is computed from the x- and y-components, while the transverse energy is defined as $E_T = E \sin\theta$.

Within the field volume are a silicon pixel and strip tracker, a crystal electromagnetic calorimeter and a brass/scintillator hadron calorimeter. The momentum resolution for reconstructed tracks in the central region is about 1.5% for non-isolated particles with Transverse Momentum (p_T) between 1 and 10 GeV and 2.8% for isolated particles with p_T of 100 GeV. The calorimeter system surrounds the tracker and consists of a scintillating lead tungstate crystal electromagnetic calorimeter and a brass/scintillator hadron calorimeter with coverage up to $|\eta|$ = 3. The quartz/steel forward hadron calorimeters extend the calorimetry coverage up to $|\eta|$ = 5. A system of gas-ionization muon detectors embedded in the steel flux-return yoke of the solenoid allows reconstruction and identification of muons in the $|\eta| < 2.4$ region. Events are recorded using a two-level trigger system. A schematic view of the detector can be visualized in Figure 2.2.

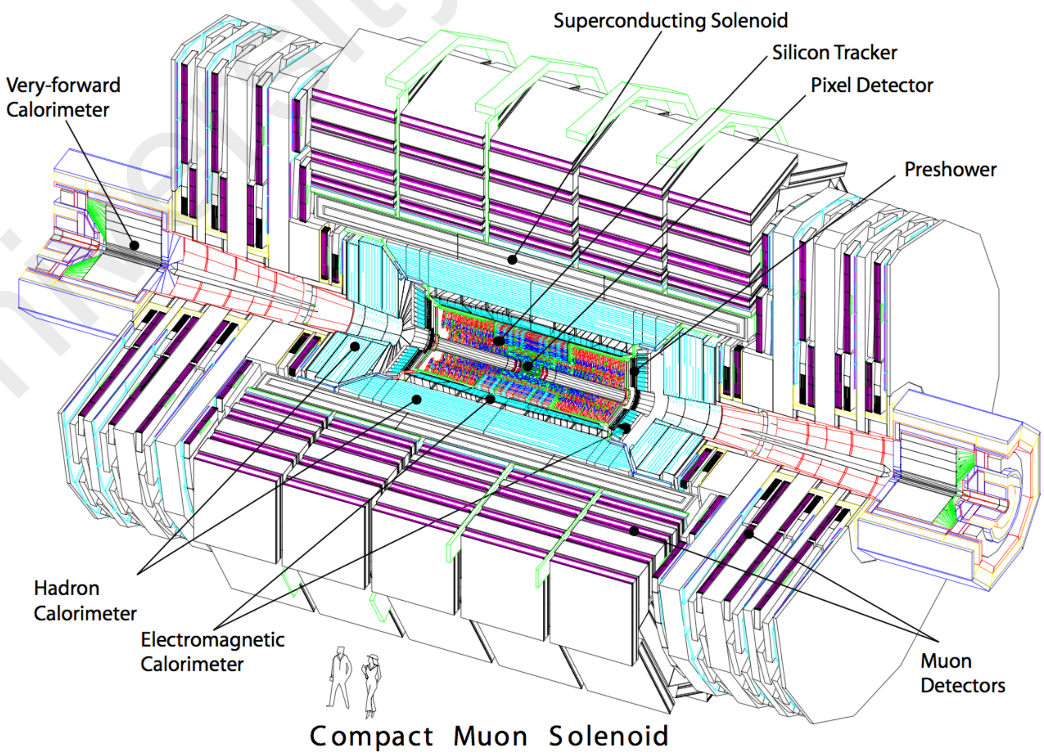

Figure 2.2: A schematic view of the CMS detector (CMS Collaboration, 2008).

2.4 Trigger System in the CMS Experiment

The LHC provides proton-proton and heavy-ion collisions at high interaction rates. For protons the beam crossing interval is 25 ns, corresponding to a crossing frequency of 40 MHz. Since it is impossible to store and process the large amount of data associated with the resulting high number of events, a drastic rate reduction has to be achieved. This task is performed by the trigger system, which is the start of the physics event selection process. The CMS trigger system reduces the events rate in two steps called Level-1 Trigger (L1) and High-Level Trigger (HLT). The L1 is designed to achieve a maximum output rate of 100 kHz and consists of custom-designed, programmable electronics while the HLT is based on software algorithms running on a large cluster of commercial processors, the event filter farm (Chatrchyan et al., 2008).

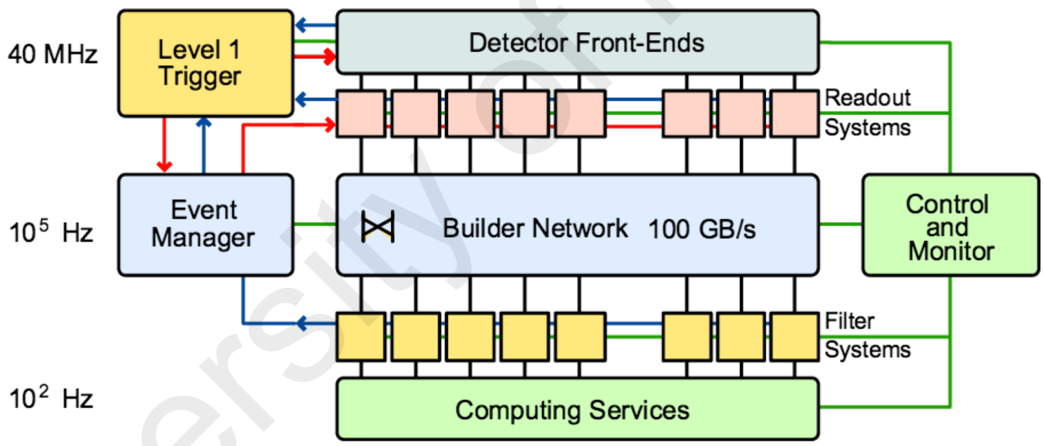

Figure 2.3: Architecture of the CMS DAQ system (Chatrchyan et al., 2008).

The CMS Data Acquisition System (DAQ) is designed to collect and analyze the detector information delivered by the CMS. Since the L1 reduces the incoming average data rate to a maximum of 100 kHz, therefore the DAQ system must sustain a maximum input rate of 100 kHz, and must provide enough computing power for a software filter system, HLT to reduce the rate of stored events by a factor of 1000. All events that passes the L1 are sent to a computer farm or an event filter that performs physics selections, using faster versions of the offline reconstruction software, to filter events and achieve the required output rate. The architecture of the DAQ is shown schematically in Figure 2.3.

If an event is accepted by the L1, the full detector information is read out by the DAQ, passed to the event filter farm and used as input for the HLT. The HLT algorithms are implemented in the same software as used for offline reconstruction and analysis and consist of subsequent reconstruction and selection steps.

2.5 Physics Object Reconstruction in the CMS Experiment

The CMS experiment uses the Particle Flow algorithm (PF) event reconstruction technique (CMS Collaboration, 2010) for reconstructing and identifying all stable particles in the event, such as electrons, photons, muons, charged hadrons and neutral hadrons. In this technique the PF algorithm making use of information from all CMS sub-detectors and reconstructing into a physics objects.

As illustrated in Figure 2.4, the tracker is the cornerstone of the PF event reconstruction considering good momentum resolution and precise measurement of the charged-particle direction at the production vertex. On the other hand, stable neutral particles such as photons and neutral hadrons are not reconstructed by the tracker thus it appears as dotted line in the tracker. The information from the calorimeters are used to find the energy and direction of the neutral particles. A specific clustering algorithm is developed for the PF event reconstruction and is performed separately in each sub-detector. After the reconstruction and identification processes, the list of the particles are then used for building the jets, determining the missing transverse energy, reconstructing the decay products and so on.

2.5.1 Muons

The muon reconstruction at the CMS is based on the track reconstruction at the tracker and muon detectors. The matched energy deposits in the calorimeters are also used in the muon reconstruction (CMS Collaboration, 2012a). The tracks in the silicon tracker and muon spectrometers are reconstructed independently and are called tracker tracks and standalone tracks, respectively. Then, two complementary approaches are used for the muon reconstruction from these tracks.

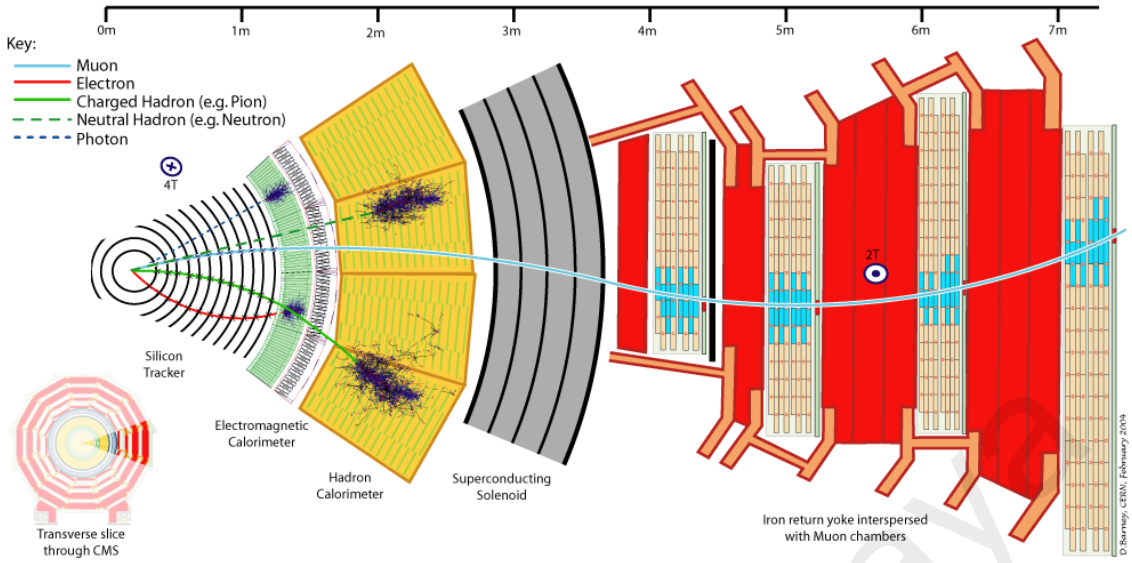

Figure 2.4: The transverse slice of the CMS detector and the interaction of each type of particle with the sub-detectors. Different type of particle interacts with different CMS components, various colored curved lines represent paths that a particular particle might take.

The Global Muon reconstruction approach perform a global fit on the matched standalone muon track and tracker track. On the other hand, the Tracker Muon reconstruction approach extrapolates all tracker tracks to the muon system and requires at least one muon segment in the muon system is matched with the tracker track. A reconstructed muon will be labeled as global muon if a match is found in the former approach, and tracker muon for the latter approach. However, if both approaches fail and a standalone track left without any tracker track, then the muon will be labeled as standalone muon. All the muon candidates are merged into a single collection, each one containing information from tracker, standalone and global fits. The candidates which are reconstructed by both approach are merged into a single candidate (CMS Collaboration, 2012a). Additional information is required for muon quality identification and selection. The isolation cut is performed to reduce contamination from muons that originate from hadronic processes. The isolation variable is calculated using PF candidates in a cone of a given size 0.4 around the muon track direction. On the other hand, in order to reduce the contribution of the muons originating from the cosmic rays, heavy flavor decays and hadronic showers, two selection working points are used, which is the tight and loose collections.

2.5.2 Jets

The original quark or gluon is never seen in its free states and they bind into colorless hadrons due to the QCD confinement. In the collision, a quark or gluon hadronises immediately after being produced and subsequently the produced spray of the hadrons travel more or less in the direction of the final state parton, collectively called a jet. The jet algorithm is invented to cluster hadrons into a well defined jet, it usually involve one or more parameters that indicate how two particles are in a same or separate jet. The CMS experiment uses anti-k_T algorithm to define the jet (Cacciari, Salam, & Soyez, 2008a). The anti-k_T is a sequential recombination algorithm which uses the following distances:

$$d_{ij} = min(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2})\frac{\Delta_{ij}^2}{R^2} \tag{2.5}$$

$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{2.6}$$

where $\Delta_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and $p_{Ti}, y_i = \frac{1}{2}ln(\frac{E+p_z}{E-p_z})$ and ϕ_i are transverse momentum, rapidity and azimuth of particle i. R is the radius parameter similar to radius in cone algorithm. To cluster a jet, the anti-k_T algorithm computes and identifies the minimum distance of d_{ij} and d_{iB} according to Equation 2.5. If the minimum distance is d_{ij}, particles i and j are recombined into a single particle. Else, declares i as a jet and remove it from the list of particles if the minimum distance is d_{iB}. The procedure will stop when no particles remain. The value of d_{ij} is determined by the p_T of particles and separation between particles, Δ_{ij}. If there is a hard particle with high p_T between soft particles, the minimum of d_{ij} occurs when i is hard particle and j is a soft particle close to the hard particle. Therefore a hard particle simply accumulate all the soft particles within a circle of radius R through the anti-k_T algorithm and leads to a perfectly conical jet. Jets are reconstructed from several types of inputs as shown in Figure 2.5:

- **gen-jets**:stable simulated particles, except for neutrinos, are clustered after hadronization and before interaction with the detector.

- **PF-jets**: all PF candidates are clustered without distinction of type and any energy threshold.

- **calo-jets**: calorimeter towers in each hadron calorimeter cells and underlying electromagnetic calorimeter crystals are clustered.

The four-momentum vectors of PF candidates are used to reconstruct jets by the anti-k_T algorithm with R=0.5 in CMS. PF jets take advantage of the excellent momentum and spatial resolutions for the charged hadrons and photons inside a jet, which together constitutes 85% of the jet energy. The PF jet momentum and spatial resolutions are greatly improved with respect to calorimeter jets. Gen-jets are used as a reference to compare the PF-jet performance to the calo-jet performance (CMS Collaboration, 2009b).

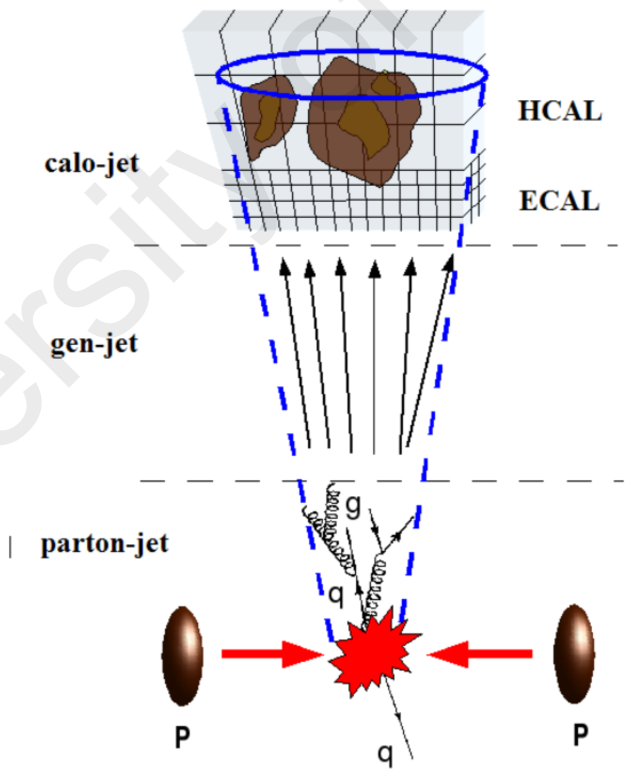

Figure 2.5: Schematics of the jet reconstruction from parton level, hadron level and caloremeter level objects. HCAL refers to hadron calorimeter and ECAL refers to electromagnetic calorimeter.

2.5.3 Missing Transverse Energy

The vector momentum imbalance in the plane perpendicular to the beam direction is known as missing transverse momentum (E_T^{miss}) and its magnitude is called missing transverse energy (MET). MET is the transverse momentum that must have been carried by something invisible such as neutrinos. E_T^{miss} is the important variable in reducing background which involves the $Z(\nu\bar{\nu})$ event and the leptonic decay of W bosons. The PF MET is reconstructed as the negative vector sum of the transverse momentum of all PF candidate particles in the event (CMS Collaboration, 2009b). The magnitude of the E_T^{miss} can be affected by various sources. The minimum energy thresholds in the calorimeters, minimum p_T threshold in tracker and non-linear response of calorimeters would vary the E_T^{miss}.

Anomalous E_T^{miss} measurements existed in data from calorimeter noise and beam halo. The beam halo phenomenon are showers of secondary particles produced through the interaction of protons with the beam collimators or the residual gas particles.

2.6 Dark Matter

Dark Matter, as its name inferred it is dark or non-luminous, is the proposed unknown entity to explain the missing mass detected in various cosmological phenomena. The history of DM discovery can be dated up to 1933. Zwicky predicted the total mass of the Coma Cluster (Zwicky, 2009), which to his surprise the mass inferred from the relative velocity is 400 times the mass of the visible star in galaxies in the cluster. The observation was soon confirmed by a similar measurement of the Virgo Cluster carried out by Smith (Smith, 1936). A consistent contradiction was further substantiated by Vera Rubin who found that most of the galaxy rotational velocity remained constant at larger distance, contradicting the Keplerian prediction (Rubin, Ford, & . Thonnard, 1980).

Despite there were theoretical efforts carried out in spirit to explain the observation by modifying the Newtonian mechanics rather than introducing a new form of matter (Sanders & McGaugh, 2002), inevitably a new form of particle is strongly motivated due to the compelling evidence of IE0657-558 (Clowe et al., 2006). Evidently the non-baryonic matter does not interact with the baryonic matter. From the total density of mass distribution measured it showed the total mass moved ballistically after the collision, indicating DM self-interactions were weak in nature.

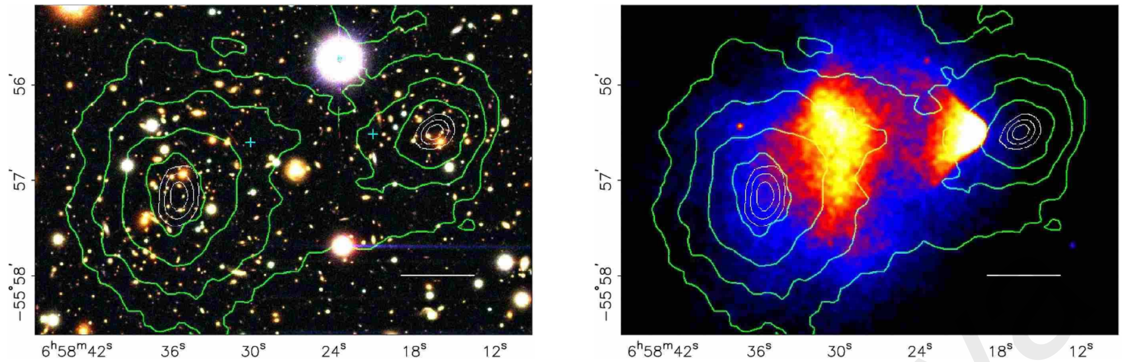

Figure 2.6: Images of the bullet cluster, 1E0657-558: optical image from the Hubble Space Telescope (left) and X-ray image from Chandra telescope (right). The mass density contours from gravitational lensing reconstruction superimposed on two images (Clowe et al., 2006).

2.7 WIMP as Dark Matter Candidate

A wide range of DM candidates such as Massive Compact Halo Object (MACHO) and Primordial Black Hole were excluded based on the cosmic microwave background and the large-structure formation study which subsequently theorized Cold Dark Matter Model (ΛCDM) (Primack, 2009). The DM particle is known to be massive due to its non-relativistic speed, invisible, electromagnetically neutral, and only interact with weak interaction and gravity. Initially the neutrino was suspected to be DM candidate but it was immediately dismissed as it was not massive enough. Consequently, the Standard Model does not provide a viable DM candidate. However in physics beyond Standard Model such particle happens to be naturally motivated by a wide spectrum of models which attempt to solve gravity and strong-CP problem. Besides sterile neutrinos and Axions, the compelling DM candidate is WIMP.

The early universe was full of radiation fueled by the constant pair-production and annihilation of particles while establishing a thermal equilibrium. The WIMP was assumed to produce and annihilate in the same fashion as other particles until it decouples from the thermal equilibrium as the annihilation process diminishes due to the subsequent universe expansion (Feng, 2010). Figure 2.7 shows the evolution of thermal relic density for WIMP with mass M_χ =100 GeV, when the

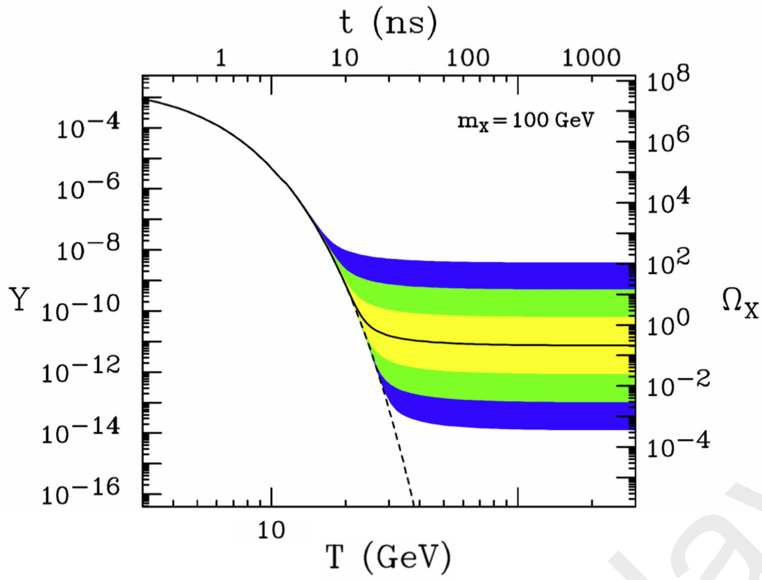

Figure 2.7: The number density Y (left) of a 100 GeV, P-wave annihilating Dark Matter particle evolves as a function of time t (top) and temperature T (bottom). The solid contour is the annihilation cross section yields the correct relic density Ω_χ (right), which is ~ 0.23 (Feng, 2010).

universe was expanded and cooled at $T < M_\chi$, the DM was "frozen-out" and left over, manifestly the comoving number density Y remained constant and produced the present thermal relic density Ω_χ. By solving the Boltzmann equation the relic density implied that the average DM annihilation cross section was $\sim <\sigma_A v>$=1pb which turns out to be the typical cross section the LHC is currently producing. Furthermore a simple dimensional analysis suggests if the correct thermal relic was made up of DM, the DM particle falls on the weak-scale mass range of 100 GeV $\sim$ 1 TeV. This connection between Cosmology and Particle Physics establishing the WIMP Miracle. Naturally the WIMP Miracle implied many model providing viable DM candidates. Example of DM candidates are the lightest neutralino from weak-scale Supersymmetry (SUSY) and Kaluza-Klein photon from extra dimensions.

2.8 Theoretical Model in Dark Matters

The evidence of DM is compelling from the cosmological studies, a cosmological observation that is strongly supported by the large-scale structure of the Universe and measurements of the cosmic microwave background. While the existence of DM thus seem well established, very little is known about the properties of the DM particles. The most logical approach to learn more about the DM is to interact with a SM particles through the thermal relic freeze-out framework, which define the core in all DM detection (Abdallah et al., 2015). In order to interpret the result from DM detection, a theory of DM is needed. As illustrated in Figure 2.8, a large number of qualitative different DM models can be constructed. Collectively these models populate the "theory space" of all possible realizations of physics beyond the SM with a particle that is a viable DM candidate.

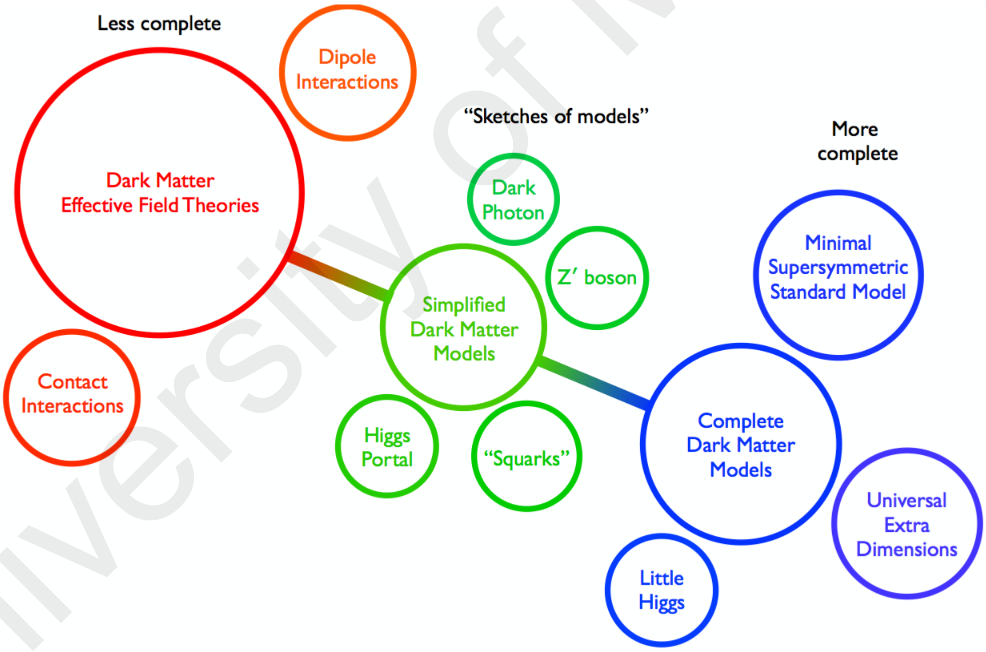

Figure 2.8: Artistic view of the DM theory space.

On the simple end of the spectrum, we have theories where the DM may be the only accessible state to our experiments. In such a case, EFT allows us to describe the interaction between DM and SM mediated by all kinematically inaccessible particles in an universal way. The EFT approach allows to derive stringent bounds on the "new-physics" scale, or Effective Contact Interaction Scale (M_{*}) suppresses the higher-dimensional operators. Since for each operator the M_{*}

encodes the information on all the heavy states of the dark sector, comparing LHC bounds to the limits following from direct and indirect DM searches is straightforward in the context of EFT.

The large energies accessible at the LHC has brought up a question on the momentum expansion underlying the EFT approximation and we can expand our level of detail toward simplified DM models. Such models are characterized by the most important state mediating the DM particle interactions with the SM, as well as the DM particle itself. Unlike the EFT, simplified models are able to describe correctly the full kinematics of DM production at the LHC, because they resolve the EFT contact interactions into single particle s-channel or t-channel exchanges. This comes with the price that they typically involve not just one, but a handful of parameters that characterize the dark sector and its coupling to the visible sector.

While simplified model captures some set of signals accurately at LHC energies, they are likely to miss important correlations between observable. Complete DM models close this gap by adding more particles to the SM, most of which are not suitable DM candidates. The classical example is the Minimal Supersymmetric Standard Model (MSSM), in which each SM particle gets its own superpartner and the DM candidate, the neutralino, is a weakly interacting massive particle. Reasonable phenomenological models in this class have of order 20 parameters, leading to varied visions of DM. At the same time, they build-in correlations from symmetry-enforcing relations among couplings, that would look like random accidents in a simplified model description. Complete DM models can in principle answer any question satisfactorily, but one might worry that their structure is so rich that it is impossible to determine unambiguously the underlying new dynamics from a finite amount of data ("inverse problem").

2.8.1 Model Independent Approach

In accordance to the thermal relic freeze out framework, the DM has to minimally interact weakly with SM in order to produce the current thermal relic. The interaction of DM with SM is heavily model-dependent because the coupling usually is restricted by gauge invariance and other symmetries especially for a higher spin DM particle. It may be useful to assume a more generic interaction rather than model-oriented in hope of understanding which is truly generic to Dark Matter physics. In the model-independent approach, with the assumption that DM exists and stable, the interactions between DM and SM are parametrized by a set of effective non-renormalizable operators which mimics the effect of heavy mediator. A realization of EFT can be visualized in figure 2.9.

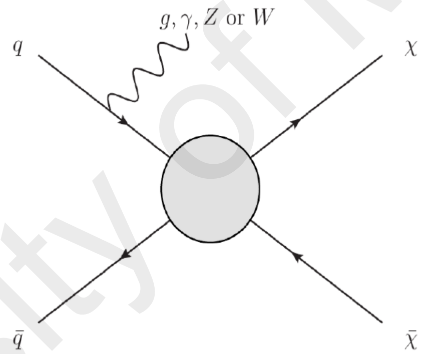

Figure 2.9: In the Mono-X topology, the interaction between DM and SM are described by a physical scale as a result of the mediator being factorized out.

The factorization of heavy mediator generates a set of DM contact operator showed in Figure 2.10 which effectively describe the interaction (Goodman et al., 2010; Busoni, De Simone, Morgante, & Riotto, 2014). The EFT approach is justified whenever there is a clear separation between the M_* and the underlying microscopic interaction of the process. For example, for indirect DM searches the annihilation of non-relativistic DM particles occurs with momentum transfer Q^2 of the order of Dark Matter mass (M_χ); in direct searches, the Q^2 is of few order of tens of keV in the scattering process of DM on heavy nuclei. Therefore it is possible to carry out an effective description in terms of DM operator with a Ultra-Violet cutoff larger than the typical Q^2 to limit on M_*.

Name	Initial state	Type	Operator
D1	qq	scalar	$\frac{m_q}{M_*^3}\bar{\chi}\chi\bar{q}q$
D5	qq	vector	$\frac{1}{M_*^2}\bar{\chi}\gamma^\mu\chi\bar{q}\gamma_\mu q$
D8	qq	axial-vector	$\frac{1}{M_*^2}\bar{\chi}\gamma^\mu\gamma^5\chi\bar{q}\gamma_\mu\gamma^\mu q$
D9	qq	tensor	$\frac{1}{M_*^2}\bar{\chi}\sigma^{\mu\nu}\chi\bar{q}\sigma_{\mu\nu}q$
D11	gg	scalar	$\frac{1}{4M_*^3}\bar{\chi}\chi\alpha_s(G^s_{\mu\nu})^2$

Figure 2.10: List of effective DM operators describing the interaction between SM in EFT assuming DM is Dirac fermion.

However for LHC searches it is dramatically different from the other DM searches. The Q^2 involved can be very high that the EFT description is no longer valid. Nevertheless under some condition the EFT description still can be valid if the energy scale involving the DM and SM process is smaller compared to the energy scale of heavy mediator (Busoni et al., 2014). A light mediator version or a simplified model where it assumes only one new particle with mass M_{med} which may come from the dark sector accurately describes the DM and SM interaction. After all EFT is an approximation of simplified model corresponding to expanding the propagator of the heavy mediator in powers of Q_{tr}^2/M_{med}^2 truncating at lowest order.

2.9 The Collider Searches

Usually the collider searches focus on leading generic Feynman Diagram responsible for DM production, specifically a pair production of WIMP pair plus the ISR/FSR of a gluon, photon or a weak gauge boson Z, $W^{\pm}$. Different channels have different coupling to the WIMP. On the detector level, the ISR/FSR particle is needed to balance the two WIMP momentum allowing the events "triggerable", visualized in Figure 2.11. Generally the selected events requires a central leading parton with high p_T and a WIMP pair with high E_T^{miss} to forms a back-to-back topology. Additionally the angular separation $\Delta\phi$ between the leading parton and E_T^{miss} has to be more than 0.5 radians to minimize the contribution from fake missing energy from jet mis-measurements (Beltran, Hooper, Kolb, Krusberg, & Tait, 2010). Given the immensity of the energies make available in the collision, most of the interactions is dominated by electroweak and QCD multi-jet processes. On event selection level most of the background processes for example the top-quark pair production, single top process, QCD multi-jet process and diboson processes were easily removed by kinematics cuts, however some dominant backgrounds are not rejected and thus require a data driven background estimation technique.

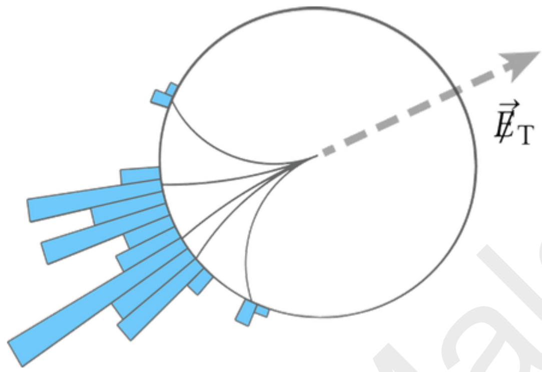

Figure 2.11: The typical collider signature of particles recoiled against a large E_T^{miss}. The manifestation of the mono-jet will be targeted to trigger the E_T^{miss} object in the events of interest.

2.9.1 Monojet

The monojet channel is expected to give the strongest coupling due to the rate of gluon and quark ISR is larger relative to other SM radiation as shown in Figure 2.12. The gluon and quarks are not a free particles which further hadronized along their projectile forming a spray of particle called jet which introducing higher uncertainties at higher E_T^{miss}. The common backgrounds to the monojet is the irreducible background $Z(\nu\bar{\nu}) + jets$ where the neutrinos decayed from the Z boson escaped the detector undetected, and the $W + jet$ where the decay lepton mis-identified and therefore out of detector acceptance or "lost". Nevertheless their contribution to the signal region can be estimated from the selected control samples from the signal events with a specific transfer function to account for kinematics differences (Askew, Chauhan, Penning, Shepherd, & Tripathi, 2014).

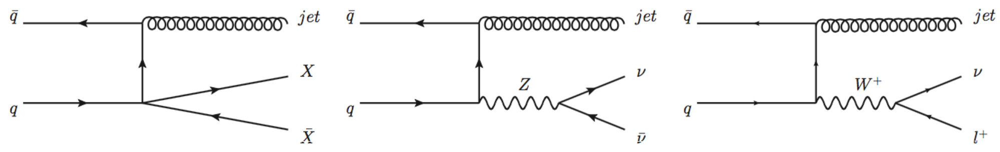

Figure 2.12: The Feynman Diagram of DM pair production in association with gluon (left), and the background $Z(\nu\nu)+jet$ (middle) and $W(l\nu)+jet$ (right).

2.9.2 Monophoton

Since the indirect detection is looking for an energetic gamma-ray, the monophoton channel is a potential channel which consist of a final state of large E_T^{miss} and an energetic photon. The photon required to pass through a quality tight cut to ensure it is not an electron faking photon. In term of backgrounds, the $Z+\gamma$, $W+\gamma$ and $\gamma+multijets$ are the background processes complicating this channel. Despite the monophoton does not couple strongly compares to monojet, the fact that photon only interact with electromagnetic force has given the advantage of being capable to be precisely measured from the electromagnetic calorimetry, thus reducing the experimental uncertainties.

2.9.3 Mono-V

The mono vector boson channel is looking for final state consisting an energetic vector gauge boson with a high E_T^{miss}. For example, the mono-W signature is characterized by a single lepton with high E_T^{miss} whereas for mono-Z which is reconstructed with a pair of charged lepton with high E_T^{miss}. The strength of gauge boson radiates off a quark pair initial state depends on parameter ξ which parametrized the relative strength of the coupling to down-quarks with respect to up-quarks. Compare to Monojet the mono-lepton channel is cleaner with small experimental systematic error and are likely to out-perform monojet searches with increased luminosity or pile-up (Bai & Tait, 2013). On the other hand, the Mono-W searches also provide a valuable input on how to disentangle WIMP couplings to up-type versus down-type quarks.

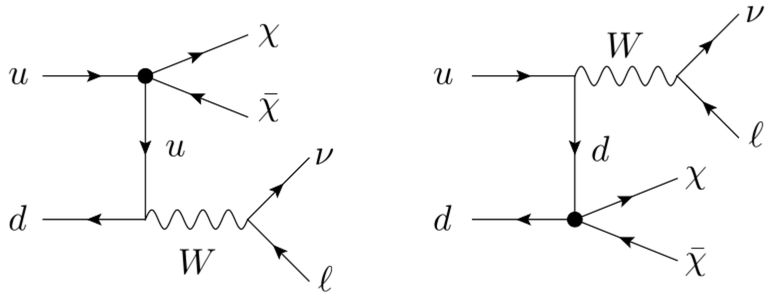

Figure 2.13: The interference of two scenarios where the W boson radiates from the u-quark or the d-quark (Bai & Tait, 2013).

2.9.4 Mono-b

The exclusive search for a b quark final state is motivated by the scalar interaction between the DM and the SM particles. The yukawa-like coupling as a consequence of Minimum Flavor Violation (MFV) has provided a direct exploitation to the signal strength by focusing on third generation quarks. Usually two channels will be used to constrain the mono-b + E_T^{miss} final state, namely the di-b final state and $t\bar{t}$ final state shows in Figure 2.14. The di-b final state is less likely to be observed in the level of analysis as the second b quark's momentum is soft, as a result only a b quark is observed. The $t\bar{t}$ decay yields a b-enriched final state, which can be exploited by b-tag the final state to reduce SM background processes (Lin et al., 2013).

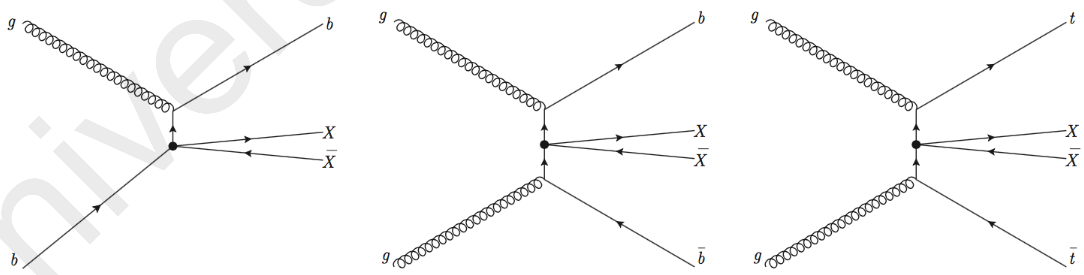

Figure 2.14: The dominant diagrams contributing to associated production of dark matter with bottom and top quarks (Lin et al., 2013).

CHAPTER 3: DATA-DRIVEN BACKGROUND ESTIMATION

3.1 Introduction

For a monojet analysis the common SM backgrounds are Z+jets and W+jets. The former has the same topology as the signal which the Z boson decays into a pair of neutrino resulted into E_T^{miss} recoiling with a jet while the latter's decayed lepton is failed to be reconstructed. Despite of the state of the art deployed in the CMS detector, it is really hard to tell whether the final state consisting E_T^{miss} is from the background or signal.

The conventional Monte Carlo (MC) generator is used for simulating the signal and the background process. However, for background modeling study it might be preferable to extract background estimate from the data themselves especially if the uncertainties on the simulation are large or the simulations do not describe the data sufficiently well. This technique used to estimate such backgrounds based on real data therefore has earned the name, data-driven background estimation. As the name suggests, the strategy is to select a control sample from the same dataset which has topological similarities with the backgrounds the analysis wishes to estimate, later the control sample is corrected in term of acceptance and efficiency to mimic the backgrounds. In this study, the $Z(\nu\bar{\nu})$ background is estimated by two different leptonic control samples, the standard candle process $Z(\mu\bar{\mu})$ event and $W(\mu\nu)$ event. Also in the context of mono b-jet channel, the study will examine the feasibility of using $W(\mu\nu)$ control sample in estimating $Z(\nu\bar{\nu})$ backgrounds by assessing the corresponding uncertainties and the number of predicted $Z(\nu\bar{\nu})$ events.

3.2 Literature Review

The $Z(\nu\bar{\nu})$ background imposes a drastic impact on the analysis. If the backgrounds are properly studied, the analysis sensitivity can be improved. A brief overview of the background estimation technique is discussed as it is indispensable for new physics searches which involving E_T^{miss} observable.

3.2.1 The Strategy of Z+jets Events Estimation

The $Z + jets$ event, refers to the production of Z boson in association with jets, is a common electro-weak process in experimental particle physics. Since the neutrino cannot be detected directly inside the detector, it requires other physics objects which will be treated as a handle to trigger such events to infer their presence in the collision. As respect to monojet signature, the $Z(\nu\bar{\nu})$ background appears as Z+jets event enters into the monojet signal region.

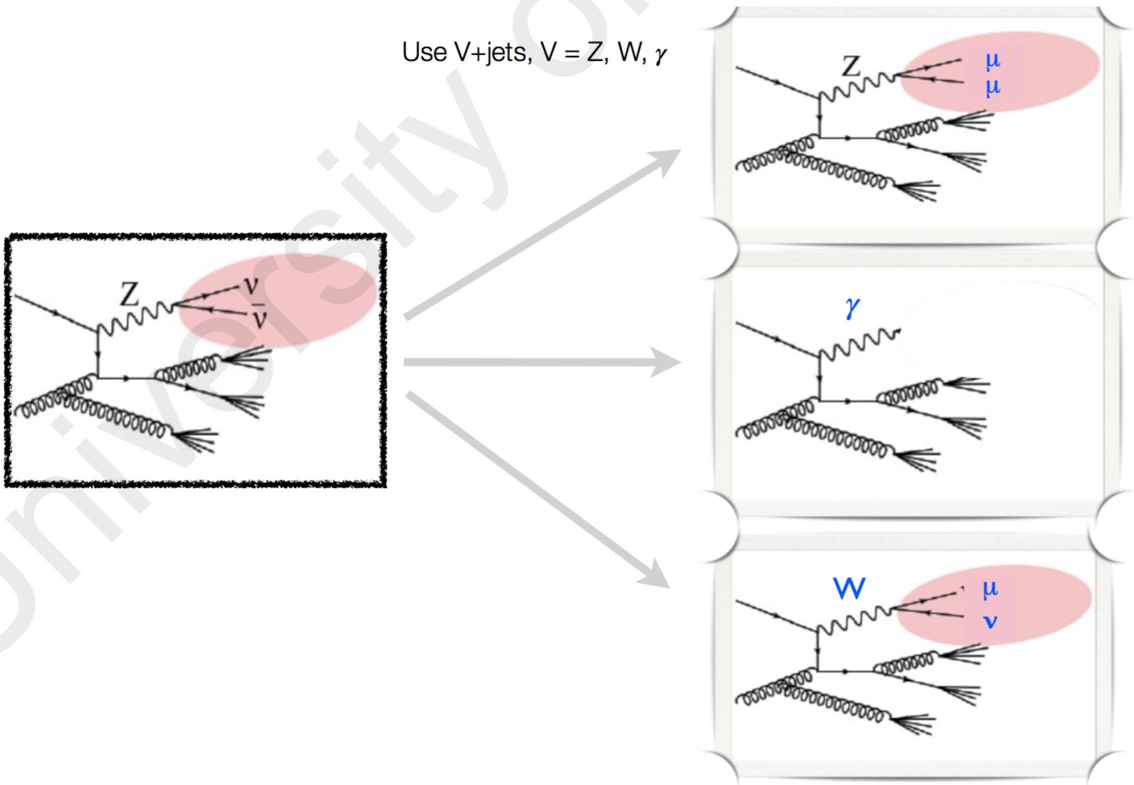

Figure 3.1: Processes with similar topology, V+jets can be used to estimate exclusively on $Z(\nu\bar{\nu}) + jets$.

The basic idea to estimate the $Z(\nu\bar{\nu})$ background from control sample originating from the same dataset is to remove the leptons in the selected control sample and recompute the total E_T^{miss} which will be used to model the $Z(\nu\bar{\nu})$ event E_T^{miss} spectrum after correcting its relative branching fractions, efficiencies, and acceptances (Malik & Watt, 2014). Since the control sample need to be topologically similar to the estimated background, the relevant processes must have a V+jets configuration where V can be W, Z and photon as illustrated in Figure 3.1.

The common control sample such as the $Z(\mu\bar{\mu})$+jets is used in most of the analysis to estimate the background relevant to new physics search such as DM (Khachatryan et al., 2015) and SUSY (CMS Collaboration, 2009a). In general, the number of $Z(\nu\bar{\nu})$ events can be predicted by the following formula (Malik & Watt, 2014):

$$N_{Z\to\nu\bar{\nu}} = N_{Z\to l^+l^-}(p_T(l^+l^-)) \times c_A \times \varepsilon \times \frac{Br(Z\to\nu\bar{\nu})}{Br(Z\to l^+l^-)} \tag{3.1}$$

where the fraction of the branching ratio $Br(Z\to\nu\bar{\nu})/Br(Z\to l^+l^-)$ can be calculated theoretically. The $l^{\pm}$ denotes the charged leptons decay from the Z boson. The correction constant c_A takes the acceptance of the detector into account and can be obtained from MC data, ε is a further correction because of the identification efficiency, which also depends on selection cuts and reconstruction efficiency. In general all events with two leptons in the final state contribute to the backgrounds for the $Z\to l^+l^-$ sample, therefore a further correction is needed to obtained a pure $Z(l^+l^-)$ sample by subtracting those relevant backgrounds, such as the W+jets, $t\bar{t}$, single t and di-boson.

However, the two competitive alternatives to the control sample $Z(\mu\bar{\mu})$+jets are the $\gamma + jets$ and W+jets. They correspond to the use of samples of events containing a high p_T photon, or W, produced with high p_T jets. The E_T^{miss} spectrum is obtained by removing the identified photon or lepton and correcting for residual differences between these events and invisible Z events. The higher statistics of these samples allows one to apply all search criteria. The two alternative control samples permit cross checks to the number of predicted $Z(\nu\bar{\nu})$+jets derived from $Z(\mu\bar{\mu})$+jets, therefore considering alternative control sample may provide a concordance approach in modeling the background process.

3.2.2 Uncertainties from Data Driven Background Estimation

It is interesting to study the systematic and statistical uncertainties of the data driven background estimation derives from different control sample which may impact the sensitivity of the favourite channel. The differing aspects among those V+jets events topology are the underlying physics properties of their final state, which defines different detection method, and thus different corresponding detector acceptance and efficiency. Attributing to the unique features of each control sample, the uncertainties incurred are of great interest to be evaluated in order to define an optimized method in background estimation. The choice of control sample is dependent on the properties of each process, the advantage and disadvantage of the three control samples are summarized below:

1. $Z(ll) + jets$. The reconstructable decay of a Z boson to dilepton is a standard candle process for many analysis. The only theoretical input is the ratio of the branching fractions for $(Z \to ll)/(Z \to \nu\bar{\nu})$ which is within 0.3% (Nakamura & Group, 2010). The dilepton events is easy to select and the sample is clear. However,this method has a large statistical uncertainty due to limited $Z(ll) + jets$ statistics mainly in the signal region.

2. $\gamma + jets$. The $\gamma + jets$ channel has a significant higher production rate than $Z(ll) + jets$ but the cost is it heavily depends on the theoretical prediction of the γ/Z cross section via gauge boson substitution (Chatrchyan et al., 2014).

3. $W(l\nu) + jets$. This channel is statistically more powerful than $Z(ll) + jets$ but usually come with non-negligible contribution from background processes such as $t\bar{t}$. It also incurs an additional systematic uncertainty from the substitution of a Z boson with a W boson, which enters in the ratio of W/Z cross sections in the regions of high p_T that are of typical searches.

Apparently the optimized background estimation technique is defined by the level of uncertainty through different control samples, fundamentally the source of each uncertainty in each analysis step needs to be understood. Those uncertainties are:

1. The statistical uncertainty in the numbers of predicted events in the data.

2. The uncertainty due to backgrounds.

3. The uncertainties in the acceptance associated with the Parton Density Function (PDF) and the size of the simulation samples.

4. The uncertainty in the selection efficiency as determined from the difference in measured efficiencies in data and simulation and the size of the simulation samples.

5. The theoretical uncertainty on the ratio of branching fractions (Nakamura & Group, 2010).

3.3 Methodology

The study will be the subset of monojet analysis (Khachatryan et al., 2015). The data was collected by CMS detector at the center of mass 8 TeV with an integrated luminosity of $19.7 fb^{-1}$. The analysis was carried out in CMSSW_5_3_11_patch6. Jets and E_T^{miss} are reconstructed using a PF technique (CMS Collaboration, 2009b) to produce unique list of particles in each event, which later used as input to the anti-k_T algorithm (Cacciari, Salam, & Soyez, 2008b) with a distance parameter of 0.5. The missing transverse energy vector is computed as the negative vector sum of the transverse momenta of all particles reconstructed in the event. In this analysis, the E_T^{miss} is calculated by excluding the muons. The standard monojet analysis code had been revised and modified to accommodate a data-driven background estimation using $W(\mu\nu)$ control sample in this study.

3.3.1 Event Selection

The standard monojet analysis deployed two triggers to collect the events, the first trigger had a E_T^{miss} threshold of 120 GeV (Khachatryan et al., 2015) , where the E_T^{miss} was computed from the calorimeter information only. On the other hand the second trigger required a PF jet with $p_T > 80$ GeV and $E_T^{miss} > 105$ GeV, the E_T^{miss} quantities for the second trigger was reconstructed using the PF algorithm with excluded muons. With this configuration, the E_T^{miss} allowed the control sample of $Z(\nu\bar{\nu})$ events used for background estimation to be collected from the same trigger as the signal sample.

In the standard monojet event selection, events were required to have a well-reconstructed primary vertex, and assumed to correspond to the hard scattering process. Instrumental and beam-related backgrounds were suppressed by rejecting events where less than 20 % of the energy of the highest p_T jet was carried by charged hadrons, or more than 70 % of this energy was carried by either neutral hadrons or photons. A jet with the highest transverse momentum (j_1) was required to have a $p_T > 110$ GeV and $|\eta| < 2.4$. Additionally, the first leading jet was being b-tagged with Combine Secondary Vertex (CSV) algorithm with medium working point (Dhingra, 2014). As signal events typically contains jets from ISR, a second jet (j_2) with $p_T > 30$ GeV and $|\eta| < 4.5$ was allowed, provided the second jet is separated from the first in azimuth (ϕ) by less than 2.5 radians, $\Delta\phi(j_1, j_2) < 2.5$. This angular requirement suppressed QCD dijet events. Events with more than two jets with $p_T > 30$ GeV and $|\eta| < 4.5$ were discarded, thereby significantly reducing background from top-quark pair and QCD multijet events. The full event selection for the data driven background estimation study is summarized in Table 3.1.

Table 3.1: The standard monojet event selection augmented with b-tagging without the lepton veto.

PFMuon + $E_T^{miss} > 200$ GeV
Noise Cleaning
Jet1 $p_T > 110$ GeV
CSV b-taging (medium WP)
Jet Multiplicity < 3
$\Delta\phi(j_1, j_2) < 2.5$
Tau veto
Control sample event selection
7 inclusive E_T^{miss} cuts

3.3.2 Control Sample Event Selection

After the monojet event selection without the lepton veto, the $Z(\mu\bar{\mu})$ events was selected by requiring at least one well identified tight muon and passing the isolated requirement or isolated muon with $p_T > 20$ GeV and Pseudorapidity (η) < 2.4. The tight muon defines as muon identified as Global Muon, and the relative combined isolation R is defined as the sum of the p_T of the charged hadrons, neutral hadrons and photon contributions computed in a cone of radius 0.4 around the lepton direction, divided by the lepton p_T. The isolation R is then resembled as (CMS Collaboration, 2012a):

$$R = \frac{\Sigma_i [p_{Ti}(\text{Charged hadron}) + p_{Ti}(\text{Neutral hadron}) + p_{Ti}(\text{Photon})]}{p_T} \tag{3.2}$$

Later, the reconstructed invariant mass, M_Z of this muon with another reconstructed muon which was identified as loose muon or Tracker Muon in the events was required to be in between 60 and 120 GeV. Additionally the pair of muons must be opposite charged towards each other.

On the other hand, the $W(\mu\nu)$ events was selected by requiring one well identified muon or tight muon with $p_T > 20$ GeV and $\eta < 2.4$, and required passing the same isolation requirement as in the di-muon case, with a reconstructed Transverse mass (M_T) between 50 and 100 GeV. The M_T variable is defined as:

$$M_T = \sqrt{2p_T^{\mu} E_T^{miss}(1 - cos\Delta\phi)} \tag{3.3}$$

where p_T^{μ} is the transverse momentum of the muon and $\Delta\phi$ is the azimuthal angle between the muon direction of flight and the negative of the sum of the p_T of all the particles reconstructed in the events. The event selection for $Z(\mu\bar{\mu})$ and $W(\mu\nu)$ control sample were summarized in Table 3.2 and 3.3.

Table 3.2: $Z(\mu\bar{\mu})$ event selection criteria in data driven background estimation.

at least one tight muon passing isolation R < 0.12
one loose muon $p_T > 20$ GeV
$\lvert\eta\rvert < 2.4$
two muons are opposite charged
60 GeV/$c^2 > M_Z >$ 120 GeV/c^2

Table 3.3: $W(\mu\nu)$ event selection criteria in data driven background estimation.

single tight muon passing isolation R < 0.12
$p_T > 20$ GeV
$\|\eta\| < 2.4$
$50\ \mathrm{GeV}/c^2 > M_T > 100\ \mathrm{GeV}/c^2$

3.3.3 Data Driven Background Estimation Analysis

The two set of event selection yield the number of observed $Z(\mu\bar{\mu})$ event and $W(\mu\nu)$ event. The data of the two observed events were then subtracted with their corresponding MC backgrounds, obtaining the data-driven $Z(\mu\bar{\mu})$ and $W(\mu\nu)$ events. In order to mimic the the pair of neutrino, a transform factor and the correction factor on the acceptance and efficiency for each control sample were weighted to account for their kinematic difference with respect to $Z(\nu\bar{\nu})$ events. Finally the estimated number of $Z(\nu\bar{\nu})$ event with their corresponding uncertainties from the two control samples was calculated expressed in 7 E_T^{miss} control regions ranging from 250 GeV to 550 GeV in step of 50 GeV. Figure 3.2 shows the flow chart on the data driven background estimation from two different control samples.

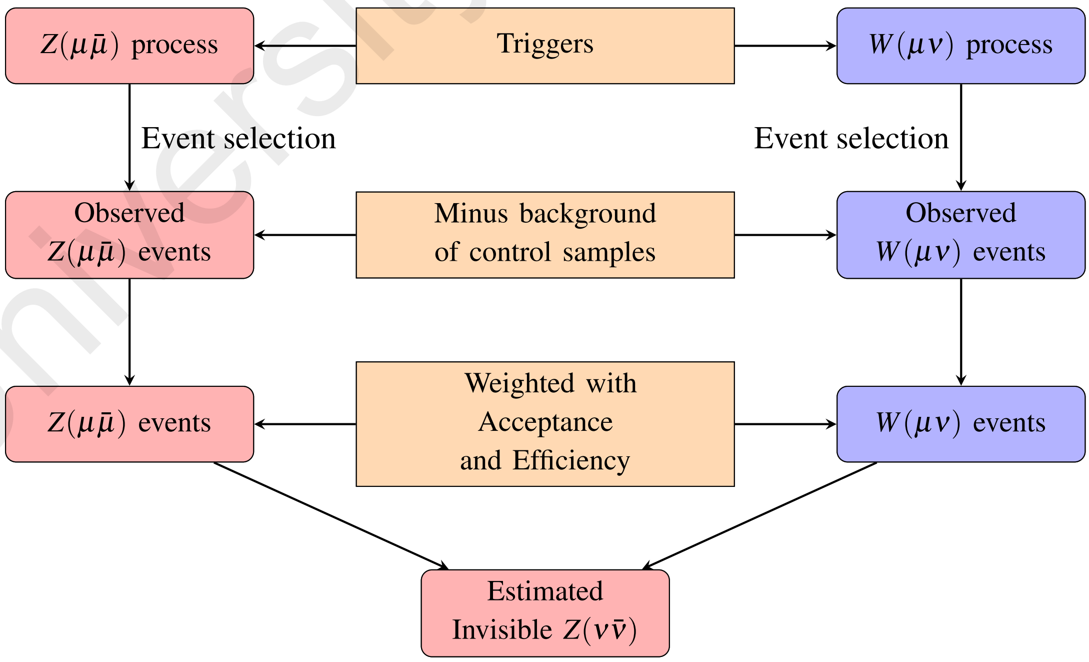

Figure 3.2: The flow chart data-driven background estimation on $Z(\nu\bar{\nu})$ events by using two different control sample in the context of b-tagged monojet event selection.

3.4 Results

3.4.1 Estimation of Z($\nu\bar{\nu}$) Background with Z+jets

The $Z(\mu\bar{\mu})$ events and $Z(\nu\bar{\nu})$ events share similar kinematic characteristics but differ only in branching ratio. By interpreting the pair of muons as E_T^{miss} during the control sample event selection, the topology of the process in which the Z boson decays to neutrinos can be reproduced. The E_T^{miss} in the $Z(\mu\bar{\mu})$ event is then used to model the $Z(\nu\bar{\nu})$ event's E_T^{miss} distribution by redefining the vector sum of the p_T of all particles excluding muons.

Table 3.4: Event yield for the Z($\mu\bar{\mu}$) data control samples and the backgrounds from MC in various E_T^{miss} cut expressed in GeV.

E_T^{miss}	W+Jets	Z+Jets	$Z(\nu\bar{\nu})$	DiBoson	$t\bar{t}$	Single t	QCD	MC	Data
250	0.0	147	0.0	13	5.5	0.5	0.0	166	185
300	0.0	62	0.0	5.8	2.6	0.0	0.0	71	85
350	0.0	29	0.0	2.5	1.8	0.0	0.0	34	34
400	0.0	10	0.0	1.6	0.0	0.0	0.0	12	14
450	0.0	6.4	0.0	0.6	0.0	0.0	0.0	7.1	4
500	0.0	2.6	0.0	0.4	0.0	0.0	0.0	3.0	2
550	0.0	1.5	0.0	0.3	0.0	0.0	0.0	1.8	0

After the control sample event selection, the number of event yield for SM background processes and the data are shown in Table 3.4. The number of $Z(\nu\bar{\nu})$ events can be predicted by using the formula below:

$$N(Z(\nu\bar{\nu})) = \frac{Nobs_{Z(\mu\bar{\mu})} - Nbgd_{Z(\mu\bar{\mu})}}{Acc \times Eff} \times \frac{Br(Z(\nu\bar{\nu}))}{Br(Z(\mu\bar{\mu}))} \tag{3.4}$$

The distribution of $Z(\nu\bar{\nu})$ event is estimated from the observed control sample consisting of dimuons, $Nobs_{Z(\mu\bar{\mu})}$ after correcting the estimated background in the dimuon sample, $Nbgd_{Z(\mu\bar{\mu})}$; differences in muon acceptance, Acc and efficiency, Eff with respect to neutrinos; and the ratio of branching fractions for the Z decay to a pair of neutrinos, and to a pair of muons, $Br(Z(\nu\bar{\nu}))/Br(Z(\mu\bar{\mu})$. The value for the branching ratio was taken from Particle Data Group (Olive et al., 2014) which is 5.942±0.019.

The acceptance Acc is defined as the fraction of all generated event where the generated muons are reconstructed with $p_T > 20$ GeV and $|\eta| < 2.1$ and requires an invariant mass within the Z boson mass window. The 0.5 factor is due to each event is required to have 2 muons.

$$\text{Acc} = \frac{0.5 \times \text{All gen stable } \mu}{0.5 \times (\text{gen stable } \mu), \, (p_T(\mu) > 20 \text{ GeV}, |\eta| < 2.1), \, (60 \text{ GeV} < M_Z < 120 \text{ GeV})} \tag{3.5}$$

On the other hand, the event selection efficiency Eff is defined as the efficiency of reconstructing a pair of opposite charge PF muons passing all the identification and isolation criteria and with an invariant mass between 60 and 120 GeV, given that they are within the detector acceptance.

$$\text{Eff} = \frac{0.5 \times (\text{gen stable } \mu \,), \, (p_T(\mu) > 20 \text{ GeV}, |\eta| < 2.1), \, (60 \text{ GeV} < M_Z < 120 \text{ GeV})}{(\text{ PF } (\mu^+\mu^-) \,), \, (\text{ } 60 \text{ GeV} < M_Z < 120 \text{ GeV})} \times \frac{1}{SF_{\mu\mu}} \tag{3.6}$$

It was corrected by factor $SF_{\mu\mu}$ = 1.092 (Khachatryan et al., 2015) to account for differences in the measured muon reconstruction efficiency in data and simulation, measured by the muon Physics Object Group (POG) (CMS Collaboration, 2012a). All The value used in Acc and Eff calculation was obtained from Z+jets MC using generator level information.

Table 3.5: Estimated number of $Z(\nu\bar{\nu})$ events from $Z(\mu\bar{\mu})$ control sample in various E_T^{miss} cut expressed in GeV.

E_T^{miss}	> 250	> 300	> 350	> 400	> 450	> 500	> 550
Nobs	185	85	34	14	4.0	2.0	0.0
Nbgd	19	8.4	4.3	1.6	0.6	0.4	0.3
Acc	0.896	0.926	0.917	0.923	0.962	0.961	0.933
Eff	0.754	0.747	0.729	0.643	0.733	0.586	0.589
$N(Z(\nu\bar{\nu}))$	1462±159	658±96	264±59	124±43	28±18	16±16	-3.7±-2.3

In Table 3.5 shows the observed $Z(\mu\bar{\mu})$ control data event yield and its corresponding MC background processes, together with the correction factors for various E_T^{miss} cuts used in Equation 3.4. Finally the number of predicted $Z(\nu\bar{\nu})$ event was obtained after subtracting the backgrounds and weighted with the correction. A summary of the fractional contributions of the uncertainties to the total error on the $Z(\nu\bar{\nu})$ background is shown in Table 3.6. The uncertainty on the $Z(\nu\bar{\nu})$ background estimation was due to statistical uncertainty in the observed number of events; uncertainty in the number of background events, 50% uncertainty assigned to each MC background, added in quadrature; and uncertainty in the acceptance and efficiency. In addition to the statistical uncertainty, 2 % error in the PDF was absorbed into the acceptance uncertainty (Chatrchyan et al., 2014), and a 2% error due to hadronization was absorbed into the efficiency uncertainties.

Table 3.6: Systematic Uncertainty (%) of $Z(\nu\bar{\nu})$ prediction from $Z(\mu\bar{\mu})$ control sample, error in PDF (2%) and hadronization (2%) are absorbed to acceptance, and efficiency uncertainty term, respectively.

E_T^{miss} (GeV)	> 250	> 300	> 350	> 400	> 450	> 500	> 550
Nobs error	8.2	12	20	30	60	91	-0.0
Nbgd error	4.2	4.1	5.2	6.5	9.5	14	-50.0
Acc error	2.4	2.6	3.3	4.5	4.3	5.8	9.7
Eff error	3.1	4.2	5.9	11	12	23	31
Data/MC S.F	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Total $Z(\nu\bar{\nu})$ error	11	15	22	35	62	96	64

3.4.2 Estimation of Z($\nu\bar{\nu}$) Background with W+jets

Alternatively, the $Z(\nu\bar{\nu})$ event can be predicted by $W(\mu\nu)$ event, where the E_T^{miss} in $W(\mu\nu)$ event is defined as the vector sum of the muon p_T and E_T^{miss} to emulate the missing energy in events.

Table 3.7: Event yield for the W($\mu\nu$) data control samples and the backgrounds from MC in various E_T^{miss} cut expressed in GeV.

E_T^{miss}	W+Jets	Z+Jets	$Z(\nu\bar{\nu})$	DiBoson	$t\bar{t}$	Single t	QCD	MC	Data
250	498	7.1	0.0	36	196	50	0.0	787	910
300	226	1.5	0.0	18	63	22	0.0	331	355
350	105	1.2	0.0	9.6	23	13	0.0	151	170
400	52	0.0	0.0	3.5	8.4	3.6	0.0	68	77
450	28	0.0	0.0	1.7	3.4	2.9	0.0	36	43
500	16	0.0	0.0	1.0	2.2	2.2	0.0	21	21
550	11	0.0	0.0	0.4	0.6	1.3	0.0	13	6

After the control sample event selection, the number of event yield for SM background processes and the data are shown in Table 3.7. Similarly the number of $Z(\nu\bar{\nu})$ event can be predicted by using Equation 3.7:

$$N(Z(\nu\bar{\nu})) = \frac{Nobs_{W(\mu\nu)} - Nbdg_{W(\mu\nu)}}{Acc' \times Eff'} \times \frac{Br(Z(\mu\bar{\mu}))}{Br(W(\mu\nu))} \times \frac{Br(Z(\nu\bar{\nu}))}{Br(Z(\mu\bar{\mu}))} \tag{3.7}$$

Here, $Nobs_{W_{\mu\nu}}$ is the total number of observed single muon control data events. Within these events there are small backgrounds, $Nbdg_{W(\mu\nu)}$ which were simulated in SM. The $W(\mu\nu)$ event is then obtained by subtracted off with $Nbdg_{W(\mu\nu)}$ and weighted appropriately with correction factor, the acceptance and efficiency respectively.

Later, the control sample was corrected for the detector acceptance Acc'. The Acc' defined as the selection of reconstructed event with well identified and isolated single muon with $p_T > 20$ GeV and $|\eta| < 2.1$ or $p_T > 10$ GeV if the single muon originates from hadronically decayed τ over all the generated stable muon.

$$Acc' = \frac{(\text{ stable } \mu, p_T(\mu) > \text{ 20 GeV }, |\eta| < 2.1\text{ }), (\mu \text{ from } \tau, p_T(\tau) > \text{ 10 GeV })}{\text{All generated stable } \mu + \text{ All } \mu \text{ from } \tau} \tag{3.8}$$

On the other hand, the efficiency Eff' for selecting for the single muon selection is defined as the number of selected reconstructed events in the W mass window given they are within the acceptance, including a data-MC normalization factor of 0.957 (Khachatryan et al., 2015). These values were taken from generator level MC.

$$Eff' = \frac{\text{Tight ID stable } \mu \text{ in W mass window } + \text{ Tight ID } \mu \text{ from } \tau \text{ in W mass window}}{\text{stable } \mu, p_T(\mu) > 20 \text{ GeV }, |\eta| < 2.1 + \mu \text{ from } \tau, p_T(\tau) > \text{ 10 GeV}} \times 0.957 \tag{3.9}$$

Note that an extra term arose which is the branching ratio of $Br(Z(\mu\bar{\mu}))/Br(W(\mu\nu))$, and empirically given as:

$$\frac{Br(Z(\mu\bar{\mu}))}{Br(W(\mu\nu))} \sim \frac{\text{Number of stable generated di-muon}}{\text{Number of stable generated single muon}} \tag{3.10}$$

The term is approximated as the number of all stable generated dimuon from Z+jets MC event divided by the number of all single generated muon from W+jets MC event. Both MC generator level information was used in the approximation. Also, all the value used in Acc' and Eff' calculation was obtained from W+jets MC using generator level information. As usual the third term which is the $Br(Z(\nu\bar{\nu}))/Br(Z(\mu\bar{\mu}))$ is required to correct the kinematics differences between $Z(\nu\bar{\nu})$ and $W(\mu\nu)$ event. Their value was taken from the Particle Data Group (Olive et al., 2014), which is 5.941 ± 0.019.

Table 3.8: Estimated number of $Z(\nu\bar{\nu})$ events from $W(\mu\nu)$ control sample in various E_T^{miss} cut expressed in GeV.

E_T^{miss}	>250	>300	>350	>400	>450	>500	>550
$Nobs_{W(\mu\nu)}$	910	355	170	77	43	21	6.0
$Nbgd_{W(\mu\nu)}$	288	105	47	16	7.9	5.3	2.3
$Br(Z(\mu\bar{\mu}))$	199	83	40	16	8.3	4.2	2.4
$Br(W(\mu\nu))$	1410	618	285	144	72	41	28
Acc'	0.886	0.902	0.906	0.893	0.929	0.964	0.985
Eff'	0.382	0.388	0.388	0.389	0.399	0.384	0.375
$N(Z(\nu\bar{\nu}))$	1542±278	567±96	293±51	113±22	66±15	26±9.0	5.2±3.7

The Table 3.8 shows the observed $W(\mu\nu)$ control data sample event yield and its corresponding MC background processes, together with the correction factors for various E_T^{miss} cuts used in Equation 3.7. Finally the predicted number of $Z(\nu\bar{\nu})$ event was calculated by subtracting the background and weighted with the correction factor. A summary of the fractional contributions of the uncertainties to the total error on the Z($\nu\bar{\nu}$) background was shown in Table 3.9. The uncertainty on the $Z(\nu\bar{\nu})$ background estimation was due to errors in the observed number of events; errors in the number of background events, 50% uncertainty assigned to background that is estimated using MC, and added in quadrature; and errors in the acceptance and efficiency. Again, 2 % error in PDF absorbed into acceptance uncertainty (Chatrchyan et al., 2014) and 2 % error due to hadronization was accounted into the efficiency uncertainty.

Table 3.9: Systematic Uncertainty (%) of $Z(\nu\bar{\nu})$ prediction from $W(\mu\nu)$ control sample, error in PDF (2%) and hadronization (2%) are absorbed to acceptance, and efficiency uncertainty term, respectively.

E_T^{miss} (GeV)	> 250	> 300	> 350	> 400	> 450	> 500	> 550
Nobs error	4.9	7.5	11	14	19	29	66
Nbgd error	17	14	11	8.0	6.8	10	20
Acc' error	2.1	2.1	2.3	2.6	2.8	2.7	2.4
Eff' error	2.9	3.7	5.1	6.9	9.2	12	15
Data/MC S.F	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Total $Z(\nu\bar{\nu})$ error	18	17	17	19	23	34	71

3.5 Discussion

After the $Z(\mu\bar{\mu})$ control sample event selection, throughout the E_T^{miss} region the number of observed control data event is diminished at a faster rate. On the other hand, the background processes towards the control region is favorable small. The angular cut between the leading jet and sub-leading jet is effective in removing all QCD background. The dominant background to the $Z(\mu\bar{\mu})$ control sample is the Di-boson process, which constitute ~ 68 % of total backgrounds, follow by the $t\bar{t}$ event where the E_T^{miss} contribution is come from the semi-leptonic decay of top quark, lastly traces of single t event. A comparison between data and MC for the dimuon invariant mass and momentum after all the selection cuts and a E_T^{miss} cut of 250 GeV is shown in Figure 3.3.

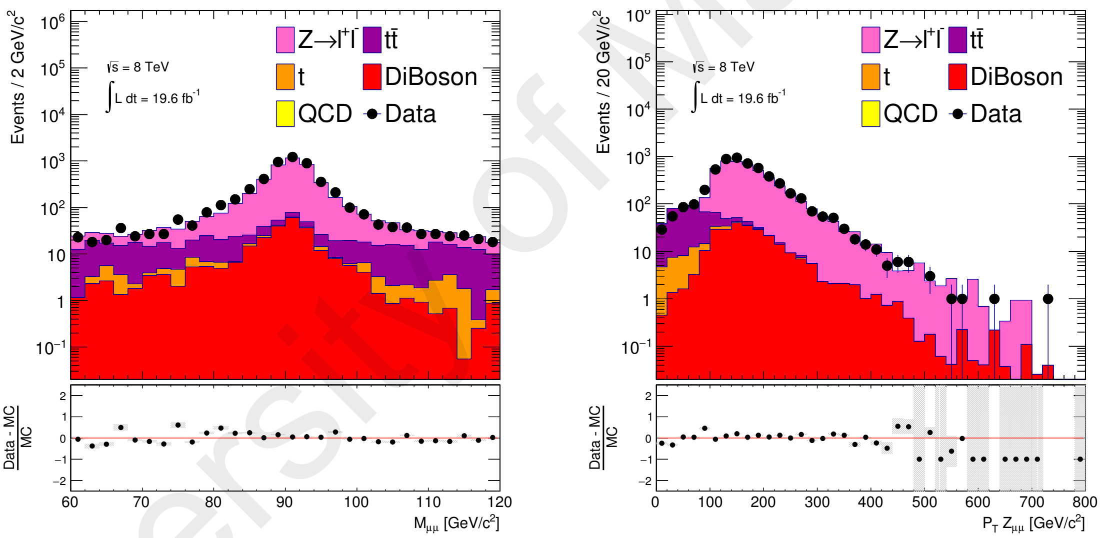

Figure 3.3: Invariant mass and transverse momentum of the di-muon pair in the $Z(\mu\bar{\mu})$ control sample.

In Table 3.5, the predicted number of $Z(\nu\bar{\nu})$ event is reasonable with 11 % uncertainty in lower E_T^{miss} region. However, higher uncertainty is observed while progressing to higher E_T^{miss} region due to the limited number of observed control data event. In order to understand further the origin of the uncertainty, a breakdown of the total uncertainty is shown in Table 3.6 has once again confirmed that the higher error incurred is statistical. It is concluded that the E_T^{miss} is cutting too hard on the $Z(\mu\bar{\mu})$ control data sample to the extend that it almost reduce most of the control data event, for example there is only 4 events left at $E_T^{miss} > 450$ GeV.

On the contrary, the number of $W(\mu\nu)$ control data sample event has more statistics compares to the $Z(\mu\bar{\mu})$ control data sample. The observed control data sample has higher number of event before $E_T^{miss} > 350$ GeV, subsequently the number of event is reduced moderately after the E_T^{miss} threshold due to higher E_T^{miss} cut. Also, the angular cut between the leading jet and sub-leading jet is effective in removing all QCD background. The dominant background for the $W(\mu\nu)$ event is the $t\bar{t}$ event which constitutes ~ 68 % of the total background, follow by single t event, Di-boson and lastly traces of Z+jets event which decay leptonically. A comparison between data and MC for the transverse mass and momentum of the W after the full selection and a E_T^{miss} cut of 250 GeV is shown in Figure 3.4, where the missing transverse energy is defined as the vector sum of the muon transverse momentum and missing transverse energy to emulate the missing energy in events where the lepton is lost.

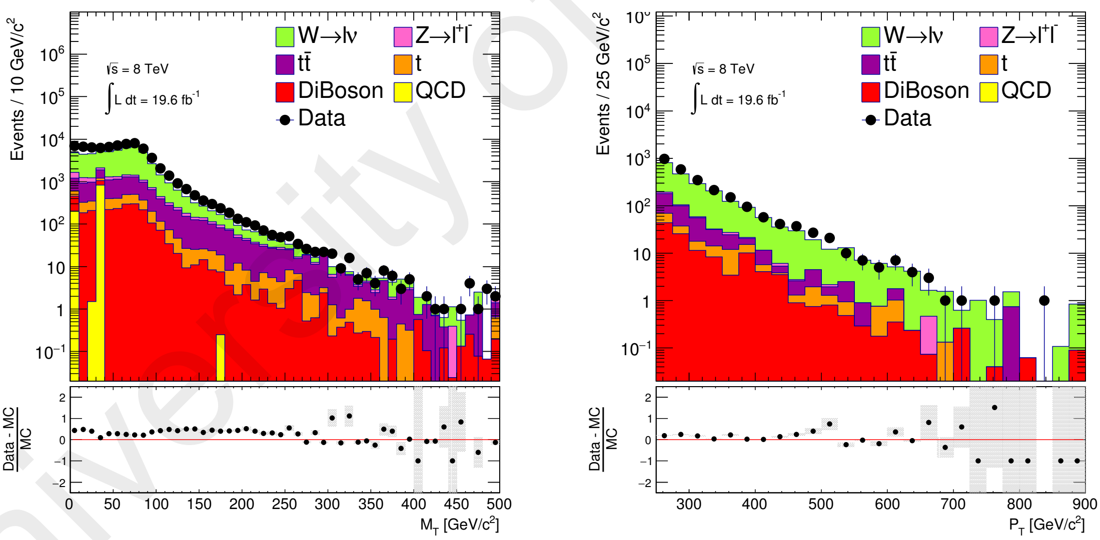

Figure 3.4: Transverse mass and transverse momentum of $W^{\pm}$ candidates in the single-muon sample.

Referring to Table 3.8, at low E_T^{miss} region the predicted $Z(\nu\bar{\nu})$ event has higher event compares to $Z(\mu\bar{\mu})$ event's based estimation, but with higher uncertainty which is 18 %. As the E_T^{miss} cuts increase, appreciable number of predicted $Z(\nu\bar{\nu})$ event is observed. For example at the higher E_T^{miss} threshold such as $E_T^{miss} > 500$ GeV, 26 events left compares to $Z(\mu\bar{\mu})$ control data event's which is 16 events left, correspondingly with total uncertainty of 34 % compares to 96 %. If we turn to the fractional uncertainty of the data-driven estimation as shown in Table 3.9, the source of the major uncertainty is contributed from the error due to the number of observed background events. As expected the choice of using $W(\mu\nu)$ event as a control sample as it yield more event has come with the cost of higher background, for this case the dominant background being $t\bar{t}$ process which decay into W boson and a b quark.

The related uncertainty for efficiency and acceptance are due to the flaw in detector simulation and the finite η coverage inside the detector. From both data-driven background estimation approach, improvement is encouraging especially at higher E_T^{miss}. The remaining uncertainty is the theoretical uncertainty, describing the limited precision of background process cross section as each of them is computed separately and thus have different uncertainty.

3.6 Conclusion

The summary of the data-driven background estimation for both control sample is shown in Table 3.10 and the total uncertainty for each control sample is shown in Table 3.6 and Table 3.9 respectively.

Table 3.10: Number of Z($\nu\bar{\nu}$) events estimated from different choices of control sample compares to MC simulated $Z(\nu\bar{\nu})$ events (first row) in various E_T^{miss} region expressed in GeV.

E_T^{miss}	> 250	> 300	> 350	> 400	> 450	> 500	> 550
$Z(\nu\bar{\nu})$	1080±540	441±221	200±100	96±48	51±26	28±14	17±8.5
$Z(\mu\bar{\mu})$	1462±159	658±96	264±59	124±43	28±18	16±16	-3.7±-2.3
$W(\mu\nu)$	1542±278	567±96	293±51	113±22	66±15	26±9.0	5.2±3.7

The alternative $W(\mu\nu)$ control sample was explored in data-driven background estimation, the predicted number of $Z(\nu\bar{\nu})$ events and the corresponding total uncertainty were investigated and compared with $Z(\mu\bar{\mu})$ control sample. Both prediction revealed an increase in the number of $Z(\nu\bar{\nu})$ events with competing uncertainties. The $Z(\mu\bar{\mu})$ control sample yields a favourable prediction at lower E_T^{miss} region accompanied with tolerable total uncertainty; while $W(\mu\nu)$ control sample is favourable on predicting $Z(\nu\bar{\nu})$ events at high E_T^{miss} region with appreciable higher number of predicted background event with tolerable total uncertainty. Both predicted number of background event has consistently reported higher number of event compare to MC simulation. The summary of the data-driven background estimation for both control sample is shown in Table 3.10 and the total uncertainty for each control sample is shown in Table 3.6 and Table 3.9 respectively.

The $Z(\mu\bar{\mu})$ control sample is historically favourable in data-driven background estimation due to the striking feature such as the dimuon event is clean and easily selected, but it is also known that it has limited size of event giving rise to high statistical uncertainty. On the other hand, the $W(\mu\nu)$ events is a popular control sample in estimating lost W+jets background. Besides of the kinematics similarity, it has larger event size but comes with a cost of higher background event, giving rise to higher uncertainty too. Based on the finding of the study, it affirms their strength and weakness. In the spirit of improving the sensitivity of a mono b-jet channel, the strategic method to improve background modeling such as $Z(\nu\bar{\nu})$ background is to optimize the application of both control samples in data-driven background estimation to describe its E_T^{miss} spectrum. For example, the $Z(\mu\bar{\mu})$ control sample is used to model the low E_T^{miss} region and the $W(\mu\nu)$ control sample used to model the high E_T^{miss} region of the $Z(\nu\bar{\nu})$ event's E_T^{miss} spectrum.

CHAPTER 4: SENSITIVITY STUDY ON Mono b-jet + MET FINAL STATE

4.1 Introduction

DM particle has been proposed to explain numerous astrophysical measurements as such the rotational curves of galaxies and gravitational lensing. Popular model of dark matter particle proposes the existence of non-relativistic particle that interacts weakly with SM particles. The model is consistent with DM thermal relic abundance if the WIMP have weak-scale mass provided if their interaction cross section with baryonic matter is of the order of electroweak cross sections (Feng, 2010). Additionally some new physics scenarios are postulated to explain the hierarchy problem also predict the existence of WIMP (Farrar & Fayet, 1978).

New physics events are usually interpreted as large E_T^{miss} quantity in the final state. Such event might contain the undetected DM particled produced with visible SM particle. There will not be a discernible signal in the CMS detector since the WIMP will not interact with the detector components. Like neutrinos, their existence can only be inferred from an imbalance of the total momentum from the reconstructed particles in the plane transverse to the beam axis. The monojet signature can be used to search for the pair production of WIMP in association with a hadronic jet, which is used to tag or trigger the such important event.

The heavy flavour quark such as bottom and top quark production in association with DM pair production can be an interesting channel compares to other physics object due to the enhancing effect on the coupling permitted by MFV (D'Ambrosio, Giudice, Isidori, & Strumia, 2002). If the WIMP Miracle is true, by interpreting the mono b-quark + E_T^{miss} final state it may provide a significant signal in LHC. Since the astrophysical evident suggested that DM is heavy due to its non-relativistic velocity (Feng, 2010), the model independent approach is suitable to describe the interaction between DM and SM particle. This study is focused on studying the sensitivity of mono b-jet channel assuming a scalar interaction between a pair of Dirac fermion DM particle and a b quark. The sensitivity of such channel is expressed in term of M_* and in term of nucleon-DM elastic scattering cross section as a function of DM mass.

4.2 Literature Review

The search for monojet signature is a powerful way to place model independent constraints on effective operators describing the coupling between DM and SM particles. The operators generated by the exchange of a scalar mediator, however, coupling to light quarks will suppress the interaction strength and thus the prospect of probing such interactions through the inclusive monojet channel at the LHC is limited.

4.2.1 Dark Matter Association Production With Heavy Flavour

The mono b-quark signature can improve the current limit as the signal arises partly from direct production of b-quarks in association with DM particles, but the dominant component is from top quark pair production in the kinematic regime where one top quark is boosted (Lin et al., 2013). From the various operators proposed within the EFT (Goodman et al., 2010), the scalar operator is particularly challenging to be constrained where interaction between DM and quarks is mediated by a heavy scalar mediator given as:

$$\mathscr{O} = \frac{m_q}{M_*^3}\bar{q}q\bar{\chi}\chi \tag{4.1}$$

summing over all quarks. The M_* characterizes the coupling between SM particle and DM particle, and the m_q is the mass of the coupled quark. The form of the interaction is fixed by MFV (D'Ambrosio et al., 2002). Scalar interactions with SM quarks are typically strongly constrained by flavour changing neutral current measurements, but in MFV these dangerous flavour violating effects are automatically suppressed. Due to the fact that the interactions are proportional to quark mass, the monojet + E_T^{miss} signal rate is suppressed due to the light quark masses.

The Feynman diagram for the heavy flavour production in association with DM pair production is shown in Figure 2.14. Particularly the direct b production occurs through b quark and gluon-initiated processes such as b g $\rightarrow \bar{\chi}\chi+$ b. In comparison to the light quark initial states, these processes are suppressed by the b-quark parton density. However, the enhancement due to the MFV form for the coupling is suffice to compensate this. Furthermore, g g $\rightarrow \bar{\chi}\chi + t\bar{t}$ seem likely to be the dominant contribution to the monojet signal. Thus, the final states are highly b-enriched (Lin et al., 2013). At the same time, by focusing on exclusive b-tagged final state reduces the SM backgrounds significantly. Therefore, an improvement in the LHC reach for the scalar operator is foreseen by requiring a b-tagged monojet.

As showed in Figure 4.1, the constraints on M_* derived from monojet channel (blue curve) is less competitive compares to the mono b-jet channel. On the other hand by requiring an exclusive b-tagged monojet (red dotted curve), the constraints is improved $\sim$ 10 compares to monojet channel. While more stringent constraint is observed if the $t\bar{t}$+ E_T^{miss} is considered together with $b\bar{b}$+ E_T^{miss} and mono b + E_T^{miss}. The same observation is seen on the direct detection's phase space.

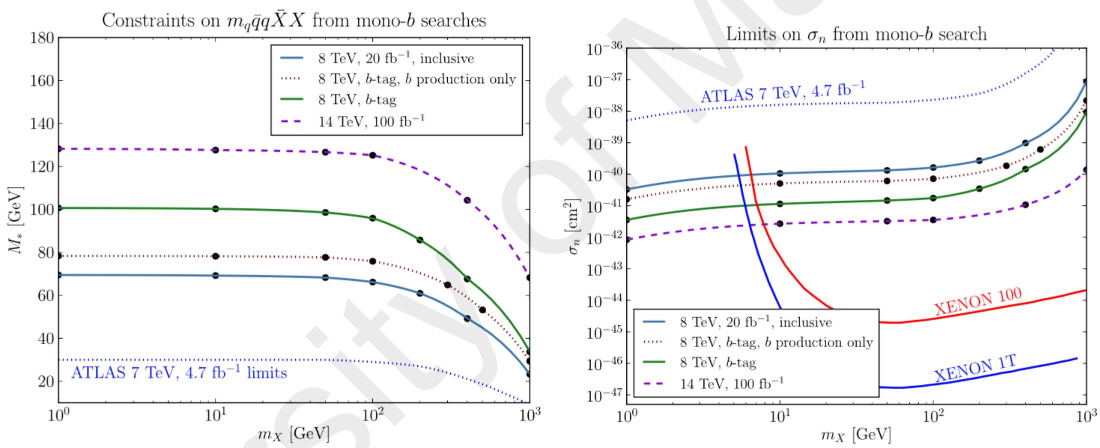

Figure 4.1: (left) The Expected 90 % Confident Level (CL) limits on D1 operator from mono b-quark search (Lin et al., 2013); (right) The corresponding constraints on the spin-independent nucleon scattering cross section, along with direct detection limits (Aprile et al., 2012; Aprile & XENON1T collaboration, 2012).

4.2.2 b-tagging Technique in CMS Experiment

The identification of jets originating from b quark is crucial in both searches for new physics and for the measurement of standard model processes. Jets originating from the hadronization of bottom quark appear in many important physics processes, such as decays of top quark, Higgs bosons, and many new particles predicted by supersymmetric models. The identification of b-jet is, therefore, a key ingredient in reducing overwhelming backgrounds to these channels from processes involving jets from gluons, light flavour quarks, and c quark fragmentation. In the high energy physics terminology, b-quark identification is often referred to b-tagging.

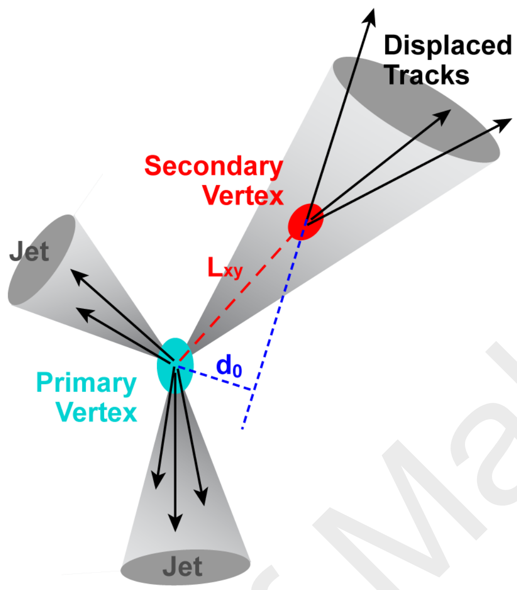

Figure 4.2: Schematic view of identifying a b-jet in the b-tagging algorithm (Abazov et al., 2009)

The b-tagging relies on the special decay of b hadrons, B^- or B^+. Owing to flavour conservation in the strong and electromagnetic force, these particles can only decay via the weak force. But these decays into an up or charm quark are suppressed in the CKM-Matrix. Therefore, they have a very long lifetime compared to other particles with higher mass. Moving almost at the speed of light, they can travel a distance that can be up to the order of cm as showed in Figure 4.2. The decay of the b hadron then leads to a secondary vertex that can be observed with the pixel detector.

A variety of b-tagging algorithm has been developed in the matter to improve the mis-identification and the tagging efficiency by exploiting the properties of b quark. The variant of b-tagging algorithm can be find here (Dhingra, 2014). For this analysis, the CSV algorithm will be used. The CSV algorithm combines secondary vertex and displaced track information to build a likelihood-based discriminator to distinguish between jets from b quarks and those from charm or light quarks and gluons. The minimum thresholds on these discriminators define loose, medium, and tight operating points with a mis-identification probability for light parton jets of close to 10%, 1%, and 0.1%, respectively (Dhingra, 2014).

4.2.3 Event Generation

In order to study a signal from SM processes or extract a signal of new physics from the SM backgrounds, one needs to generate and simulate the signal events similar to what is expected in real data. At high energy colliders like LHC, different issues make this procedure challenging. In each hard interactions hundreds of SM or Beyond SM particles can be produced with momenta range over many orders of magnitude. The calculation of Matrix Element (ME) is too laborious at higher orders of perturbation theory. At low energies, all soft hadronic phenomena (like hadronization and the underlying event) must rely upon QCD inspired models and cannot be computed from first principles. Many divergences and near divergences issues should be addressed after calculation of ME. Finally, the ME must be integrated over a final state phase space with huge dimensions in order to obtain predictions of experimental observables (Buckley et al., 2011).

There is a very broad spectrum of event generators from general purpose ones to ME generators. The general purpose MC event generators such as HERWIG (Corcella et al., 2001), Pythia (Sjöstrand, Mrenna, & Skands, 2006) and Sherpa (Gleisberg et al., 2004) provide a comprehensive list of Leading Order (LO) ME of the SM and some Beyond SM processes. In addition to the LO ME, multi-purpose MC generators contain theory and models for a number of physics aspects, such as hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation and decay. In order to compute the hard process ME at higher order and cope with arbitrary final state, ME generators have therefore been constructed. Parton level events generated by the ME generators are processed by general purpose event generators to do the remained steps. The most widely used ME generators in CMS are ALPGEN (Mangano, Piccinini, Polosa, Moretti, & Pittau, 2003), POWHEG (Alioli, Nason, Oleari, & Re, 2010) and MADGRAPH (Maltoni & Stelzer, 2003).

4.3 Methodology

The sensitivity study was conducted by simulating the signal process of direct b-quark production in association with a pair of DM production in the EFT framework. A MLM jet matching description (Mangano, Moretti, Piccinini, & Treccani, 2007) was implemented on the simulated samples since the samples contained a ISR or FSR; On the other hand, the monojet standard analysis was reiterated with additional b-tagging cut and performing the cut-and-count analysis on the set of MC simulated monojet SM backgrounds and on the collected CMS data. Dominant background processes such as $Z(\nu\bar{\nu})$ event and lost W+jets event were estimated by a set of control data sample collected by the same set of monojet triggers used in collecting CMS data. Later a 90% CL limit was derived by interpreting the mono b-jet channel on the M_*. Owing to the benefit of EFT which is model independent, the collider limit will be translated into the Nucleon-DM scattering cross section ($\sigma_{\chi-nucleon}$) by correcting the relevant kinematics. The series of analysis involved for the sensitivity study is summarized in flow chart shows in Figure 4.3, each step will be elaborated in the following section.

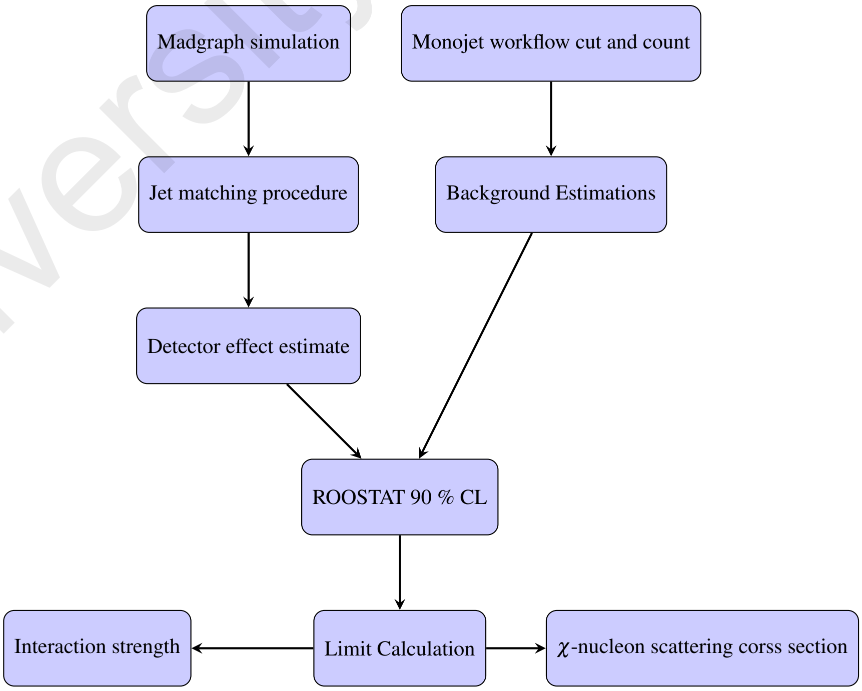

Figure 4.3: The flow chart of the sensitivity study on mono-b + E_T^{miss} channel within the standard monojet framework.

4.3.1 Madgraph 5 Generation

The signal generation was carried out in MADGRAPH5_aMC@NLO (Alwall, Herquet, Maltoni, Mattelaer, & Stelzer, 2011), with version MG5_aMC_v2_1_0. The mono b-quark with a pair of DM was simulated with the EffDM MADGRAPH model (Lin et al., 2013) at LO computation for the scalar operator D1. The EffDM model describes the interaction between the Dirac fermion DM with quarks or gluons in term of EFT, allowing one to capture a wide class of theories of DM in which the particles mediating the interaction are somewhat heavier than the energies of interest. The model is suitable for the study particularly it allows one to map these interactions from collider observable into the parameter space of direct detection of DM. The EffDM model requires a threshold on the interaction scale which is referred as M_*. In this study, the M_* is given the value of 1000 GeV and a total of 9 different DM mass points M_χ were simulated.

During the user interface session, the SM MAGRAPH model file is pre-loaded. After appropriately imported the model file, the initial proton state was redefined to include b quark in the proton definition. By this definition the b quark is assumed massless as it is originating from the PDF, enabling better prediction of cross section and modeling for the second b quark as it usually carry small momentum. A b-jet is defined as a b quark or anti b quark to emulate the b-tagging algorithm. After that, the generated signal required at least one b quark in the final state and a pair of DM ($\chi\chi\sim$), together with D1=1 indicating a scalar interaction. A second jet was allowed to retain events with ISR, where the second jet can either be a b-jet, light quarks or gluon jet. The command defined the simulation in MADGRAPH is summarized in Figure 4.4.

```
import model EffDM_UFO -modelname
define p = g u c d s u~ c~ d~ s~ b b~
define j = g u c d s u~ c~ d~ s~ b b~
define bjet = b b~
generate p p > chi chi~ bjet QED=0 D1=1 D2=0 D3=0 D4=0 D5=0 D6=0
D7=0 D8=0 D9=0 D10=0 D11=0 D12=0 D13=0 D14=0
add process p p > chi chi~ bjet j QED=0 D1=1 D2=0 D3=0 D4=0 D5=0
D6=0 D7=0 D8=0 D9=0 D10=0 D11=0 D12=0 D13=0 D14=0
```

Figure 4.4: The MADGRAPH command for mono b-jet + E_T^{miss} final state simulation.

The generated signal events were imposed a series of generator cuts showed in Table 4.1. The M_* and M_χ were set in the param_card.dat file supplied together with the model file. The generator kinematics cuts is defined in run_card.dat file. The minimum p_T of the leading jet (ptj1) is set above 80 GeV, and the minimum p_T of the sub-leading jet (ptj2) set above 20 GeV. Additionally a jet matching cut (xqcut) of 20 GeV was implemented at parton level. The jet matching cut is the measure of required parton separation at ME calculation level. Since the b quark is assumed to be massless, the maximum flavour scheme (Maxjetflavor) to be considered as a jet was assigned to 5 in order to consider the b quark in jet clustering step.

Table 4.1: Parameter used to generate the signal events this study.

M_* = 1000 GeV
M_χ = [0.1, 1, 10, 100, 200, 300, 400, 700, 100] GeV
ptj1 $>$ 80 GeV
ptj2 $>$ 20 GeV
xqcut = 20 GeV
Maxjetflavor = 5

At the end of the events generation, a Les Houches Event Files (LHE) which encoded the simulated events were produced. By using ROOT (Brun & Rademakers, 1997), a ROOT macro was written to ntuple the LHE file into ROOT format for ease of being read in the study. Subsequently another analysis-based ROOT macro was written to extract the relevant kinematics information of the generated signal event and the detector acceptance was calculated with consideration of 50% Hadronization and 70 % b-Tag efficiency for accounting detector effect and the b-tagging algorithm defect. The analysis was carried out in 7 different E_T^{miss} regions similar to previous study in Chapter 3. Lastly a bash script was written to loop over a set of 9 samples with different DM mass point on the analysis-base ROOT macro.

Since the study is intended to produce a preliminary result in order to evaluate the sensitivity of the mono-b + E_T^{miss} channel, a parton-level LHE file will be used in the subsequent step throughout the analysis. There the MLM-jet matching is only implemented in ME generator step.

4.3.2 Mono b-jet Event Selection

The analysis was carried out in CMSSW_5_3_11_patch6, the standard monojet analysis code has been revised and modified to implement the b-tagging CSV algorithm (Dhingra, 2014). In the standard monojet analysis, events were collected using two triggers, the trigger efficiencies are measured to be nearly 100% for all signal regions (Khachatryan et al., 2015). The sensitivity study followed the same event selection depicted in Section 3.3.1. On top of that, the lepton veto was added into the event selection showed in Table 3.1 to remove processes producing leptons, such as W and Z production, di-bosons, and top-quark decays. Events with well reconstructed and isolated electrons with $p_T > 10$ GeV, reconstructed muons (CMS Collaboration, 2012b) with $p_T > 10$ GeV and well-identified (CMS Collaboration, 2012c) hadronically decaying tau leptons with $p_T > 20$ GeV and $|\eta| < 2.3$ were suppressed and rejected. Electrons and muons were considered isolated if the scalar sum of the p_T of the charged hadrons, neutral hadrons and photon contributions computed in a cone of radius $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ about the lepton direction, divided by the electron or muon p_T, is less than 0.2 (Khachatryan et al., 2015). The analysis was performed in 7 inclusive regions of E_T^{miss}:$E_T^{miss} > 250, 300, 350, 400, 450, 500, 550$ GeV. The Figure 4.2 shows the mono b-jet event selection for selecting a b-tagged monojet signal event.

Table 4.2: The summary of mono b-jet event selection.

PFMuon + $E_T^{miss} > 200$ GeV
Noise Cleaning
Jet1 $p_T > 110$ GeV
CSV b-taging (medium WP)
Jet Multiplicity < 3
$\Delta\phi(j_1, j_2) < 2.5$
Lepton veto
Tau veto
7 inclusive E_T^{miss} cuts

4.3.3 W+jets Data-Driven Background Estimation

After the signal event selection, there are two dominant backgrounds remained. Particularly the Z+jets which the Z boson decay into a pair of neutrinos, $Z(\nu\bar{\nu})$ event; and W+jets with the W boson decaying leptonically, $W(l\nu)$, where l stands for a charged lepton, and can be replaced by e, μ or τ to denote specific decays to electron, muon, or tau, respectively.

In the spirit of improving the sensitivity of the mono b-jet channel, a data-driven background estimation was carried out to estimate the two backgrounds. The $W(\mu\nu)$ control sample will be used extensively to estimate Z+jets and W+jets processes. The former background had been predicted in Chapter 3 while the latter will be covered in this chapter.

As usual, the $W(\mu\nu)$ control sample was selected by going through the event selection described in Table 3.1 and Table 3.3. Later, the $Z(\nu\bar{\nu})$ event is predicted by taking account of the kinematics difference between the control data sample and the estimated sample as described in Section 3.4.2. Similarly the number of lost muon, lost electron and tau lepton were estimated and corrected with their respective kinematics differences from the same control data sample. Finally by combining the the lost lepton event contribution to the predicted number of lost W+jets event, the total lost W+jets background is predicted.

4.3.4 Limit Derivation at 90 % CL

After taken into account of the estimated SM background events and the mono b-jet + E_T^{miss} final state signal generation, a lowest expected limit at 90 % CL (Read, 2002) was derived on the M_*. In order to compute the expected limit, the ROOSTAT routine (Schott, 2012) required the calculated acceptance from the mono b-jet + E_T^{miss} final state from the MADGRAPH simulation, and the number of estimated background events, error of the background events, and the number of the data events after the mono b-jet event selection. Since our interest is to study the sensitivity of mono b-jet channel's contribution to the DM signal, the number of data event after the event selection was made the same as the total number of MC simulated SM background events. A lowest expected limit was determined by examining the 7 inclusive E_T^{miss} of 9 DM mass points. Finally a corresponding 90 % limit was obtained on the $\sigma_{\chi-nucleon}$.

4.4 Results

4.4.1 Parton-Level Signal Event Kinematics

A total of 9 samples were generated according to the process configuration. The leading jet and sub-leading jet were defined by the highest p_T and second highest p_T after the jet candidates were sorted with p_T-ordered. Since our signal event consists of a b-jet, the distribution of the leading jet p_T, sub-leading jet p_T and the leading b-jet p_T were plotted separately to explore the correlation between them. Besides, the DM p_T distribution was plotted since it recoiled with the b-jet. The parton-level kinematics information of the generated process was summarized in figure 4.5.

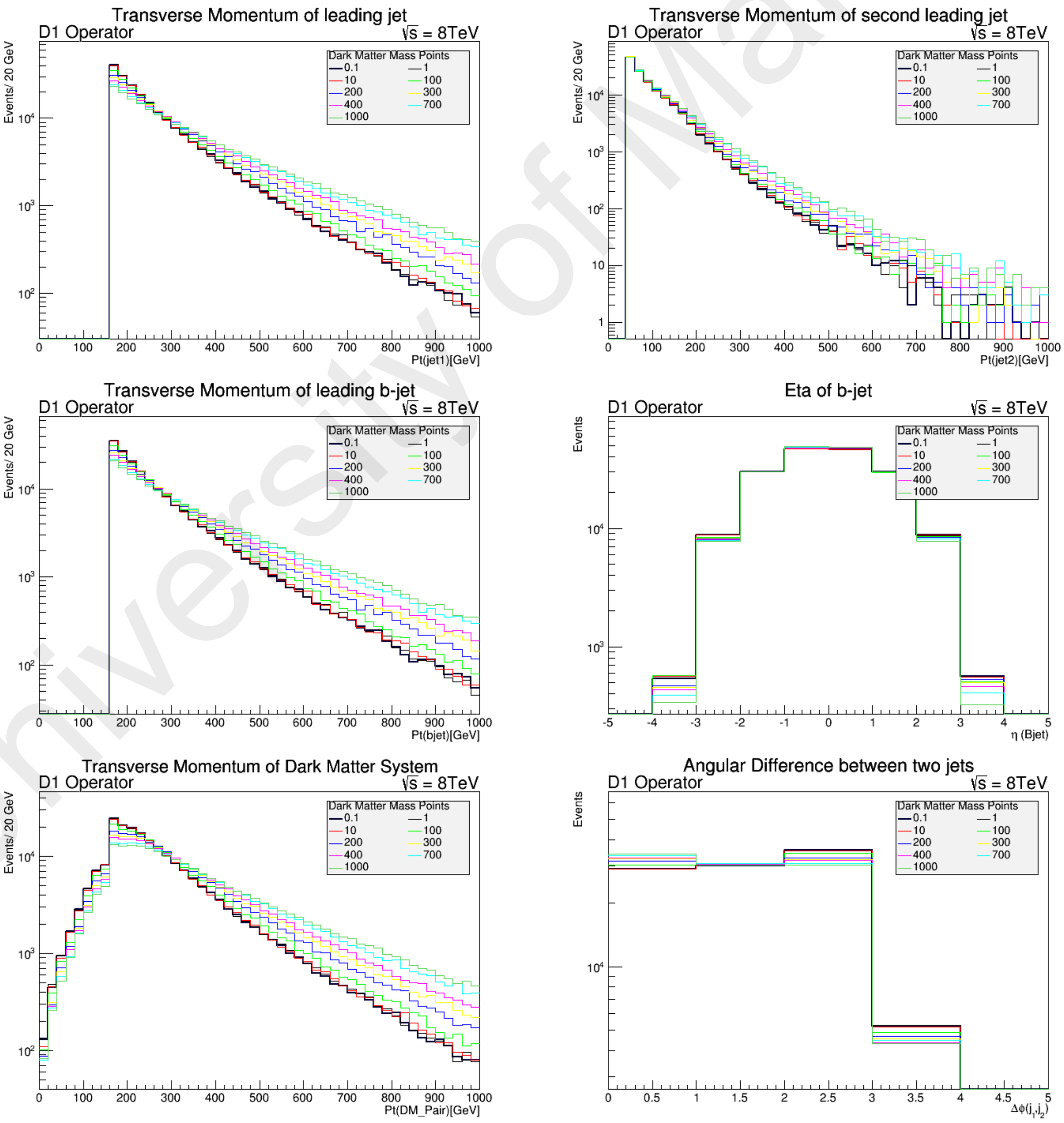

Figure 4.5: The basic kinematics for the generated signal process consisting two leading jets and a DM system.

The distribution of the generated signal event reflected the imposed generator level cuts during event generation. The p_T distribution of the DM system showed a large tail at all 9 mass points as expected. Similarly the same was observed on both p_T distribution of the first leading jet and the b-jet. The leading b-jet p_T distribution had been observed to have similar distribution as b-jet p_T, as described in the ME generation, implying most of the leading jet is b-jet. On the other hand, the sub-leading jet had less energetic momentum compared to the first leading jet, nevertheless it had appreciable high p_T and harder distribution since the two leading jets were described by ME in the generator. Since a second jet was allowed through $\chi\bar{\chi}$ + b-jet + jet, the contribution at lower p_T region in DM system was characterized by the soft sub-leading jet which had lower p_T. The mixing from events consisting of sub-leading jet together with events only have leading one b-jet significantly improves the mono b-jet $+E_T^{miss}$ final state events, therefore a distinguishing peak at 80 GeV was visible in DM system with harder spectrum.

It was clearly seen that the kinematics of the jet was strongly dependent on the mass of the pair of DM, due to the fact that the b-jet was the only particle recoiling the pair of DM. Because of the conservation of momentum, the same strong dependent on M_χ was observed on the DM pair system. However, the second allowed jet p_T distribution did not change much as the M_χ increases. Similarly, the centrally produced b-jet as indicated by its η distribution did not show a strong dependent on M_χ. Likewise the angular separation between the two leading jets was not affected by the mass of the DM pair.

Table 4.3: LO cross section computed by Madgraph5 assuming $M_* = 1000$ GeV.

Mass Point (GeV)	Cross Section (pb)
0.1	1.2706e-07
1	1.2729e-07
10	1.2633e-07
100	7.6861e-08
200	3.5074e-08
300	1.5568e-08
400	6.9567e-09
700	6.4741e-10
1000	6.2584e-11

Heavier DM mass point is difficult to produce in particle collision as it is limited by the center of mass energy. The simulated signal event's cross section was decreasing while moving from low DM to higher DM. The cross section for each generated mass point computed at LO was reported in table 4.3. The acceptance for each E_T^{miss} region of each DM mass point was computed by taking the ratio of events passing each E_T^{miss} cuts over the total number of generated events in MADGRAPH. Also a hypothesized hadronization efficiency amount to 50 % due to smearing effect during showering and 70 % of b-tagging efficiency were considered. The calculated acceptance for each E_T^{miss} cuts is shown in table 4.4. Since kinematics of the signal process is strongly dependent on the M_χ, more event will be expected to pass at higher E_T^{miss} cuts for high M_χ sample, thus the higher the acceptance.

Table 4.4: Acceptance includes 50 % of hadronization, and 70 % of b-tagging efficiency at different E_T^{miss} cut.

Mass Point (GeV)	250	300	350	400	450	500	550
0.1	2.2169	1.1622	0.6507	0.3679	0.2193	0.1342	0.0812
1	2.2698	1.2005	0.6587	0.3586	0.2086	0.1236	0.0756
10	2.2876	1.2282	0.6790	0.3875	0.2184	0.1306	0.0756
100	2.8632	1.5622	0.8880	0.5182	0.3024	0.1860	0.1166
200	3.7237	2.1166	1.2453	0.7478	0.4480	0.2734	0.1714
300	4.4527	2.6094	1.5801	0.9718	0.6036	0.3707	0.2349
400	5.0985	3.1227	1.9387	1.2189	0.7828	0.4960	0.3222
700	6.3970	4.0968	2.6653	1.7747	1.1788	0.7882	0.5315
1000	7.1269	4.7072	3.1526	2.1140	1.4319	0.9777	0.6692

4.4.2 Mono b-jet Event Selection

As a result of the mono b-jet event selection, the Table 4.5 shows the number of MC and data event survived in each cut presented in the cut-flow table. A cut flow table is useful to understand the sensitivity of the particular cuts towards each of the number of background event. The b-tagging implemented on the leading jet together with a p_T cut had drastically reduced most of the number of event to ~ 96 %. This is due to the stringent criteria prescribed by the medium working point in b-tagging (Dhingra, 2014). The jet multiplicity cut (NJet) which required a 2 jets event had been effective on reducing single t background as the inclusive decay produced two jets in its final state. The $\Delta\phi$ cut was effective to bring down ~ 94 % of QCD-multijet process. On the other hand, the lepton veto was able to minimize most of the background processes producing lepton in their final state, it was effective on W+jets, $t\bar{t}$ and single t. However after the $\Delta\phi$ cut the $Z(\nu\bar{\nu})$ background barely reduced more then ~ 3 %.

In order to evaluate the total number of SM background contributions to the signal region, a comparison between MC simulated SM background and data after the event selection was studied, as shows in figure 4.6. As expected, the most dominant backgrounds were the $Z(\nu\bar{\nu})$ events which made up of ~ 56 %, and $W(l\nu)$ events ~ 25 %, followed by di-boson and $t\bar{t} \sim 8$ %, QCD-multijets, and traces of Z boson leptonic decay events. On the kinematics of the signal region, the p_T distribution of the two leading jets were agreed with the data at lower p_T region, but higher uncertainty observed at higher p_T region. The two leading jets mostly being produced in the center region of the detector as it had a fair bell shape distribution centered at η=0, which were found to be consistent with the data. On the other hand, the signal region which is the observable, E_T^{miss} distribution agreed with the data within the 7 inclusive E_T^{miss} cut regions. Meanwhile at high E_T^{miss}, higher uncertainty was observed as it was reported in the monojet analysis at 8 TeV (Khachatryan et al., 2015). Despite of the noise cleaning criteria and the event selection requirement, backgrounds such as $W(l\nu)$, $Z(ll)$ and $t\bar{t}$ events were observed as residual at lower E_T^{miss} region.

Table 4.5: The cut flow table for standard monojet event selection with additional b-tagging

The cut flow table of standard monojet selection with b-tagging added on various SM Backgrounds.

Selection	W+Jets	Z+Jets	$Z(\nu\bar{\nu})$	DiBoson	$t\bar{t}$	Single t	QCD	MC	Data
Pre-selection	2399311	178008	1058395	3458593	461413	77284	5429269	13062273	23801498
E_T^{miss} >200 GeV	304520	28580	124954	31764	63174	9289	87605	649887	1520129
Noise cleaning	282679	26662	116256	26178	59565	8546	81890	601776	635498
Jet1>110 GeV&CSV>0.679	8429	957	4241	1172	14469	2052	5745	37063	40013
NJet<3	5779	691	3141	839	3208	686	2351	16694	19105
DeltaPhi(j1,j2)<2.5	5239	625	2928	779	2712	610	149	13042	16264
Muon veto	1977	26	2928	485	906	186	134	6641	8372
Electron veto	1426	15	2926	435	518	116	134	5571	7200
Tau veto	1366	14	2906	422	482	105	133	5429	7017
E_T^{miss} >250 GeV	473	3.3	1080	150	140	24	44	1916	2323
E_T^{miss} >300 GeV	179	2.1	441	63	50	8.1	3.7	747	887
E_T^{miss} >350 GeV	81	1.6	200	32	21	3.1	2.7	341	374
E_T^{miss} >400 GeV	39	1.3	96	15	7.7	0.3	0.4	160	196
E_T^{miss} >450 GeV	15	0.5	51	6.4	0.8	0.0	0.3	74	98
E_T^{miss} >500 GeV	7.9	0.0	28	3.6	0.0	0.0	0.2	39	53
E_T^{miss} >550 GeV	3.7	0.0	17	2.3	0.0	0.0	0.1	23	34

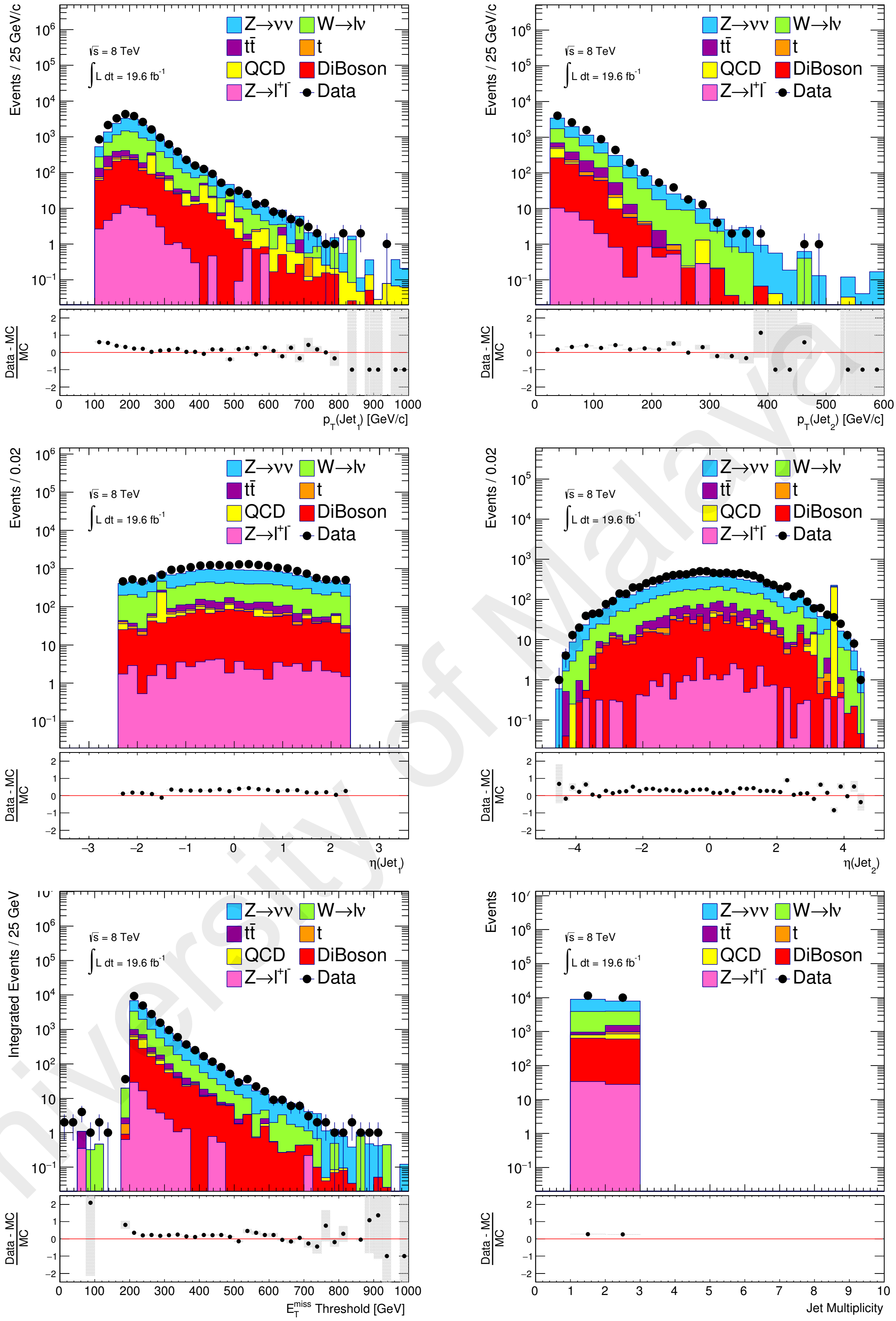

Figure 4.6: The kinematics plots after the mono-b event selection with b-tagging operating at medium working point.

4.4.3 Data-Driven Backgrounds Estimation on Z+jets and W+jets

As the study conducted in Section 4.4.2 suggested, a large amount of $Z(\nu\bar{\nu})$ event and some appreciable W+jets event had survived the event selection. Since the configuration of the triggers allowed events containing leptons, estimating backgrounds with leptonic events can be realized by selecting these lepton according to the choice of the control sample to define the control region. Therefore a data-driven background estimation is needed to model these backgrounds, and a $W(\mu\nu)$ event will be selected as a control sample to estimate both Z+jets and W+jets backgrounds.

Since the $Z(\nu\bar{\nu})$ events prediction had been carried out in Section 3.4.2, the W+jets study will be covered in this section. The W+jets estimation technique will be fully adopted from the monojet analysis (Khachatryan et al., 2015). The control sample of $W(\mu\nu)$ was selected and corrected by the similar approach described in Section 3.3, but only differ in the method estimating the total number of single muon event:

$$N^{\mu}_{tot} = \frac{N_{obs} - N_{bgd}}{Acc'\varepsilon'} \tag{4.2}$$

The required correction for background contamination of the control sample, the acceptance and efficiency are taken from MC simulation. The Acc' and ε''s definition were similar as Equation 3.8 and Equation 3.9.

After we obtained the total number of $W(\mu\nu)$ control data sample which was reported in Table 3.7, the total number of $W(\mu\nu)$ events that was out of the acceptance and not identified or isolated can be calculated by:

$$N_{lost,\mu} = N^{\mu}_{tot} \times (1 - A_{\mu}\varepsilon_{\mu}) \tag{4.3}$$

The lost muon was due to inefficiencies in the reconstruction or because they have trajectories outside the muon system acceptance. The value of A_{μ} and ε_{μ} were aslo taken from MC simulation. There were similar contributions from W decays to electrons and tau leptons. These contributions were also estimated based on the $W(\mu\nu)$ control data sample. Similarly, the total number of

$W(e\nu)$ event that was out of the acceptance and not identified or isolated and $W(\tau_{had}\nu)$ event can be written as:

$$N_{lost,e} = N^{\mu}_{tot} \times f_e \times (1 - A_e \varepsilon_e) \tag{4.4}$$

$$N_{lost,\tau} = N^{\mu}_{tot} \times f_\tau \times (1 - A_\tau \varepsilon_\tau) \tag{4.5}$$

The ratio of $W(l\nu)$ event to $W(\mu\nu)$ event passing the selection steps prior to the lepton veto was taken from simulation, separately for each lepton flavor. The term f_e and f_τ were simply the ratio of $W(\mu\nu)$ and $W(e\nu)$, $W(\tau\nu)$ event predicted at generator level MC (Khachatryan et al., 2015). Again, the acceptance and efficiency were taken from each respective MC simulation. The same procedure as that used in the muon case was then applied to obtain the background contribution to the signal region. The number of the lost muon, electron and τ backgrounds in W+jets result is shown in table 4.6 and table 4.7.

Table 4.6: Number in calculation of lost muon background in W+jets.

E_T^{miss}	> 250	> 300	> 350	> 400	> 450	> 500	> 550
$Nobs_{W(\mu\nu)}$	910	355	170	77	43	21	6.0
$Nbgd_{W(\mu\nu)}$	288	105	47	16	7.9	5.3	2.3
$A'\varepsilon'$	0.338	0.350	0.351	0.348	0.370	0.370	0.369
N^{μ}_{tot}	1838	716	351	177	95	42	10
$A_\mu\varepsilon_\mu$	0.903	0.923	0.932	0.926	0.942	0.954	0.985
$N_{lost\mu}$	179	55	24	13	5.5	2.0	0.1

Table 4.7: Number in calculation of lost electron and τ background in W+jets.

E_T^{miss}	> 250	> 300	> 350	> 400	> 450	> 500	> 550
$A_e\varepsilon_e$	0.548	0.590	0.618	0.681	0.714	0.780	0.767
f_e	0.277	0.263	0.280	0.299	0.289	0.265	0.238
$N_{lost,e}$	230	77	38	17	7.8	2.5	0.6
$A_\tau\varepsilon_\tau$	0.063	0.054	0.057	0.000	0.000	0	0
f_τ	0.122	0.113	0.119	0.102	0.070	0.088	0.058
$N_{lost,\tau}$	211	77	39	18	6.6	0	0

All of the above can then be summarized in a master equation for estimating the lost W background which is the sum of Equation 4.3, Equation 4.4 and Equation 4.5:

$$N_{lost} = \frac{N_{obs} - N_{bgd}}{A'\varepsilon'} \times [(1 - A_{\mu}\varepsilon_{\mu} + f_e \times (1 - A_e\varepsilon_e) + f_{\tau} \times (1 - A_{\tau}\varepsilon_{\tau}))] \tag{4.6}$$

The uncertainty on the W+jets background estimation was due to errors in the observed number of events; errors in the number of background events with 50% uncertainty assigned to background that was estimated using MC added in quadrature; and errors in the acceptance and efficiencies. A 2% error in the PDF was absorbed into the acceptance uncertainty A', and a 2% error due to hadronization was absorbed into the efficiency uncertainty ε'. After appropriately accounting for the uncertainty, the total number of W+jets events where a lepton has been lost was computed by Equation 4.6, and the total uncertainty is just $N_{lost} \times Error(W_{lost})$. The estimated lost W+jets background was summarized in table 4.8. Finally a summary of the predictions and corresponding uncertainties for all the SM backgrounds after the mono b-jet event selection is reported in table 4.9.

Table 4.8: Estimated number of lost W+jets events in each E_T^{miss} regions.

E_T^{miss}	> 250	> 300	> 350	> 400	> 450	> 500	> 550
$N_{lost\mu}$	179	55	24	13	5.5	2.0	0.1
$N_{lost,e}$	230	77	38	17	7.8	2.5	0.6
$N_{lost,\tau}$	0.122	0.113	0.119	0.102	0.070	0.088	0.058
Total Error	18	17	18	20	26	0	0
$W_{lost} + jets$	620±113	209±37	101±18	48±9.8	20±5.1	0.0±0.0	0.0±0.0

Table 4.9: Table of the total SM backgrounds prediction for the numbers of events passing the selection requirements, for various E_T^{miss} thresholds. The uncertainties include both statistical and systematic components.

E_T^{miss}	>250	>300	>350	>400	>450	>500	>550
$Z(\nu\bar{\nu})$	1542±278	567±96	293±51	113±22	66±15	26±9.0	5.2±3.7
W+jets	620±113	209±37	101±18	48±9.8	20±5.1	0.0±0.0	0.0±0.0
$t\bar{t}$	140±70	50±25	21±10	7.7±3.9	0.8±0.4	0.0±0.0	0.0±0.0
Z+jets	3.3±1.6	2.1±1.1	1.6±0.8	1.3±0.6	0.5±0.3	0.0±0.0	0.0±0.0
Single t	24±12	8.1±4.0	3.1±1.5	0.3±0.2	0.0±0.0	0.0±0.0	0.0±0.0
QCD	44±22	3.7±1.9	2.7±1.3	0.4±0.2	0.3±0.2	0.2±0.1	0.1±0.0
DiBoson	150±75	63±31	32±16	15±7.7	6.4±3.2	3.6±1.8	2.3±1.2
Total SM	2524±318	901±110	453±58	186±26	94±16	0.0±0.0	0.0±0.0

4.4.4 90% CL Limit on Dark Matter

The good agreement between the signal sample and the SM prediction is used to set limits on the dark matter production cross sections. The limits on dark matter production are set as a function of the dark matter candidate mass, M_χ, for spin independent scenario. These limits are converted into lower limits on the cut off scale M_*, which are used to also derive upper limits on the $\sigma_{\chi-nucleon}$.

A modified RooStatsCl95 package (Read, 2002) was used, together with 30% uncertainty assigned for the signal, the 90 % expected limit was derived on DM quantity. Since the data sample was blinded in the mono b-jet event selection, the number of data survived after the 7 E_T^{miss} cuts were made equal to the number of survived SM background in order to assess the sensitivity of the channel.

The expected limit was optimized by running the ROOSTAT on the same DM mass point with different E_T^{miss} cuts. The values obtained in Table 4.4 and Table 4.9 were used as an input to the routine. The lowest expected limit was observed on $E_T^{miss} > 400$ GeV on each DM mass points. Thus the 90 % expected lower limit was derived on M_* as a function of M_χ. Also, the 90% CL expected upper limit on $\sigma_{\chi-nucleon}$ as a function of M_χ was obtained by using the Equation 4.7 assuming a spin independent scenario. The respective limits are reported in Figure 4.7 .

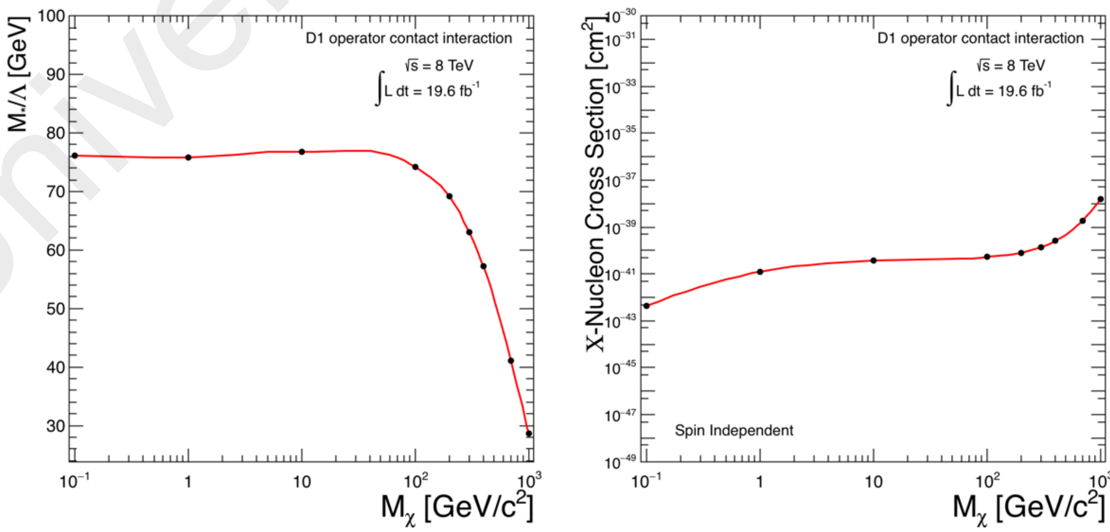

Figure 4.7: (left) The 90% CL lower limit on M_* and (right) the 90 % CL upper limit on $\sigma_{\chi-nucleon}$ as a function of M_χ assuming a D1 scalar operator in mono-b + E_T^{miss} channel.

4.5 Discussion

The 90% CL was chosen to enable a direct comparison with the results from the direct detection experiments. By interpreting the mono b-jet channel, the lower expected limit on M_* has showed an improvement of a factor of $\sim$ 6 GeV compares to inclusive monojet channel at 8 TeV which is $\sim$ 70 GeV (Lin et al., 2013) as showed in Figure 4.1. A plateau is observed in the range of low M_χ region up to 100 GeV, with M_* limited at 76 GeV. As the mass of DM moving to higher M_χ region, the contact interaction is diminished resulting into a slop observed in higher DM region which is consistently with the idea that higher mass of DM is hard to produce since it requires more energy to create.

The collider limit on the strength of interactions between DM and hadrons can be translated into the constraints on the possible contributions to direct detection cross sections for each of those interactions or operators. Those possible effective operators are listed in figure 2.9 provided they are operated in the limit of low momentum transfer (Goodman et al., 2010). For D1 operator, by combining the kinematics of WIMP-nucleon scattering, the cross section can be computed by the formula (Goodman et al., 2010):

$$\sigma_o^{D1} = 1.60 \times 10^{-37} cm^2 (\frac{\mu_\chi}{1GeV})^2 (\frac{20GeV}{M_*})^6 \tag{4.7}$$

where μ_χ is the reduced mass of the WIMP-nucleon system and the M_* is the probed interaction scale. Note that the behavior at low WIMP masses is affected strongly by the spin of the WIMP itself. By comparing to the result from inclusive monojet at 8 TeV together with direct detection experiments (Khachatryan et al., 2015) which is showed in Figure 4.8, the derived limit in this study has persistently showed that the collider limit is capable to probe into light DM region effectively compares to the DM direct detection and consistently excludes most of the low M_χ. This is attributing to the fact that the direct detection suffers from small momentum threshold in the interaction between the WIMP and nucleon scattering, and it is further severed by the extremely small scattering rate, disallowing the experiment to probe into extremely light DM.

On the contrary, the collider limit excludes poorly at higher M_χ due to the energy threshold is limited at probing into higher masses of DM. The direct detection limits accompanied with the collider limits in figure 4.8 are taken from CoGeNT (Aalseth et al., 2011), SIMPLE (Felizardo et al., 2012), COUPP (Behnke et al., 2012), CDMS-II (Z. Ahmed, 2010), XENON100 (Aprile et al., 2011) and LUX (Akerib et al., 2014) collaboration.

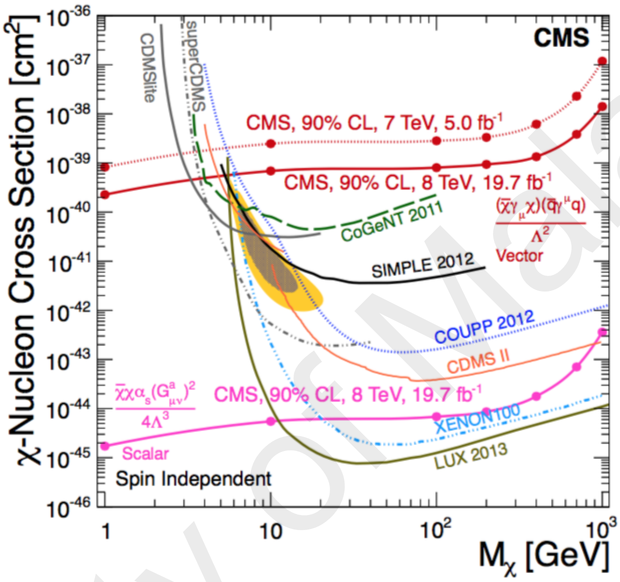

Figure 4.8: Upper limits for vector and scalar operators interpreted in inclusive monojet channel on the $\sigma_{\chi-nucleon}$, at 90% CL, plotted against DM particle mass and compared with previously published results.

Based on the result of the study, the collider-derived limit on $\sigma_{\chi-nucleon}$ especially interpreted form mono b-jet channel with scalar coupling has excluded more phase space compares to monojet channel assuming a vector (D5) and axial-vector (D6) coupling at 8 TeV due to the effective SM background reduction by the b-tagging. At higher M_χ the mono b-jet channel yields a looser limit compares to monojet channel, this is due to the higher error observed *sim* 13 % on number of SM backgrounds event prediction at higher E_T^{miss} showed in Figure 4.9.

However, the inclusive monojet channel with gluon-induced coupling (D11) has outperformed both the current result with D1 coupling from the study and the D5 coupling. That is because the D11 operator taking into account of the loop process in the monojet and thus expecting to yield a higher limits.

Given the high center of mass energies that are being probed by the LHC, it is important to consider the possibility that the effective theory is not always valid. The validity of the effective theory has been discussed in (Busoni et al., 2014). It is pointed out in the literature that for theories to be perturbative the product of the couplings $g_\chi g_q$ is typically required to be smaller than 4π, and this condition is likely not satisfied for the entire region of phase space probed by the collider searches. For future study on the DM searches in EFT, the evaluation on the validity of the EFT in a given kinematical condition is encouraged in order to complement the EFT results.

4.6 Conclusion

The sensitivity study on the mono b-jet + E_T^{miss} final state is investigated on the CMS data collected at 19.7 fb^{-1} in 8 TeV over a range of DM mass point. In the EFT interpretation, the mono-b + E_T^{miss} channel assuming a scalar coupling has excluded parameter M_* at 76 GeV given M_χ is 100 GeV at 90% CL. Similarly, the collider-derived limit at 90% CL has consistently excluded the DM low mass region in $\sigma_{\chi-nucleon}$ parameter space compares to DM direct detection experiment, and outperformed collider limits derived from inclusive monojet channel with vector coupling in spin independent scenario.

At higher M_χ, the mono-b + E_T^{miss} channel excludes poorly compares to inclusive monojet derived collider limit due to the limitation of energy delivered by LHC and high uncertainty from background estimation. The sensitivity study highlights the plausibility of interpreting mono b-jet + E_T^{miss} final state as a search channel in DM search at LHC. Although the mono b-jet channel with scalar operator does not yield a stringent limit compares to other search channel, the channel is highly motivated as it might be contributing to the anomalies observed on the measurements of gamma rays that originate from the galactic centre (Calore, Cholis, McCabe, & Weniger, 2015).

Also, in order for the EFT to hold especially in a collider experiment, the EFT validity study is encouraged for DM search interpreted in mono b-jet channel.

CHAPTER 5: CONCLUSION

To summarize, both study is served as an extension to the monojet channel analysis (Khachatryan et al., 2015) in the spirit to study the sensitivity of mono b-jet channel. The analysis aspect such as the background modeling and mono b-jet + E_T^{miss} final state MC simulation study are explored by attempting different alternative methods in order to yield a higher limit to constrain on DM process. Although the centrally produced SM background samples were monojet-like, interpreting mono b-jet channel on those samples will not be any major change since the process topology is similar to monojet.

The data-driven background estimation constitutes an important sub-study in most of the high energy physics analysis especially searching for new physics process. The $W(\mu\nu)$'s derived estimation on $Z(\nu\bar{\nu})$ background event is well motivated especially for an analysis involving b-tagging which usually reduce significant amount of event during event selection. The study has examined the feasibility of $W(\mu\nu)$ control sample as its offered sizable control sample compares to $Z(\mu\bar{\mu})$ control sample in an effort to constrain $Z(\nu\bar{\nu})$ background events. So far the estimation has successfully predicting the backgrounds with tolerable uncertainties, but the predictive power is still limited by appreciable uncertainty incurred from various correction factors. Therefore one can consider an optimized approach to modeling the $Z(\nu\bar{\nu})$'s E_T^{miss} spectrum by using both Control samples.

Under the assumption of EFT, the magnitude of DM coupling to SM particles is proportional to the quark mass, as described by Equation 4.1. Therefore the heavy flavour quark has become the important channel to be searched in collider experiment. With the assumption of 50 % efficiency from Hadronization and 70 % b-tagging, the mono b-jet + E_T^{miss} signal generated with MLM-matching scheme at the generator level is served as an exploratory step to assess the potential improvement on the signal acceptance by minimizing the possibility of double counting. As usual the second jet is allowed to retain mono-b events with ISR/FSR. In the analysis level, the sensitivity of mono b-jet channel can be assessed by including a b-tagging cuts in the monojet event selection (Khachatryan et al., 2015). With an augmented monojet event selection, the b-tagging will impose a tighter cut on the passing events thus the analysis is prone to higher statistical uncertainty. A $W(\mu\nu)$ control sample is then used to estimate the two dominant backgrounds to the mono b-jet channel in order to constrain the uncertainty.

With the similar POG recommended prescription on the physics objects in Monojet analysis (Khachatryan et al., 2015), the mono b-jet channel yields an exclusion on M_* at 76 GeV with 90% CL given the DM mass is 100 GeV. The result is consistent with the prediction reported in figure 2.14 (Lin et al., 2013). Also, like other search channel in collider experiment, the mono b-jet channel consistently probe into the low DM mass region and capable to exclude the $\sigma_{\chi-nucleon}$ phase space. On the other hand, it has also show that at higher M_χ the exclusion power is limited by the energy reach of the LHC.

The sensitivity study has concluded for the first time, that an empirical search for direct b quark production in association with a pair of DM can be a promising channel in excluding DM phase space. Therefore a dedicated analysis on mono b-jet channel is highly motivated in the spirit to improve the discovery potential of DM at LHC, given the utmost priority is focusing on improving uncertainties arises from b-tagging due to fake b-jet rate (Dhingra, 2014), correction factors such as acceptance and efficiency in background estimation, jet-matching efficiency and so on.

REFERENCES

Aalseth, C. E., Barbeau, P. S., Bowden, N. S., Cabrera-Palmer, B., Colaresi, J., Collar, J. I., ... Yocum, K. M. (2011). Results from a search for light-mass dark matter with a *p*-type point contact germanium detector. *Phys. Rev. Lett.*, *106*, 131301. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.106.131301

Abazov, V. M., et al. (2009). Observation of Single Top Quark Production. *Phys. Rev. Lett.*, *103*, 092001.

Abdallah, J., Araujo, H., Arbey, A., Ashkenazi, A., Belyaev, A., Berger, J., ... Zurek, K. (2015). Simplified models for dark matter searches at the {LHC}. *Physics of the Dark Universe*, *9–10*, 8 - 23. Retrieved from http://www.sciencedirect.com/science/article/pii/S2212686415000163

Akerib, D. S., Araújo, H. M., Bai, X., Bailey, A. J., Balajthy, J., Bedikian, S., ... Zhang, C. (2014). First results from the lux dark matter experiment at the sanford underground research facility. *Phys. Rev. Lett.*, *112*, 091303. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.112.091303

Alioli, S., Nason, P., Oleari, C., & Re, E. (2010). A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *Journal of High Energy Physics*, *2010*(6), 1–58. Retrieved from http://dx.doi.org/10.1007/JHEP06(2010)043

Alwall, J., Herquet, M., Maltoni, F., Mattelaer, O., & Stelzer, T. (2011). Madgraph 5: going beyond. *Journal of High Energy Physics*, *2011*(6), 1–40. Retrieved from http://dx.doi.org/10.1007/JHEP06(2011)128

Aprile, E., Alfonsi, M., Arisaka, K., Arneodo, F., Balan, C., Baudis, L., ... Weinheimer, C. (2012). Dark Matter Results from 225 Live Days of XENON100 Data. *Physical Review Letters*, *109*(18), 181301.

Aprile, E., Arisaka, K., Arneodo, F., Askin, A., Baudis, L., Behrens, A., ... Weinheimer, C. (2011). Dark matter results from 100 live days of xenon100 data. *Phys. Rev. Lett.*, *107*, 131302. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.107.131302

Aprile, E., & XENON1T collaboration. (2012). The XENON1T Dark Matter Search Experiment. *ArXiv e-prints*.

Askew, A., Chauhan, S., Penning, B., Shepherd, W., & Tripathi, M. (2014). Searching for Dark Matter at Hadron Colliders. *Int.J.Mod.Phys.*, *A29*, 1430041.

Bai, Y., & Tait, T. M. P. (2013). Searches with Mono-Leptons. *Phys. Lett.*, *B723*, 384-387.

Behnke, E., Behnke, J., Brice, S. J., Broemmelsiek, D., Collar, J. I., Conner, A., ... Vázquez-Jáuregui, E. (2012). First dark matter search results from a 4-kg cf$_3$**I** bubble chamber operated in a deep underground site. *Phys. Rev. D*, *86*, 052001.

Retrieved from http://link.aps.org/doi/10.1103/PhysRevD.86.052001

Beltran, M., Hooper, D., Kolb, E. W., Krusberg, Z. A., & Tait, T. M. (2010). Maverick dark matter at colliders. *JHEP*, *1009*, 037.

Brun, R., & Rademakers, F. (1997). {ROOT} — an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *389*(1–2), 81 - 86. Retrieved from http://www.sciencedirect.com/science/article/pii/S016890029700048X (New Computing Techniques in Physics Research V)

Buckley, A., Butterworth, J., Gieseke, S., Grellscheid, D., Höche, S., Hoeth, H., ... Webber, B. (2011). General-purpose event generators for {LHC} physics. *Physics Reports*, *504*(5), 145 - 233. Retrieved from http://www.sciencedirect.com/science/article/pii/S0370157311000846

Busoni, G., De Simone, A., Morgante, E., & Riotto, A. (2014). On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC. *Phys.Lett.*, *B728*, 412-421.

Cabibbo, N. (1963). Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, *10*, 531–533. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.10.531

Cacciari, M., Salam, G. P., & Soyez, G. (2008a). The anti- k t jet clustering algorithm. *Journal of High Energy Physics*, *2008*(04), 063. Retrieved from http://stacks.iop.org/1126-6708/2008/i=04/a=063

Cacciari, M., Salam, G. P., & Soyez, G. (2008b). The Anti-k(t) jet clustering algorithm. *JHEP*, *0804*, 063.

Calore, F., Cholis, I., McCabe, C., & Weniger, C. (2015). A tale of tails: Dark matter interpretations of the fermi gev excess in light of background model systematics. *Phys. Rev. D*, *91*, 063003. Retrieved from http://link.aps.org/doi/10.1103/PhysRevD.91.063003

CERN. (2014). *LHC.* Retrieved from http://home.web.cern.ch/topics/large-hadron-collider

Chatrchyan, S., Khachatryan, V., Sirunyan, A. M., Tumasyan, A., Adam, W., Bergauer, T., ... Swanson, J. (2014). Measurement of inclusive w and z boson production cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, *112*, 191802. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.112.191802

Chatrchyan, S., et al. (2008). The CMS experiment at the CERN LHC. *JINST*, *3*, S08004.

Clowe, D., Bradac, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., et al. (2006). A direct empirical proof of the existence of dark matter. *Astrophys.J.*, *648*, L109-L113.

CMS Collaboration. (2008). The cms experiment at the cern lhc. *Journal of Instrumentation*, *3*(08), S08004. Retrieved from http://stacks.iop.org/1748-0221/3/i=08/a=S08004

CMS Collaboration. (2009a). Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search.

CMS Collaboration. (2009b). *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET* (Tech. Rep. No. CMS-PAS-PFT-09-001). 2009. Geneva: CERN. Retrieved from https://cds.cern.ch/record/1194487

CMS Collaboration. (2010). Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector.

CMS Collaboration. (2012a). Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7$ tev. *Journal of Instrumentation*, *7*(10), P10002. Retrieved from http://stacks.iop.org/1748-0221/7/i=10/a=P10002

CMS Collaboration. (2012b). Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, *7*, 2P.

CMS Collaboration. (2012c). Performance of τ-lepton reconstruction and identification in CMS. *Journal of Instrumentation*, *7*, 1001.

Corcella, G., Knowles, I. G., Marchesini, G., Moretti, S., Odagiri, K., Richardson, P., ... Webber, B. R. (2001). Herwig 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *Journal of High Energy Physics*, *2001*(01), 010. Retrieved from http://stacks.iop.org/1126-6708/2001/i=01/a=010

D'Ambrosio, G., Giudice, G. F., Isidori, G., & Strumia, A. (2002). Minimal flavor violation: An Effective field theory approach. *Nucl. Phys.*, *B645*, 155-187.

Dhingra, N. (2014). Performance of b-tagging algorithms at the CMS experiment with pp collision data at $\sqrt{s}$=8 TeV. In *Proceedings, 2nd Conference on Large Hadron Collider Physics Conference (LHCP 2014).* Retrieved from http://inspirehep.net/record/1313117/files/arXiv:1409.0251.pdf

Farrar, G. R., & Fayet, P. (1978). Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Physics Letters B*, *76*(5), 575 - 579. Retrieved from http://www.sciencedirect.com/science/article/pii/0370269378908584

Felizardo, M., Girard, T. A., Morlat, T., Fernandes, A. C., Ramos, A. R., Marques, J. G., ... Marques, R. (2012). Final analysis and results of the phase ii simple dark matter search. *Phys. Rev. Lett.*, *108*, 201302. Retrieved from http://link.aps.org/doi/10.1103/PhysRevLett.108.201302

Feng, J. L. (2010). Dark Matter Candidates from Particle Physics and Methods of Detection. *Annual Review of Astronomy and Astrophysics*, *48*, 495-545.

Gleisberg, T., Höche, S., Krauss, F., Schälicke, A., Schumann, S., & Winter, J.-C. (2004). Sherpa 1.alpha, a proof-of-concept version. *Journal of High Energy Physics*, *2004*(02), 056. Retrieved from http://stacks.iop.org/1126-6708/2004/i=02/a=056

Goodman, J., Ibe, M., Rajaraman, A., Shepherd, W., Tait, T. M. P., & Yu, H.-B. (2010). Constraints on Dark Matter from Colliders. *Phys. Rev.*, *D82*, 116010.

Khachatryan, V., et al. (2015). Search for dark matter, extra dimensions, and unparticles in monojet events in proton–proton collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, *C75*(5), 235.

Kobayashi, M., & Maskawa, T. (1973). CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, *49*, 652-657.

Lin, T., Kolb, E. W., & Wang, L.-T. (2013). Probing dark matter couplings to top and bottom quarks at the LHC. *Phys. Rev.*, *D88*(6), 063510.

Malik, S. A., & Watt, G. (2014). Ratios of W and Z cross sections at large boson p_T as a constraint on PDFs and background to new physics. *JHEP*, *02*, 025.

Maltoni, F., & Stelzer, T. (2003). Madevent: automatic event generation with madgraph. *Journal of High Energy Physics*, *2003*(02), 027. Retrieved from `http://stacks.iop.org/1126-6708/2003/i=02/a=027`

Mangano, M. L., Moretti, M., Piccinini, F., & Treccani, M. (2007). Matching matrix elements and shower evolution for top-pair production in hadronic collisions. *Journal of High Energy Physics*, *2007*(01), 013. Retrieved from `http://stacks.iop.org/1126-6708/2007/i=01/a=013`

Mangano, M. L., Piccinini, F., Polosa, A. D., Moretti, M., & Pittau, R. (2003). Alpgen, a generator for hard multiparton processes in hadronic collisions. *Journal of High Energy Physics*, *2003*(07), 001. Retrieved from `http://stacks.iop.org/1126-6708/2003/i=07/a=001`

Nakamura, K., & Group, P. D. (2010). Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, *37*(7A), 075021. Retrieved from `http://stacks.iop.org/0954-3899/37/i=7A/a=075021`

Olive, K. A., et al. (2014). Review of Particle Physics. *Chin. Phys.*, *C38*, 090001.

Primack, J. R. (2009). Dark Matter and Galaxy Formation. In F. Roig, D. Lopes, R. de La Reza, & V. Ortega (Eds.), *American institute of physics conference series* (Vol. 1192, p. 101-137).

Read, A. L. (2002). Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, *28*(10), 2693. Retrieved from `http://stacks.iop.org/0954-3899/28/i=10/a=313`

Rubin, V. C., Ford, W. K. J., & . Thonnard, N. (1980). Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys.J.*, *238*, 471-487.

Sanders, R. H., & McGaugh, S. S. (2002). Modified Newtonian dynamics as an alternative to dark matter. *Ann.Rev.Astron.Astrophys.*, *40*, 263-317.

Schott, G. (2012). RooStats for Searches.

Sjöstrand, T., Mrenna, S., & Skands, P. (2006). Pythia 6.4 physics and manual. *Journal of High Energy Physics*, *2006*(05), 026. Retrieved from http://stacks.iop.org/1126-6708/2006/i=05/a=026

Smith, S. (1936). The Mass of the Virgo Cluster. *Astrophys.J.*, *83*, 23.

TE-EPC-LPC. (2012). *TE-EPC-LPC in LHC te-epc-lpc in lhc.* Retrieved 2015-07-03, from http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm

Z. Ahmed, e. a. (2010). Dark matter search results from the cdms ii experiment. *Science*, *327*(5973), 1619–1621. Retrieved from http://science.sciencemag.org/content/327/5973/1619

Zwicky, F. (2009). Republication of: The redshift of extragalactic nebulae. *General Relativity and Gravitation*, *41*(1), 207-224. Retrieved from http://dx.doi.org/10.1007/s10714-008-0707-4

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Abdallah, J., Araujo, H., Arbey, A., Ashkenazi, A., Belyaev, A., Berger, J., ... Zurek, K. (2015). Simplified models for dark matter searches at the {LHC}. *Physics of the Dark Universe*, *9–10*, 8 - 23. Retrieved from http://www.sciencedirect.com/science/article/pii/S2212686415000163 doi: http://dx.doi.org/10.1016/j.dark.2015.08.001

Abercrombie, D., et al. (2015). Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum.

Hoh, S. Y., Komaragiri, J. R., & Wan Abdullah, W. A. T. (2016). Dark matter searches at the large hadron collider. *AIP Conference Proceedings*, *1704*. Retrieved from http://scitation.aip.org/content/aip/proceeding/aipcp/10.1063/1.4940063 doi: http://dx.doi.org/10.1063/1.4940063